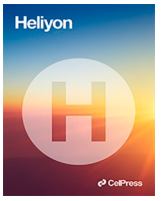
Heliyon 10 (2024) e37284

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Gender-specific microbial signatures in saliva: Unveiling the association between the oral microbiome and the pathogenesis of glioma

Hao Qin [a,*,1], Jie Liu [b,1], Yang-Yang Li [c], Ya-Lan Xu [d], Yi-Fang Yan [e,f,g,h,**]

[a] 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
[b] Department of Medical Records, Air Force Medical Center, PLA, Air Force Medical University, 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 Female Fertility Promotion, Center for Reproductive Medicine, Department of Obstetrics and Gynecology, Peking University Third Hospital, Beijing, China
[f] National Clinical Research Center for Obstetrics and Gynecology (Peking University Third Hospital), Beijing, China
[g] Key Laboratory of Assisted Reproduction (Peking University), Ministry of Education, Beijing, China
[h] Beijing Key Laboratory of Reproductive Endocrinology and Assisted Reproductive Technology, Beijing, China

ARTICLE INFO

Keywords:
Salivary flora
Glioma
Machine learning
Disease markers
Risk model

ABSTRACT

The intricate interplay between the human oral microbiome and systemic health is increasingly being recognized, particularly in the context of central nervous system pathologies such as glioblastoma. In this study, we aimed to elucidate gender-specific differences in the salivary microbiome of glioma patients by utilizing 16S rRNA sequencing data from publicly available salivary microbiome datasets. We conducted comprehensive bioinformatics analysis, