



Research article

Exploring prognostic microbiota markers in patients with endometrial carcinoma: Intratumoral insights

Jie Liu ^{a,1}, Yi Qu ^{b,1}, Yang-Yang Li ^c, Ya-Lan Xu ^d, Yi-Fang Yan ^{b,**}, Hao Qin ^{e,*}^a Department of Medical Records, Air Force Medical Center, PLA, Air Force Medical University, Beijing, China^b State Key Laboratory of Female Fertility Promotion, Center for Reproductive Medicine, Department of Obstetrics and Gynecology, Peking University Third Hospital, Beijing, China^c Medical Center for Human Reproduction, Beijing Chao-Yang Hospital, Capital Medical University, Beijing, China^d Department of Obstetrics and Gynecology, Peking Union Medical College Hospital, Beijing, China^e State Key Laboratory of Molecular Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

ARTICLE INFO

Keywords:

Endometrial carcinoma
TCGA-UCEC
Risk score
Microbial biomarker
Prognostic signature

ABSTRACT

Endometrial cancer, a leading gynecological malignancy, is profoundly influenced by the uterine microbiota, a key factor in disease prognosis and treatment. Our study underscores the distinct microbial compositions in endometrial cancer compared to adjacent non-cancerous tissues, revealing a dominant presence of *p_Actinobacteria* in cancerous tissues as opposed to *p_Firmicutes* in surrounding areas. Through comprehensive analysis, we identified 485 unique microorganisms in cancer tissues, 26 of which correlate with patient prognosis. Employing univariate Cox regression and LASSO regression analyses, we devised a microbial risk scoring model, effectively stratifying patients into high and low-risk categories, thereby providing predictive insights into their overall survival. We further developed a nomogram that incorporates the microbial risk score along with age, grade, and clinical stage, significantly enhancing the accuracy of our clinical prediction model for endometrial cancer. Moreover, our study delves into the differential immune landscapes of high-risk and low-risk patients. The low-risk group displayed a higher prevalence of activated B cells and increased T cell co-stimulation, indicative of a robust immune response. Conversely, high-risk patients showed elevated tumor immune dysfunction and exclusion scores, suggesting less favorable outcomes in immunotherapy. Notably, the efficacy of IPS-CTLA4 and PD1/PD-L1/PD-L2 blockers was substantially higher in the low-risk group, pointing to a more responsive immunotherapeutic approach. In summary, our research elucidates the unique microbial patterns in endometrial cancer and adjacent tissues, and establishes both a microbial risk score model and a clinical prediction nomogram. These findings highlight the potential of uterine microbiota as a biomarker for customizing treatment strategies, enabling precise interventions for high-risk patients while preventing overtreatment in low-risk cases. This study emphasizes the microbiota's role in tailoring immunotherapy, offering a novel perspective in the treatment and

* Corresponding author. State Key Laboratory of Molecular Oncology, Cancer Institute Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, No.17 Panjiayuan Nanli, Chaoyang District, Beijing, 100021, China.

** Corresponding author. State Key Laboratory of Female Fertility Promotion, Center for Reproductive Medicine, Department of Obstetrics and Gynecology, Peking University Third Hospital, No.49 North Huayuan Road, Haidian District, Beijing, 100191, China.

E-mail addresses: jieliusunny@163.com (J. Liu), pku_quyi@163.com (Y. Qu), lydialyy@pku.edu.cn (Y.-Y. Li), xuyalan0391@163.com (Y.-L. Xu), yanyifangyanyf@hotmail.com (Y.-F. Yan), haoqin@cicams.ac.cn (H. Qin).

¹ These authors contributed equally to this work.

<https://doi.org/10.1016/j.heliyon.2024.e27879>

Received 7 October 2023; Received in revised form 3 March 2024; Accepted 7 March 2024

Available online 13 March 2024

2405-8440/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

prognosis of endometrial cancer. Significantly, our study's expansive sample analysis from the TCGA-UCEC cohort, employing linear discriminant analysis effect size methodology, not only validates but also enhances our understanding of the microbiota's role in endometrial cancer, paving the way for novel diagnostic and therapeutic approaches in its management.

1. Introduction

Endometrial carcinoma (EC) is a form of cancer affecting the inner epithelial lining of the uterus and stands as the second most prevalent gynecological cancer among women. In 2020, it was estimated that there were 417,367 new cases of EC, resulting in 97,370 deaths [1]. EC primarily manifests as two main types: Type I and Type II. Type I EC is linked to environments with elevated estrogen levels and is often associated with high hormone receptor levels and obesity. Type I EC accounts for 80% of all EC cases. In contrast, Type II EC is more common in older, non-obese women, carries a poor prognosis, and constitutes 10–20% of all EC cases. Notably, Type II EC shows no significant correlation with estrogen stimulation [2,3]. Remarkably, bacterial cells make up approximately 1–3% of a human's body weight; this is comparable in quantity to human cells [4]. Recent studies have identified the presence of bacteria in tumor tissues, challenging the previous notion of tumors being sterile. In 2020, Knight et al. reported the existence of microbial RNA and DNA in 33 types of cancer, with unique microbial communities associated with each type of cancer [5]. These microorganisms within tumors are now recognized as crucial contributors to tumor development, drug resistance, and the tumor microenvironment [6, 7].

Presently, it is acknowledged that bacteria can infiltrate tumor cells through various pathways, including the digestive tract, lymphatic system, and bloodstream. The composition, attributes, and diversity of microorganisms within tumors are closely intertwined with the gut microbiota [8,9]. Numerous studies have linked microbial imbalances to various diseases, including breast cancer, EC, ovarian cancer, and inflammatory bowel disease [10–13]. Several female health conditions are associated with disruptions in the microbiota of the reproductive tract, which plays a vital role in regulating the barrier function of the reproductive tract. This understanding demonstrates that the uterus is not a sterile cavity [14]. To date, multiple studies have explored the microbial composition of EC. For instance, Walther-Antônio et al. identified a close association between *Atopobium vaginae* and *Porphyromonas* sp. With EC, particularly under high pH conditions [15]. In another study, Lu et al. found that *Micrococcus* was correlated with inflammation in EC patients, who also exhibited reduced microbiota diversity compared to those with benign uterine lesions [16]. However, Walsh et al. reported contrasting results, indicating that the main known risk factor for EC was an increased microbiome diversity in the reproductive tract [17]. Disruptions in the microbiota of the female reproductive tract are closely linked to the onset and progression of gynecological cancers. Thus, targeted interventions to modulate the microbiota hold promise for enhancing primary and secondary prevention strategies for gynecological cancers.

A distinctive aspect of the current study in relation to previous research is the extensive scale of our sample analysis. We utilized microbiota data from the TCGA-UCEC cohort to investigate potential correlations between the abundance of microbiota and patient outcomes in EC. Our analysis encompassed a significantly larger sample size compared to earlier studies, thus yielding a comprehensive dataset that provides unique insights and strengthens the validity of our conclusions regarding the role of the microbiota in EC. By employing linear discriminant analysis effect size (LEfSe) methodology, we identified distinct variations in the microbiota between EC and adjacent cancer tissues. Our methodology encompassed both univariate Cox regression and LASSO regression analyses to assess the prognostic relevance of diverse microbiota biomarkers. In addition, we developed a risk factor calculation model based on our findings. Furthermore, our research investigated the influence of the microbiome on the response to immune therapy and immune checkpoint expression in tumor cells. This study aimed to unravel the intricate interactions between tissue-resident microbiota and tumors, potentially laying the groundwork for novel diagnostic and prognostic approaches in the management of EC.

2. Materials and methods

2.1. Prognostic microorganism screening and identification

In this study, we screened prognostic microorganisms identified in EC patients and data obtained from the Cancer Genome Atlas Uterine Corpus Endometrial Carcinoma (TCGA-UCEC) database, which includes microbiome abundance data (http://ftp.microbio.me/pub/cancer_microbiome_analysis/TCGA/Kraken/Kraken-TCGA-Voom-SNM-Plate-Center-Filtering-Data.csv), microbiome clinical data (http://ftp.microbio.me/pub/cancer_microbiome_analysis/TCGA/Kraken/Metadata-TCGA-Kraken-17625-Samples.csv), and gene expression data (https://gdc-hub.s3.us-east-1.amazonaws.com/download/TCGA-UCEC.htseq_fpkms.tsv.gz). The survival data and clinical data from the TCGA-UCEC were obtained from the paper published by Liu et al. [18]. The microsatellite instability (MSI) score was downloaded from the cBioPortal website (https://www.cbioportal.org/study/clinicalData?id=ucec_tcg_pan_can_atlas_2018). Information relating to the tumor mutation burden (TMB) and homologous recombination deficiency (HRD) score in the TCGA was collated from PubMed Identifier (PMID) 29617664. The Ensembl database (<http://www.ensembl.org/info/data/ftp/index.html>) was used to acquire human gtf files (Homo_sapiens.GRCh38.99.gtf.gz). In addition, an immune cell gene set was retrieved from the List of Pan-cancer Immune Metagenes (PMID 28052254).

2.2. Bacterial microbiota atlas and LEfSe analysis

As microbial community samples are obtained from both the genome and transcriptome levels, they are merged by the TCGA_ID of the sample (taking the maximum value). The microbial abundance of each sample was obtained from the microbiome and the relative abundances of identified taxa were displayed from kingdom to genus levels. Box plots showed the relative abundance of the top 10 microbes. The LEfSe method was then used to detect particular enhanced microbial biomarkers in EC samples and paracancerous samples by Tutools (<https://www.cloudtutu.com>), a free online data analysis platform. The threshold score of linear regression analysis (LDA) was set at 3.0 to identify the main characteristic microorganisms. LEfSe analysis necessitates the input of a species' taxon, and each taxon must be associated with relevant information; a few species with no registered taxonomic information were removed and only the species with complete taxonomic information for each taxon were preserved for analysis.

2.3. Construction and validation of potential prognostic-related microbial markers

The TCGA-UCEC trial collected tumor samples that were subsequently divided into training and testing cohorts at a 5:5 ratio. Univariate Cox regression analysis was conducted using the survival (v3.2-7) and survminer (v0.4.8) packages in the R environment; this allowed us to assess the association between overall survival (OS) and the abundance of characteristic microbial markers in the training dataset. Microbial signatures linked to OS in EC patients were identified, with a significance threshold of $P < 0.05$. Subsequently, we employed LASSO Cox regression analysis, utilizing the “glmnet (v4.0-2)” package in R, to identify the most pertinent prognostic microbiological markers and construct a risk score model. To enhance the precision of the model, a 10-fold cross-validation method was employed to select the lambda parameter. The risk score for each sample was determined using the formula: $Rscore_i = \sum_{j=1}^n abundance_{ji} \times \beta_j$, where “abundance” represents the microbial marker's abundance, “ β ” signifies the marker's regression coefficient derived from LASSO regression, “RScore” is the sum of the product of abundance and coefficient for each significant marker in a given sample, “ i ” denotes the sample, and “ j ” corresponds to the microbial marker. To assess the effectiveness of the model, we computed risk scores for each patient and categorized them into high-risk and low-risk groups based on the median cutoff. Kaplan-Meier (KM) survival curves were plotted using OS and progression-free interval (PFI) data, and p -values were calculated. A statistically significant difference between the high and low-risk groups was determined when $P < 0.05$. The sample risk score served as the model's prediction outcome, and the area under the ROC curve (AUC) was calculated based on survival data. The ROC curve was generated using the training dataset, and the prognostic microbial biomarkers exhibited diagnostic efficacy exceeding 0.6 for the AUC at 2, 4, 6, and 8 years. To validate the performance of the model, risk scores were calculated in the validation set using the same formula applied in the training set. This validation process included plotting KM survival curves and ROC curves to ensure the reliability of the novel model.

2.4. Nomogram construction

A nomogram integrates multiple predictive variables to accurately estimate the individual probability of clinical events and provide a quantitative tool for predicting outcomes in patients with a disease of interest [19]. This entire process involves the use of well-packaged R packages for analysis. Integrating this with the additional information, a nomogram assimilates various predictive variables to accurately determine the individual likelihood of clinical events, serving as a quantitative instrument for forecasting outcomes in patients with specific diseases. First, we used the “rms (v6.1-0)” and “survival (v3.2-7)” packages in R to generate a nomogram according to independent prognostic factors. Then, we performed Cox proportional hazards regression analysis by function `cph` using the following covariates: age, grade, and clinical stage. Then, survival probabilities were calculated using the `Surv` function. Finally, we generated a nomogram to predict OS at the 2-, 4-, 6-, and 8-year stages using the function `nomogram`; this was displayed by the `plot` method. In addition, we computed the C-index for both the training and validation datasets and generated statistical charts to illustrate the results.

2.5. Prognostic signature-related immune cells and function

Next, we explored disparities in tumor-immune cell infiltration and immune functionality in both normal and EC tissues. By applying single-sample gene set enrichment analysis, we computed enrichment scores for 28 immune infiltrating cell types in cancer samples by utilizing the GSVA R package (version 1.34.0). Following this, we standardized the data using the `scale` function and generated box plots to illustrate distinctions between the low- and high-risk score cohorts. Furthermore, we determined the enrichment scores for immune function in the samples, standardizing the data with the `scale` command in R, and presenting the findings with box plots. To assess differentially expressed immune checkpoints, we procured MSI score data from the cBioPortal database. Variations in MSI across groups were calculated, and violin plots were constructed to provide a visual representation of data distributions. Information on TMB and HRD scores for tumor samples was sourced from PMID: 29617664. To gauge the prognostic significance of these scores, we obtained tumor immune dysfunction and exclusion (TIDE) scores from an online database. In addition, we retrieved immunophenoscores (IPS) data from the TCGA-UCEC cohort through the TCIA database. For insights into immune checkpoint expression disparities between groups, we referenced PMID: 32814346, quantified the variations in immune checkpoint expression, and illustrated these distinctions with box plots.

3. Results

3.1. Population classification

Table 1 presents the characteristics of the 239 participants whose samples were included in the TCGA-UCEC dataset, including 90 patients over 65 years old and 149 patients aged 65 years or younger.

3.2. Screening and verification of prognostic microorganisms

Microbiome data arising from EC samples and control tissues were compiled using the TCGA_ID number. We collected 239 cancer samples and 116 paracancerous samples. The average abundance was then aggregated based on kingdom, phylum, class, order, family, and genus; then, the top 10 microorganisms were selected for presentation. Analysis revealed that the proportion of each grade was uniformly distributed in cancer and paracancerous samples (Fig. 1A-F).

3.3. Application of the LEfSe method to identify significant microbial biomarkers

Next, LEfSe analysis was performed on cancer and paracancer samples; this identified a total of 827 microbial markers with significant differences between the two groups. Of these markers, 417 belonged to the tumor group and had a higher LDA, while the remaining 411 belonged to the normal group with a lower LDA (Fig. 2A and Supplementary Table 1). The cladogram generated by LEfSe and shown in Fig. 2B demonstrated that the p_Actinobacteria and p_Bacteroidetes phyla played an important role in the EC group, while the p_Firmicutes and p_Euryarchaeota phylum played a key role in the paracancerous group. There were 827 differential microbial markers encompassing various taxonomic levels. Following integration, a total of 485 differential microorganisms were identified at the genus level.

3.4. Construction and validation of a prognostic signature

We initiated our analysis by conducting univariate Cox regression to investigate the connection between individual microbial indicators and the survival rates of patients with EC. We calculated the hazard ratios (HR) for each bacterial marker to ascertain their potential roles in the progression and prognosis of EC patients. Initially, we divided the data randomly into two sets, with a 50:50 ratio. The training set comprised 120 samples, while the validation set contained 119 samples. Subsequently, employing information on 485 distinct microbial markers at the genus level and OS data, we performed univariate Cox regression analysis to pinpoint microbial markers significantly linked to survival. By applying a correlation threshold, we identified 26 microbial markers that exhibited a significant association with prognosis, as illustrated in Fig. 3A. Our analysis of these microbial markers revealed that six markers exhibited an HR greater than 1, thus indicating that an increase in their abundance might be correlated with a poor prognosis in EC patients. Conversely, 20 markers exhibited an HR less than 1, suggesting that a reduction in their abundance could be linked to a poorer prognosis, as depicted in Fig. 3B and Supplementary Table 2. To further investigate the significance of microbial markers with notable *p* values, we conducted KM survival curve analyses, as depicted in Fig. 3C.

3.5. LASSO Cox regression analysis

Next, we performed LASSO Cox regression analysis and integrated microbial markers to formulate a prognostic model. Subsequently, we computed risk scores for the training dataset. Using LASSO regression to reduce the dimensionality of the results obtained from univariate Cox regression analysis, we identified ten microbial markers to establish a risk-scoring model. The risk score was generated using the following equation: Risk Score = *g_Marinobacterium* * (-0.3314) + *g_Nitrobacter* * (-0.1936) + *g_Mycetocola* * (-0.1220) + *g_Zobellia* * (-0.1203) + *g_Ottowia* * (-0.0935) + *g_Leifsonia* * (-0.0608) + *g_Cyclobacterium* * (-0.0106) + *g_Nitri-liruptor* * 0.1306 + *g_Saccharicrinis* * 0.2628 + *g_Buchnera* * 0.3265 (Fig. 4A). Based on their risk scores, we classified the samples into high- and low-risk groups and plotted a KM curve using OS data (Fig. 4B). Our analysis revealed a significant disparity in survival

Table 1
Clinical information relating to samples from the TCGA-UCEC dataset.

Clinical data of TCGA-UCEC tumor samples (239)		
Age	>65	90
	≤65	149
Grade	G1	71
	G2	79
	G3	84
	High Grade	5
Clinical_stage	Stage I	163
	Stage II	17
	Stage III	46
	Stage IV	13

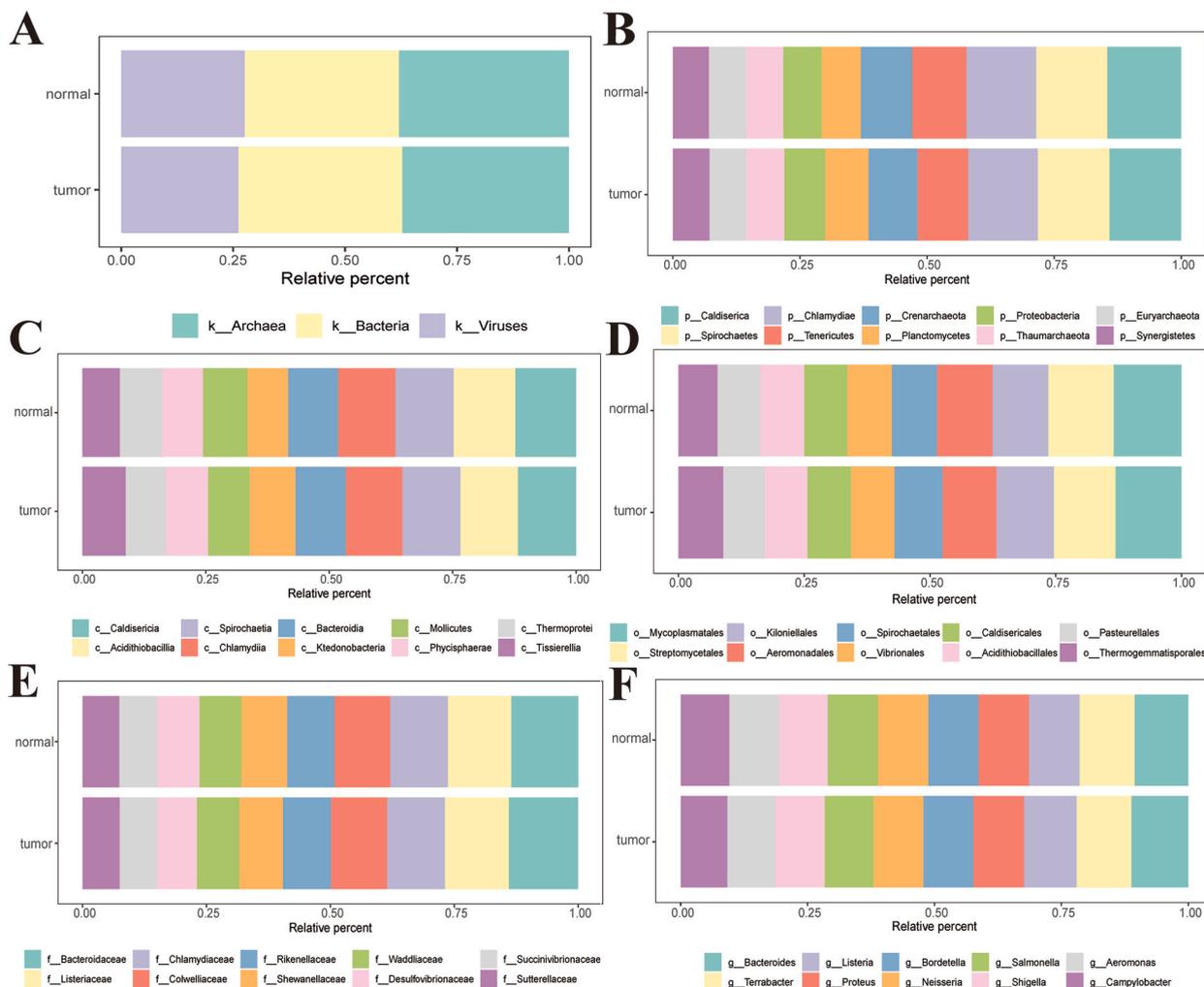


Fig. 1. A comprehensive comparison of microbiota distribution between normal and EC samples. The distribution of microbiota proportions in normal and EC samples was analyzed at the kingdom (A), phylum (B), class (C), order (D), family (E), and genus (F) levels. Each panel visually depicts the relative abundance of the microbiota at its respective taxonomic level, thus offering a detailed insight into the microbial differences between normal and EC tissues.

curves between the two groups ($P < 0.0001$). Nonetheless, the distinction based on PFI data was found to be non-significant (Fig. 4B, $P = 0.24$). Subsequently, we employed the risk score as a predictive model and calculated the AUC value in conjunction with survival data. The calculated AUC values for the model at 2, 4, 6, and 8 years all exceeded 0.6, signifying robust predictive performance (Fig. 4C).

3.6. Model testing on the validation set

Subsequently, the model's performance underwent further validation using the validation dataset. The sample risk scores were computed following the previously outlined procedure, and the division into high-risk and low-risk groups was determined using the median as the cutoff threshold. KM survival curves were generated based on OS data, revealing a statistically significant distinction between the two groups (Fig. 5A, $P = 0.022$). In addition, the KM survival curves stratified by PFI also revealed a significant difference (Fig. 5A, $P = 0.017$). Utilizing the sample risk scores as model predictions, the AUC values for 2, 4, 6, and 8-year survival predictions all exceeded 0.6 when correlated with actual survival outcomes, thus demonstrating strong model performance (Fig. 5B).

3.7. Construction and validation of a nomogram model

Subsequently, we utilized the risk score, age, grade, and clinical stage data from both the training and validation datasets to construct nomograms intended to guide clinical decision-making (Fig. 6A, B). In addition, we generated ROC curves based on these nomograms. Our analysis revealed that the AUC for 2, 4, 6, and 8-year predictions exceeded 0.65 in both the training and validation

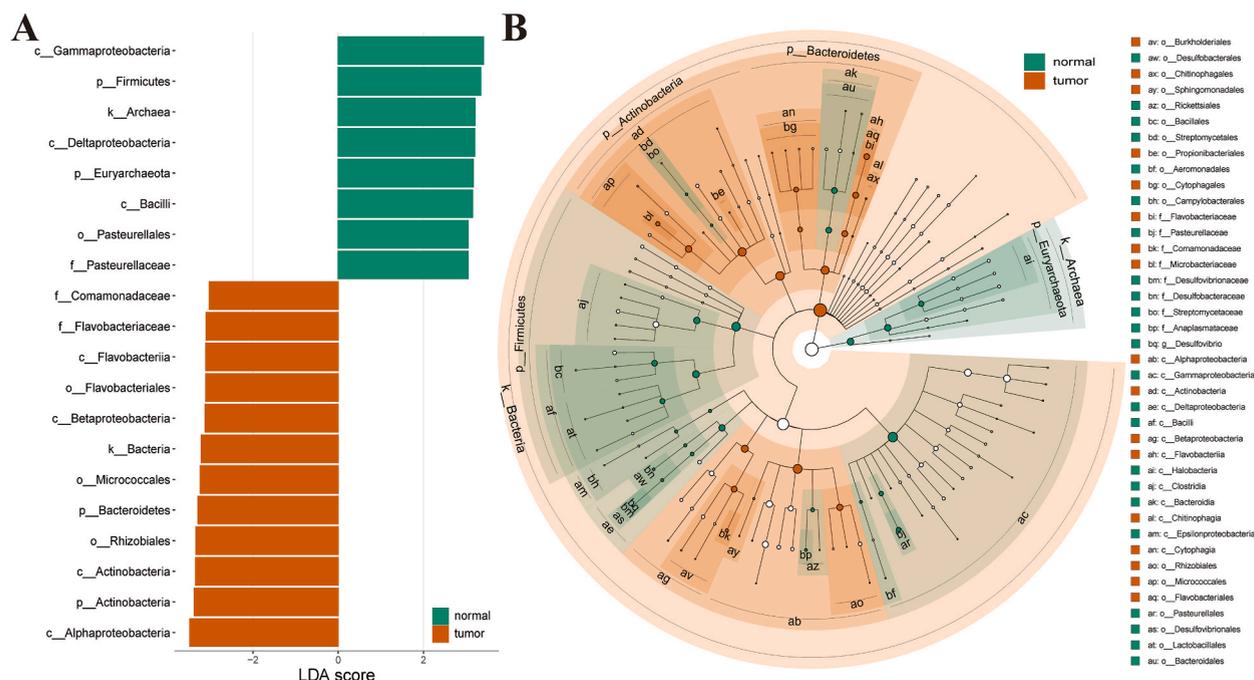
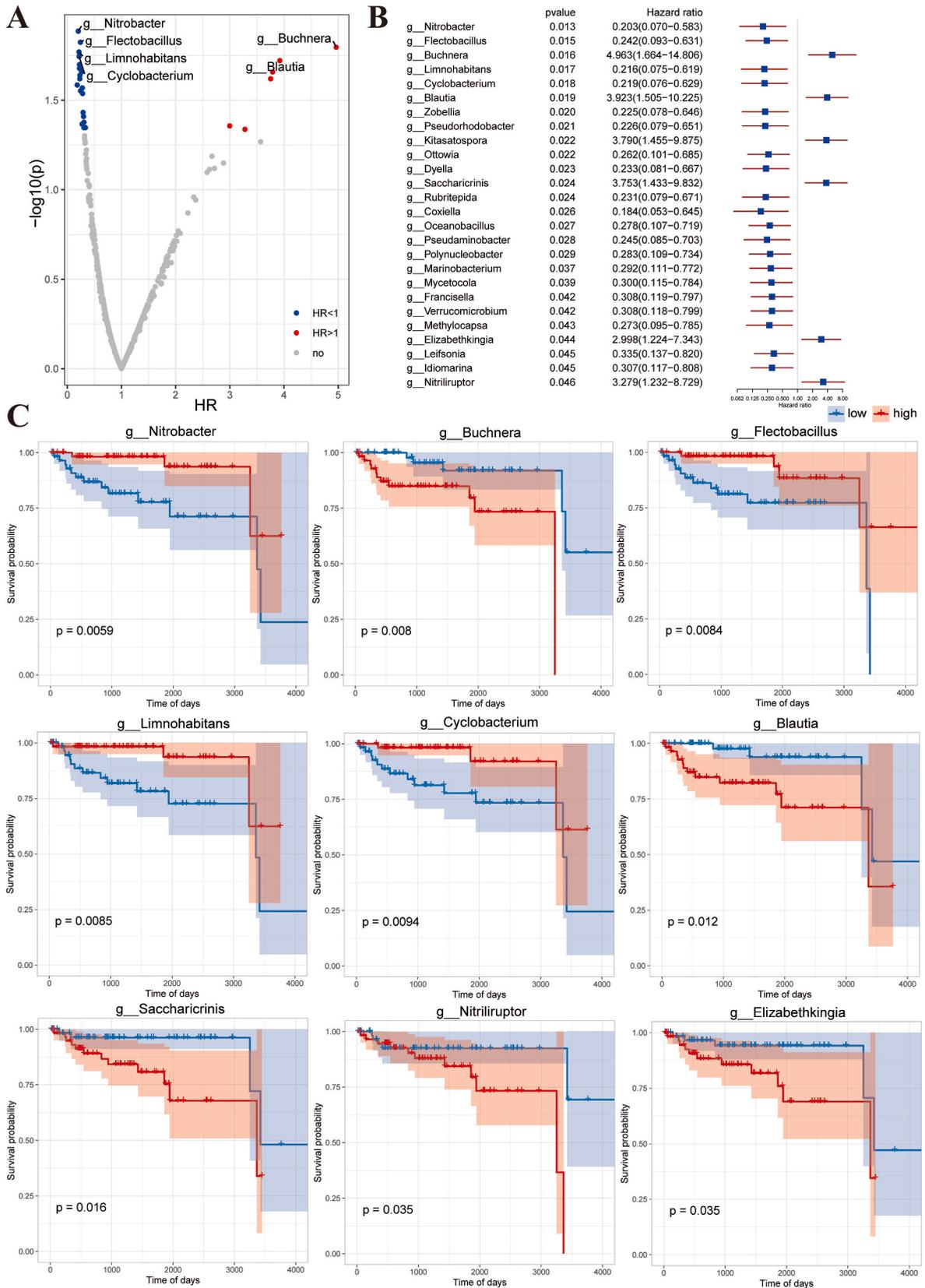


Fig. 2. Comparative analysis of bacterial taxa in paracancerous control and EC samples. (A) LefSe analysis revealed distinct bacterial taxa abundances between paracancerous control samples and EC samples. Horizontal bars illustrate the effect size for each taxon, with green denoting taxa enriched in the control group and yellow denoting taxa enriched in the EC group. (B) A cladogram showing the microbial taxa exhibiting substantial differences in abundance between paracancerous and EC samples. In the cladogram, taxonomic hierarchies were sorted from inside to outside (from genus to kingdom).

sets. Notably, the predictive performance of the nomogram for the training set closely resembled that of the risk score alone, as depicted in Fig. 6C. However, when validated with the independent validation set, the nomogram exhibited significantly superior prognostic performance when compared to the risk score alone, as illustrated in Fig. 6C. Furthermore, we computed the C-Index for both the nomogram and the risk score within both the training and validation sets. In both cases, the C-Index exceeded 0.5, indicating favorable predictive performance. It is worth noting that the C-Index of the nomogram in the training set (0.8327 [95% CI: 0.7813–0.8842]) was marginally lower than that of the risk score (0.9027 [95% CI: 0.8719–0.9335]). However, in the validation set, the nomogram exhibited significantly superior performance (C-Index: 0.7698 [95% CI: 0.7101–0.8295]) compared to the risk score (C-Index: 0.6029 [95% CI: 0.5383–0.6675]), as shown in Fig. 6D. Furthermore, we conducted comprehensive analysis of both the training and validation sets using single-factor and multi-factor regression models that incorporated the risk score, age, grade, and clinical stage groupings. Our findings consistently demonstrated that the risk score stood out as a highly significant prognostic factor in both datasets and outperformed other clinical feature groupings in predicting patient outcomes (Fig. 6E, F).

3.8. Microbial communities with unique characteristics and associations with hosts

Next, we conducted statistical analysis on the correlation between the abundance of microbial biomarkers within the model and protein-coding gene expression in EC samples. Analysis indicated that there was a correlation between the characteristic model of microbial species abundance in EC samples and sample gene expression. First, we obtained 19597 protein-coding genes; then, based on a correlation threshold, we identified 422 groups of significantly correlated combinations, including 397 protein-coding genes. The correlation heatmap showed that in the model, the proportion of positive correlation between microbial species (*g_Nitrospirillum*, *g_Saccharicrinis*, and *g_Buchnera*) and related genes was higher, while the proportion of negative correlation between other microorganisms and related genes was higher (Fig. 7A and Supplementary Table 3). Subsequently, we conducted an analysis of host gene functional enrichment for the 397 protein-coding genes. The Gene Ontology enrichment analysis was divided into three components: biological processes (BPs), cellular components (CCs), and molecular functions (MFs). The enriched pathways for BPs were mainly ribonucleoprotein complex biogenesis, ncRNA processing, mRNA processing, and RNA splicing. The enriched pathways for CCs were mainly the mitochondrial inner membrane, mitochondrial protein-containing complex, and mitochondrial matrix. The enriched pathways for MFs were mainly GTP binding and guanyl nucleotide binding. The enriched pathways identified by KEGG analysis were mainly pathways associated with neurodegeneration-multiple diseases, Huntington disease, and Alzheimer’s disease (Fig. 7B).



(caption on next page)

Fig. 3. Univariate Cox regression analysis was employed to identify microbial markers linked to patient survival. (A) A volcano plot visualizing the predictive significance of differential microbial markers based on their abundance, providing a comprehensive overview of those with potential relevance for patient survival. (B) The forest plot shows the hazard ratios, confidence intervals, and statistical significance of each microbial marker, as determined by univariate Cox regression analysis, offering insights into their individual impact on survival outcomes. (C) KM survival curves are shown for the selected microbial markers, thus demonstrating their distinct survival probabilities over time. These curves are accompanied by *p*-values, highlighting their statistical significance in the context of patient survival.

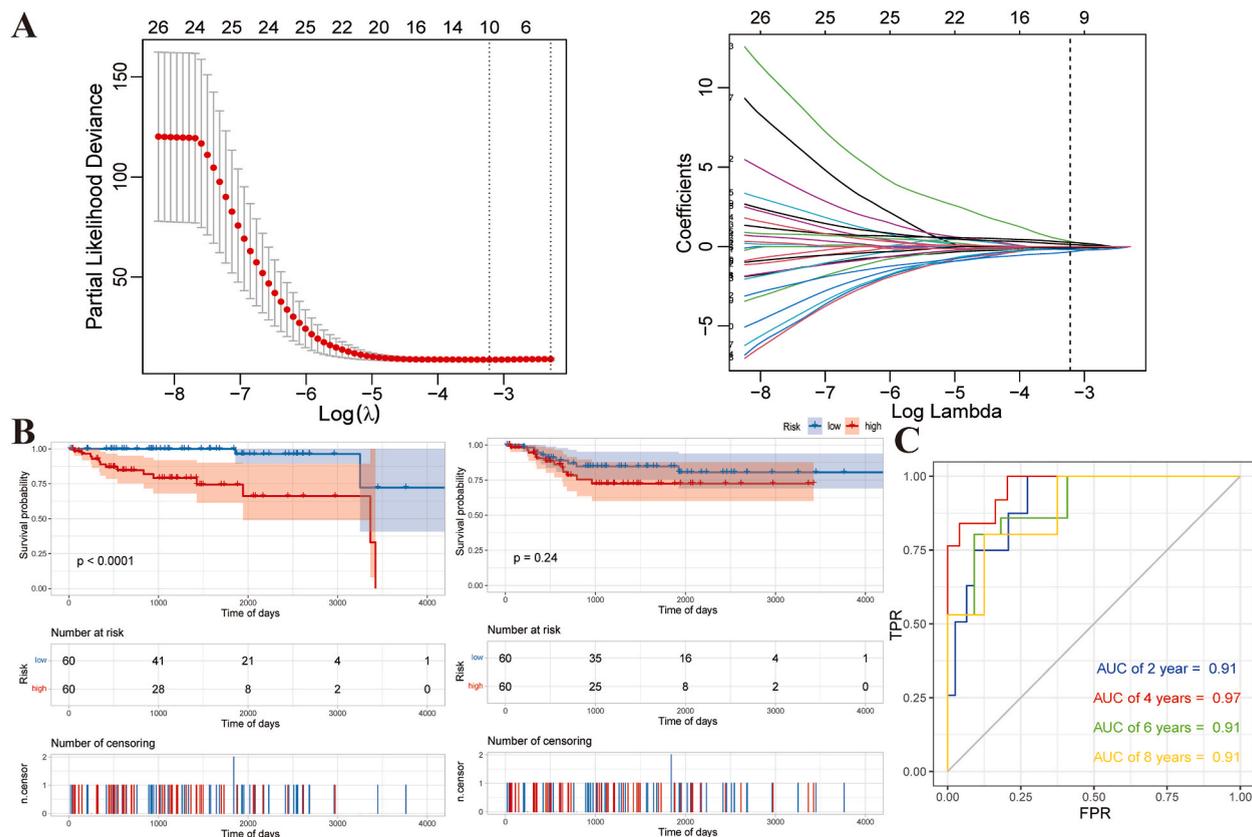


Fig. 4. Comprehensive visualization of model parameters and performance. (A) The panel illustrates the tuning parameter lambda (λ) determined through a 10-fold cross-validation process, showcasing the profiles of LASSO coefficients for 26 microbial markers. (B) In addition, the figure also included KM curves, thus providing a visual representation of survival probabilities over time for high- and low-risk groups. (C) ROC curves are presented, demonstrating the model's predictive accuracy in distinguishing between high- and low-risk groups within the training set.

3.9. Immune cells and functionality associated with prognostic signature models

To assess immune cell infiltration and demonstrate functional immune disparities, we employed gene set variation analysis to probe the association between risk scores and immune infiltration among high- and low-risk EC patients. Notable variations were identified (35.7%; 10/28) in the proportions of immune infiltrating cell types between the high-risk and low-risk cohorts (Fig. 8A). Analysis showed that the high-risk group exhibited a notably greater abundance of memory B cells compared to the low-risk cohort. Conversely, the low-risk group showed significantly elevated proportions of activated B cells, activated CD8 T cells, CD56^{dim} NK cells, macrophages, MDSCs, monocytes, natural killer T cells, and plasmacytoid dendritic cells in comparison to the high-risk group. Furthermore, we conducted supplementary analyses to scrutinize the discrepancies in the scores of thirteen immune cell-associated functions between the high-risk and low-risk groups. These analyses revealed that four immune cell-related functions, encompassing cell cycle regulation, checkpoint mechanisms, T cell co-stimulation, and type II IFN response, exhibited significantly higher enrichment scores within the low-risk group (Fig. 8B).

3.10. Differences in immune therapy indicators and checkpoint expression

We conducted a comprehensive statistical analysis to compare the expression levels of immune indicators and checkpoints between two distinct risk groups: high-risk and low-risk. Our analysis revealed that there were no statistically significant differences observed in

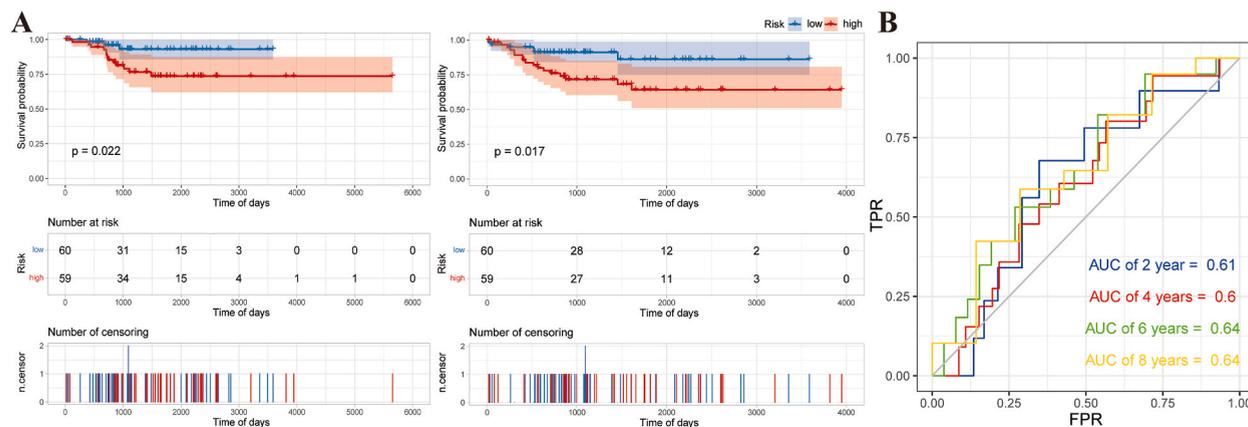


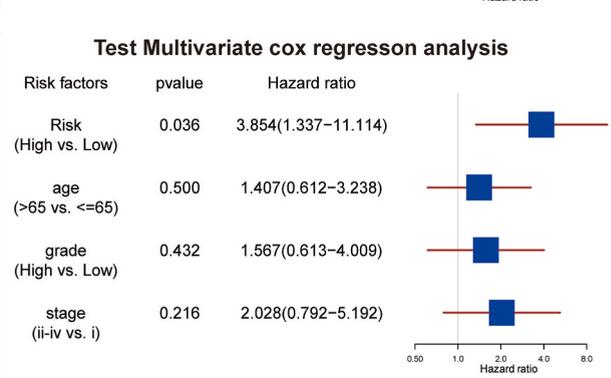
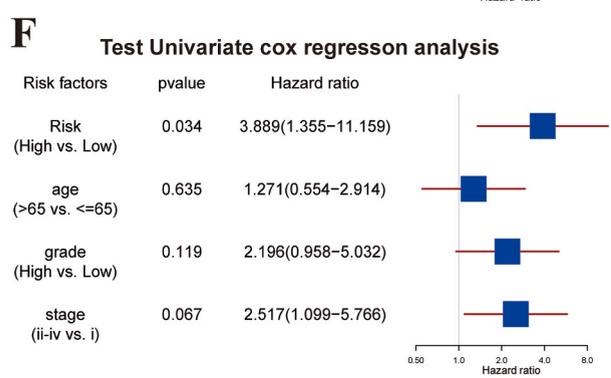
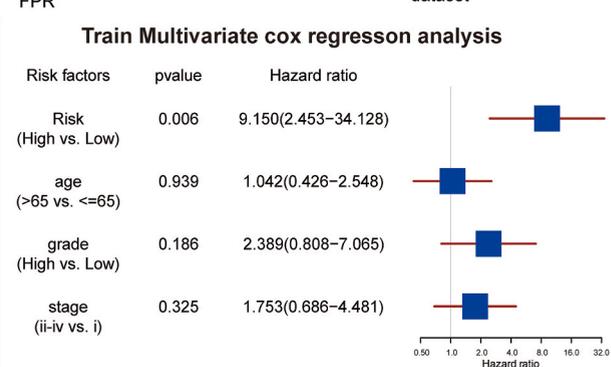
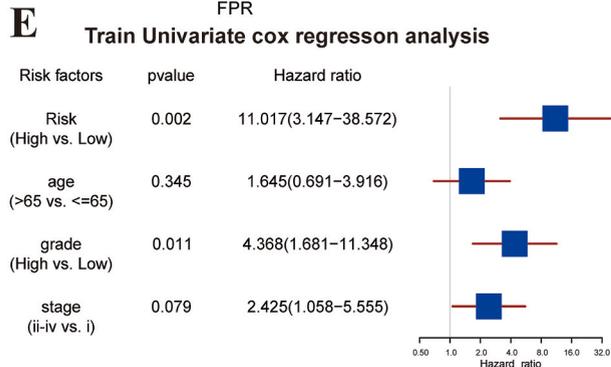
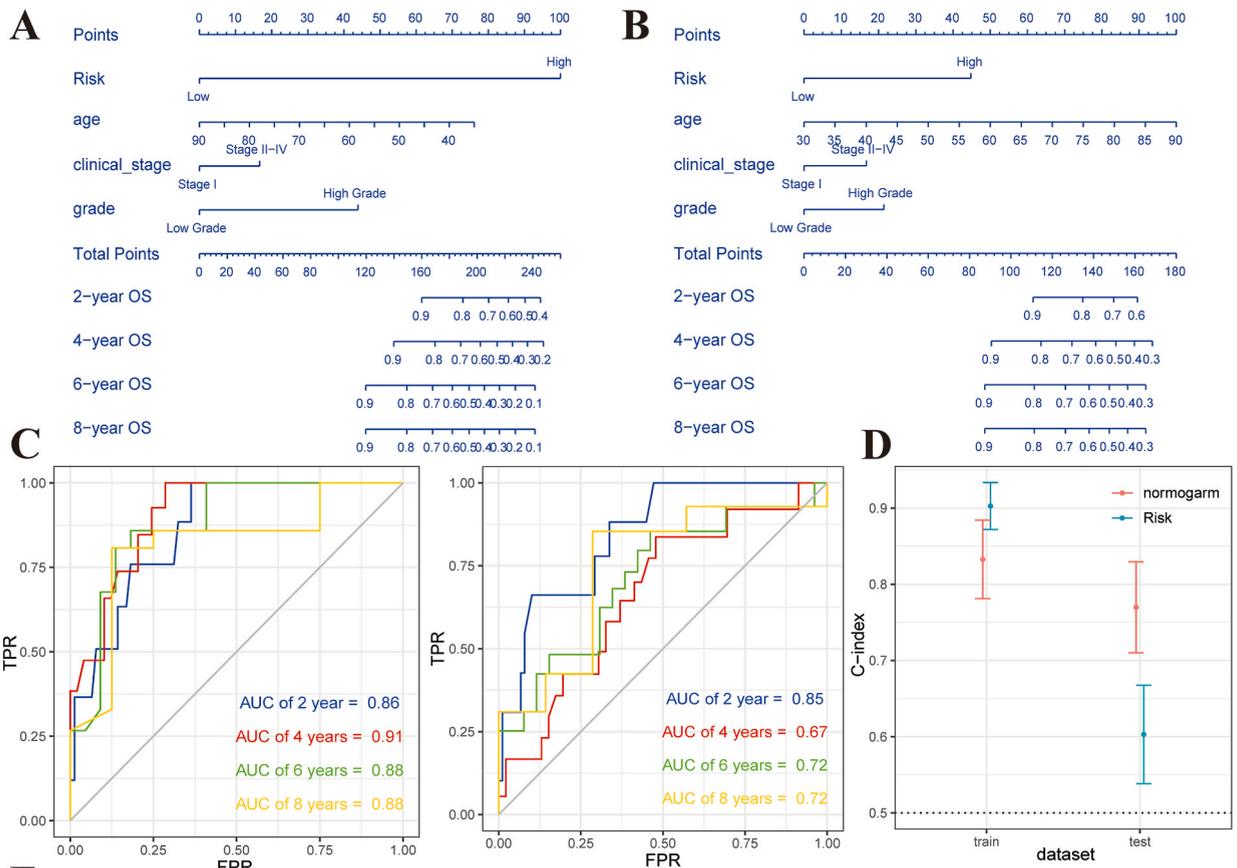
Fig. 5. KM survival analyses and time-dependent ROC curves derived from the risk score in the validation datasets. (A) KM curves demonstrating survival probabilities for high- and low-risk groups over time in the validation set. (B) ROC curves illustrating the diagnostic performance of the risk score in distinguishing between high- and low-risk groups within the validation set.

MSI, TMB, or HRD between these two risk groups. To evaluate a patient's potential response to immune checkpoint blockade therapy, we employed the TIDE score, a recognized biomarker. A lower TIDE score signifies a reduced likelihood of immune escape, suggesting a more favorable response to immunotherapy [20]. Therefore, we utilized the TIDE score (accessible at <http://tide.dfci.harvard.edu/>) to assess the prognostic relevance of our immune response prediction signature for the success of immunotherapy within both low-risk and high-risk EC subgroups. Our study revealed a significant variation in TIDE scores between the high-risk and low-risk groups (Fig. 9A, $P < 0.01$). This variation suggests that high-risk group members might have less positive outcomes from immunosuppressive treatments compared to their low-risk counterparts. Moreover, individuals in the low-risk group appeared to be more responsive to immunotherapy. In addition, the scores for IPS-CTLA4 and PD1/PD-L1/PD-L2 blockers were substantially higher in the low-risk group ($P < 0.05$), suggesting a greater potential benefit from immune therapy for these patients (Fig. 9A). In summary, the microbiome-based risk score model effectively categorizes patients, pinpointing those likely to gain from immunotherapy. We also observed substantial distinctions in the expression levels of 17 out of 64 immune checkpoints between the high-risk and low-risk groups. Of these, three checkpoints exhibited significantly higher expression in the high-risk group, while the remaining 14 showed significantly higher expression in the low-risk group, which included notable markers, including CD27 and TNFSF14 (Fig. 9B).

4. Discussion

Mounting evidence underscores the critical role of the tumor microbiome in influencing cancer progression and metastasis, as well as modulating patient responses to immunotherapy, which consequently impacts survival outcomes. The microbial communities within the tumor microenvironment are crucial in shaping immune responses [21,22]. The links between specific bacteria, such as *Atopobium vaginae* and *Porphyromonas* sp., and EC, particularly under high pH conditions, pose intriguing questions about the vaginal and uterine microbiota's role in gynecological health [15]. We observed a significant correlation between microbiota variations and EC, highlighting the role of the microbiome in cancer development and progression. The expansive sample size of the TCGA-UCEC cohort enhances the credibility of our findings, including the identification of distinct microbial communities in EC and adjacent non-cancerous tissues. This discovery not only reinforces the concept of the uterus as a non-sterile environment but also suggests microbial imbalances as key factors in EC onset and development. Our study provides groundbreaking insights into the complex interactions between the microbiota and EC, opening new paths for understanding and managing this prevalent gynecological malignancy.

In an innovative approach, we employed single-factor Cox regression analysis, revealing an association between increased *Nitrospirillum* bacterial abundance and poorer EC prognosis. In addition, LASSO regression analysis led to the generation of a risk scoring model for EC, incorporating the *Nitrospirillum* microbiome. Previous research has linked *Nitrospirillum* to colorectal cancer (CRC), revealing a strong correlation between *Nitrospirillum* abundance and the expression of host genes in advanced-stage CRC [23,24]. Moreover, a forest plot derived from univariate COX regression analysis demonstrated that a higher abundance of *Blautia* bacteria correlated with a poorer clinical prognosis in EC patients (HR, 3.923; 95% CI: 1.505–10.225; $P = 0.019$). Previous research has linked the increased presence of *Blautia* sp. in the intestines of early-stage breast cancer patients to clinical staging and histological prognostic grading, with significantly higher levels observed in clinical stage groups II/III compared to clinical stages 0/I. Moreover, *Blautia* sp. has also been associated with worse clinical outcomes and advanced clinical stages, potentially due to its involvement in estrogen and phytoestrogen metabolism [25]. Given the strong estrogen environment in the majority of EC cases, our results suggest that *Blautia* bacteria might serve as a microbial marker for EC patients, partly by participating in the estrogen signaling pathway. Notably, in our current study, *Elizabethkingia* microorganism emerged as a prognostic marker in EC patients (HR, 2.998; 95% CI: 1.224–7.343; $P = 0.044$), as determined by univariate Cox proportional hazard regression analysis in the TCGA-UCEC cohort. A previous case report described a 56-year-old patient with nasopharyngeal carcinoma, who developed bacterial encephalitis complicated by an



(caption on next page)

Fig. 6. Comprehensive nomograms for predicting OS in EC Patients. Detailed nomograms for predicting OS in EC patients, separately presented for the training (A) and validation (B) datasets. (C) ROC curves for evaluating the predictive accuracy of the nomograms in both the training and validation sets. (D) C-Index values for the nomograms and corresponding risk scores, presented for both the training and validation datasets. Analysis of independent prognostic factors in the EC patient cohorts, with results shown for the training (E) and validation (F) datasets.

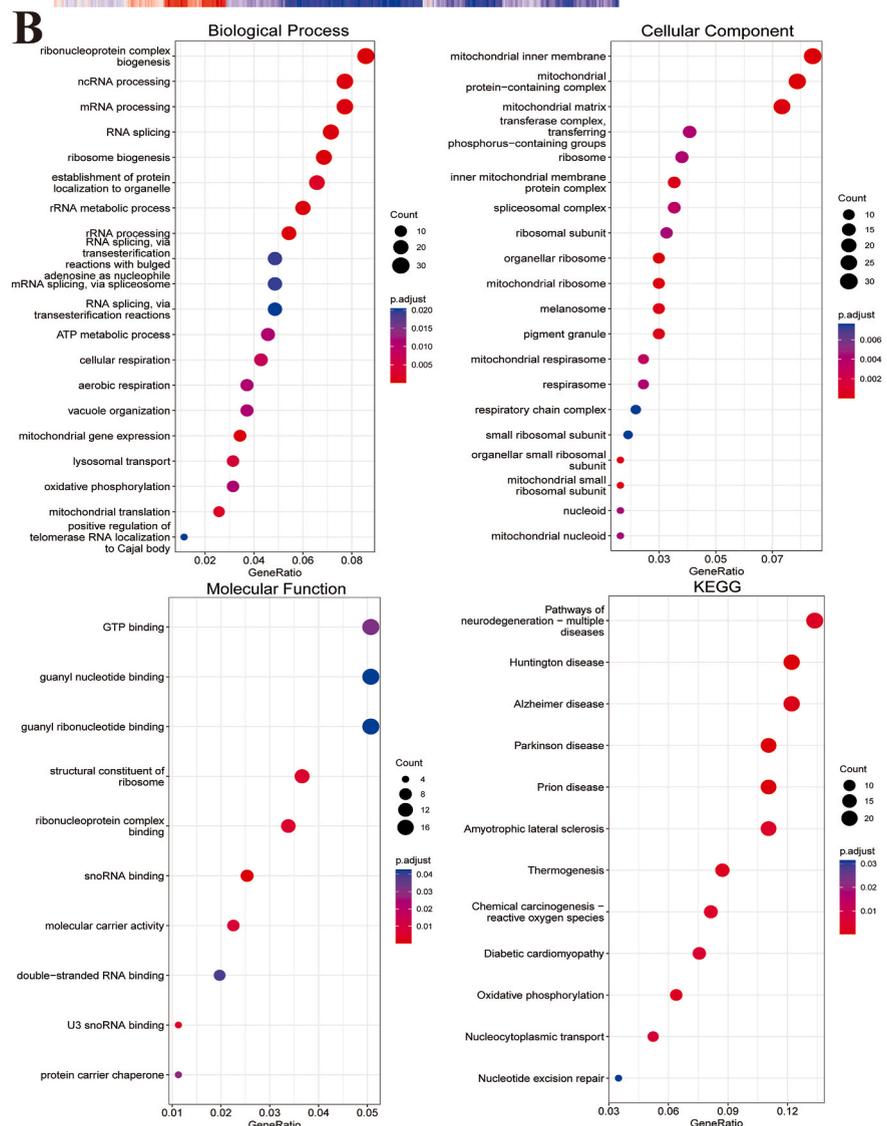
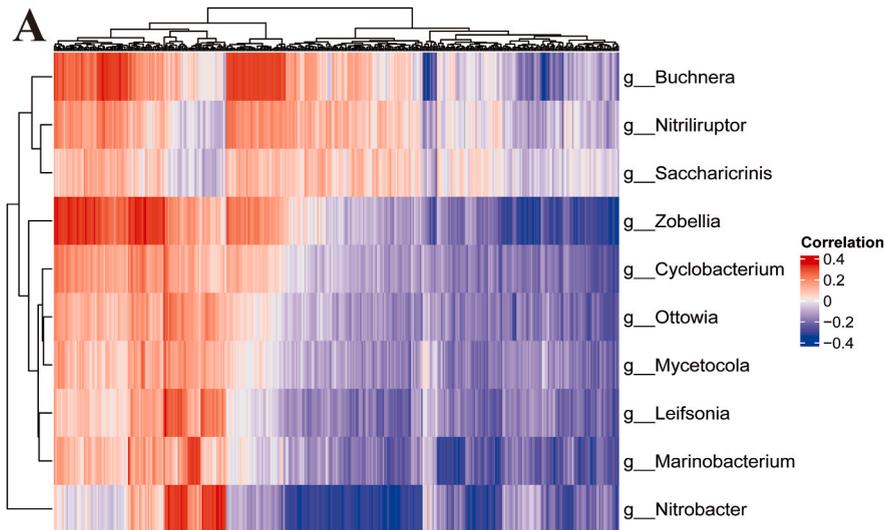
Elizabethkingia miricola infection, which was successfully treated with antibiotics [26]. *Elizabethkingia miricola* is commonly linked to adverse outcomes in patients with other comorbidities [27]. Moreover, this aligns with findings from previous studies, which have underscored that microbiome-induced inflammation can accelerate tumorigenesis by activating tumor-associated inflammatory signaling pathways [28,29].

Our study revealed significant disparities in immune cell infiltration and immune function between high-risk and low-risk microbial groups in EC patients. Analysis revealed a notably higher proportion of tumor-infiltrating immune cells, including activated B cells, activated CD8⁺ T cells, macrophages, mast cells, MDSCs, monocytes, natural killer T cells, and plasmacytoid dendritic cells, in the low-risk group compared to the high-risk group. Prior research has emphasized the pivotal role of B cell antigen presentation in activating tumor-specific CD8⁺ T cells [30]. Moreover, an association between infiltrating CD8⁺ T cells and prognosis has been observed in various solid malignancies, including colorectal and ovarian cancer [31]. The lower infiltration of CD8⁺ T cells has been linked to a poorer prognosis and is essential for disrupting tumor-associated blood vessels [32]. Analysis also indicated an elevated number of B cells and activated B cells in non-metastatic tumors, with their infiltration in primary tumors correlating with improved survival rates in colorectal cancer patients with liver metastasis. In mouse models, tumor-infiltrated activated B cells in colorectal cancer reduce liver metastasis, while increased tumor infiltration of activated B cells diminishes CRC development and liver metastasis [33]. Our study identified a significant increase in activated B cells and activated CD8⁺ T cells in the low-risk EC group, suggesting a more potent anti-tumor effect and a more favorable prognosis. Moreover, our findings emphasize the importance of the immune landscape in EC for developing personalized treatment strategies to improve patient outcomes, including immunotherapy response prediction, the variation in TIDE scores and higher scores for immune checkpoint blockers (such as IPS-CTLA4 and PD1/PD-L1/PD-L2) in the low-risk group suggesting that these patients are more likely to benefit from immunotherapy. This is pivotal for tailoring treatment strategies, as it indicates a higher potential for a positive response to immune checkpoint blockade therapy in low-risk EC patients. Furthermore, the microbiome-based risk score model effectively categorizes EC patients into high-risk and low-risk groups. This stratification is crucial for personalizing treatment approaches, particularly in deciding who may benefit most from immunotherapy. Thus, the observed distinctions in immune function and cell infiltration between high-risk and low-risk groups hold considerable implications for immunotherapy strategies.

The present study highlighted the potential of microbes to enhance both anti-tumor immunity and the efficacy of immunotherapy. By applying the TCGA dataset on EC microbiomes, our findings not only identify disparities in microbiota composition between tumor and normal tissues but also generate a survival model based on microbiota abundance. Subsequently, we utilized the risk score, age, grade, and clinical stage data from both the training and validation datasets to construct nomograms intended to guide clinical decision-making. This represents a significant step in applying the insights from our study on microbiome-based risk scoring, transforming them into practical tools for clinicians. Research on microbial signatures associated with EC not only provides a fresh pathway for personalized medicine but also revolutionizes the personalization of gynecological oncology treatment. By harnessing microbiota data to develop survival models and construct these nomograms, clinicians can improve the prediction of patient outcomes, enabling them to customize treatment strategies more effectively. This approach opens up new possibilities for impacting the screening, risk assessment, treatment, and monitoring of EC patients. The integration of microbiota profiling, alongside traditional clinical parameters in these nomograms, could significantly alter the standard of treatment for EC, paving the way for more efficient and tailored treatment plans. Novel approaches, such as probiotic and microbiota transplants, will provide a new frontier in EC therapy by modifying the composition of the endometrial microbiota. These alternative treatments present a promising yet complex challenge, potentially revolutionizing our approach to uterine microorganism modulation and thereby enhancing treatment efficacy and patient outcomes in the battle against EC. Our study, while comprehensive in its approach to understanding the role of microbiota in EC, has several limitations that must be acknowledged. The observational nature of our research implies a limitation in establishing causality between microbiota changes and cancer progression. We also recognize that our findings are derived from a specific dataset (TCGA-UCEC cohort), which may limit the generalizability of our results to all EC populations. Overall, the robust prognostic performance of the microbiome-based risk score model underscores its significance for future investigations into microbial communities in cancer patients. These findings offer fresh insights into the intricate relationship between microbiota and tumors. The molecular mechanisms uncovered hold promise for uncovering innovative therapeutic targets and biomarkers for EC patients.

5. Conclusion

This study illuminates the significant impact of uterine microbiota on EC, revealing distinct microbial patterns in cancerous and adjacent tissues. Our findings, employing extensive data analysis and risk score modeling, highlight the potential of microbiota as a biomarker for tailoring EC treatment strategies. The research underscores the crucial role of microbiota in immunotherapy efficacy, offering novel insights for personalized medicine in gynecological oncology.



(caption on next page)

Acknowledgements

None declared.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.heliyon.2024.e27879>.

References

- [1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *CA A Cancer J. Clin.* 71 (2021) 209–249, <https://doi.org/10.3322/caac.21660>.
- [2] N. Colombo, C. Creutzberg, F. Amant, T. Bosse, A. González-Martín, J. Ledermann, C. Marth, R. Nout, D. Querleu, M.R. Mirza, C. Sessa, ESMO-ESGO-ESTRO consensus conference on endometrial cancer: diagnosis, treatment and follow-up, *Ann. Oncol.* 27 (2016) 16–41, <https://doi.org/10.1093/annonc/mdv484>.
- [3] K. Nakayama, N. Nakayama, M. Ishikawa, K. Miyazaki, Endometrial serous carcinoma: its molecular characteristics and histology-specific treatment strategies, *Cancers* 4 (2012) 799–807, <https://doi.org/10.3390/cancers4030799>.
- [4] R. Sender, S. Fuchs, R. Milo, Revised estimates for the number of human and bacteria cells in the body, *PLoS Biol.* 14 (2016) e1002533, <https://doi.org/10.1371/journal.pbio.1002533>.
- [5] G.D. Poore, E. Kopylova, Q. Zhu, C. Carpenter, S. Fraccacio, S. Wandro, T. Kosciolk, S. Janssen, J. Metcalf, S.J. Song, J. Kanbar, S. Miller-Montgomery, R. Heaton, R. McKay, S.P. Patel, A.D. Swafford, R. Knight, Microbiome analyses of blood and tissues suggest cancer diagnostic approach, *Nature* 579 (2020) 567–574, <https://doi.org/10.1038/s41586-020-2095-1>.
- [6] C.M. Dejea, P. Fathi, J.M. Craig, A. Boleij, R. Taddese, A.L. Geis, X. Wu, C.E. DeStefano Shields, E.M. Hechenbleikner, D.L. Huso, R.A. Anders, F.M. Giardiello, E. C. Wick, H. Wang, S. Wu, D.M. Pardoll, F. Housseau, C.L. Sears, Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria, *Science* 359 (2018) 592–597, <https://doi.org/10.1126/science.aah3648>.
- [7] J. Liu, Y. Zhang, Intratumor microbiome in cancer progression: current developments, challenges and future trends, *Biomark. Res.* 10 (2022) 37, <https://doi.org/10.1186/s40364-022-00381-5>.
- [8] L. Yang, A. Li, Y. Wang, Y. Zhang, Intratumoral microbiota: roles in cancer initiation, development and therapeutic efficacy, *Signal Transduct. Targeted Ther.* 8 (2023) 35, <https://doi.org/10.1038/s41392-022-01304-4>.
- [9] W. Zhu, J.Z. Wang, Z. Liu, J.F. Wei, The bacteria inside human cancer cells: mainly as cancer promoters, *Front. Oncol.* 12 (2022) 897330, <https://doi.org/10.3389/fonc.2022.897330>.
- [10] A. Sipos, G. Ujlaki, E. Mikó, E. Maka, J. Szabó, K. Uray, Z. Krasznai, P. Bai, The role of the microbiome in ovarian cancer: mechanistic insights into oncobiome and to bacterial metabolite signaling, *Mol. Med.* 27 (2021) 33, <https://doi.org/10.1186/s10020-021-00295-2>.
- [11] M. Sobstyl, P. Brecht, A. Sobstyl, P. Mertowska, E. Grywalska, The role of microbiota in the immunopathogenesis of endometrial cancer, *Int. J. Mol. Sci.* 23 (2022), <https://doi.org/10.3390/ijms23105756>.
- [12] A. Fu, B. Yao, T. Dong, Y. Chen, J. Yao, Y. Liu, H. Li, H. Bai, X. Liu, Y. Zhang, C. Wang, Y. Guo, N. Li, S. Cai, Tumor-resident intracellular microbiota promotes metastatic colonization in breast cancer, *Cell* 185 (2022) 1356–1372, <https://doi.org/10.1016/j.cell.2022.02.027>.
- [13] M. Lee, E.B. Chang, Inflammatory bowel diseases (IBD) and the microbiome-searching the crime scene for clues, *Gastroenterology* 160 (2021) 524–537, <https://doi.org/10.1053/j.gastro.2020.09.056>.
- [14] C. Simon, Introduction: do microbes in the female reproductive function matter? *Fertil. Steril.* 110 (2018) 325–326, <https://doi.org/10.1016/j.fertnstert.2018.06.041>.
- [15] M.R. Walther-Antônio, J. Chen, F. Multinu, A. Hokenstad, T.J. Distad, E.H. Cheek, G.L. Keeney, D.J. Creedon, H. Nelson, A. Mariani, N. Chia, Potential contribution of the uterine microbiome in the development of endometrial cancer, *Genome Med.* 8 (2016) 122, <https://doi.org/10.1186/s13073-016-0368-y>.
- [16] W. Lu, F. He, Z. Lin, S. Liu, L. Tang, Y. Huang, Z. Hu, Dysbiosis of the endometrial microbiota and its association with inflammatory cytokines in endometrial cancer, *Int. J. Cancer* 148 (2021) 1708–1716, <https://doi.org/10.1002/ijc.33428>.
- [17] D.M. Walsh, A.N. Hokenstad, J. Chen, J. Sung, G.D. Jenkins, N. Chia, H. Nelson, A. Mariani, M.R.S. Walther-Antonio, Postmenopause as a key factor in the composition of the endometrial cancer microbiome (ECbiome), *Sci. Rep.* 9 (2019) 19213, <https://doi.org/10.1038/s41598-019-55720-8>.
- [18] J. Liu, T. Lichtenberg, K.A. Hoadley, L.M. Poisson, A.J. Lazar, A.D. Cherniack, A.J. Kovatich, C.C. Benz, D.A. Levine, A.V. Lee, L. Omberg, D.M. Wolf, C. D. Shriver, V. Thorsson, H. Hu, An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics, *Cell* 173 (2018) 400–416, <https://doi.org/10.1016/j.cell.2018.02.052>.
- [19] V.P. Balachandran, M. Gonen, J.J. Smith, R.P. DeMatteo, Nomograms in oncology: more than meets the eye, *Lancet Oncol.* 16 (2015) e173–e180, [https://doi.org/10.1016/S1470-2045\(14\)71116-7](https://doi.org/10.1016/S1470-2045(14)71116-7).
- [20] P. Jiang, S. Gu, D. Pan, J. Fu, A. Sahu, X. Hu, Z. Li, N. Traugh, X. Bu, B. Li, J. Liu, G.J. Freeman, M.A. Brown, K.W. Wucherpfnennig, X.S. Liu, Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response, *Nat. Med.* 24 (2018) 1550–1558, <https://doi.org/10.1038/s41591-018-0136-1>.
- [21] L. Yang, A. Li, Y. Wang, Y. Zhang, Intratumoral microbiota: roles in cancer initiation, development and therapeutic efficacy, *Signal Transduct. Targeted Ther.* 8 (2023) 35, <https://doi.org/10.1038/s41392-022-01304-4>.
- [22] F. Gao, B. Yu, B. Rao, Y. Sun, J. Yu, D. Wang, G. Cui, Z. Ren, The effect of the intratumoral microbiome on tumor occurrence, progression, prognosis and treatment, *Front. Immunol.* 13 (2022) 1051987, <https://doi.org/10.3389/fimmu.2022.1051987>.
- [23] I. Uriarte-Navarrete, E. Hernández-Lemus, G. de Anda-Jáuregui, Gene-microbiome Co-expression networks in colon cancer, *Front. Genet.* 12 (2021) 617505, <https://doi.org/10.3389/fgene.2021.617505>.
- [24] I. Romero-Garmendia, K. Garcia-Etxebarria, Host genetics and microbiota interactions in colorectal cancer: shared or independent risk? *Microorganisms* 10 (2022) <https://doi.org/10.3390/microorganisms10112129>.
- [25] T.H. Luu, C. Michel, J.-M. Bard, F. Dravet, H. Nazih, C. Bobin-Dubigeon, Intestinal proportion of *Blautia* sp. is associated with clinical stage and histoprogenic grade in patients with early-stage breast cancer, *Nutr. Cancer* 69 (2017) 267–275, <https://doi.org/10.1080/01635581.2017.1263750>.
- [26] X. Zhuo, Y. Zhou, L. Liu, Acute bacterial encephalitis complicated with recurrent nasopharyngeal carcinoma associated with *Elizabethkingia miricola* infection: a case report, *Front. Neurol.* 13 (2022) 965939, <https://doi.org/10.3389/fneur.2022.965939>.
- [27] P. Zdziański, M. Paściak, K. Rogala, A. Korzeniowska-Kowal, A. Gamian, *Elizabethkingia miricola* as an opportunistic oral pathogen associated with superinfectious complications in humoral immunodeficiency: a case report, *BMC Infect. Dis.* 17 (2017) 763, <https://doi.org/10.1186/s12879-017-2886-7>.
- [28] Y. Wang, G. Yang, L. You, J. Yang, M. Feng, J. Qiu, F. Zhao, Y. Liu, Z. Cao, L. Zheng, T. Zhang, Y. Zhao, Role of the microbiome in occurrence, development and treatment of pancreatic cancer, *Mol. Cancer* 18 (2019) 173, <https://doi.org/10.1186/s12943-019-1103-2>.
- [29] S. Hibino, T. Kawazoe, H. Kasahara, S. Itoh, T. Ishimoto, M. Sakata-Yanagimoto, K. Taniguchi, Inflammation-induced tumorigenesis and metastasis, *Int. J. Mol. Sci.* 22 (2021) 5421, <https://doi.org/10.3390/ijms22115421>.
- [30] R. Tan, M. Nie, W. Long, The role of B cells in cancer development, *Front. Oncol.* 12 (2022) 958756, <https://doi.org/10.3389/fonc.2022.958756>.

- [31] I.K. Jung, S.S. Kim, D.S. Suh, K.H. Kim, C.H. Lee, M.S. Yoon, Tumor-infiltration of T-lymphocytes is inversely correlated with clinicopathologic factors in endometrial adenocarcinoma, *Obstet Gynecol Sci* 57 (2014) 266–273, <https://doi.org/10.5468/ogs.2014.57.4.266>.
- [32] C. Jackaman, C.S. Bundell, B.F. Kinnear, A.M. Smith, P. Fillion, D. van Hagen, B.W. Robinson, D.J. Nelson, IL-2 intratumoral immunotherapy enhances CD8+ T cells that mediate destruction of tumor cells and tumor-associated vasculature: a novel mechanism for IL-2, *J. Immunol.* 171 (2003) 5051–5063, <https://doi.org/10.4049/jimmunol.171.10.5051>.
- [33] Y. Xu, Z. Wei, M. Feng, D. Zhu, S. Mei, Z. Wu, Q. Feng, W. Chang, M. Ji, C. Liu, Y. Zhu, L. Shen, F. Yang, Y. Chen, Y. Feng, J. Xu, D. Zhu, Tumor-infiltrated activated B cells suppress liver metastasis of colorectal cancers, *Cell Rep.* 40 (2022) 111295, <https://doi.org/10.1016/j.celrep.2022.111295>.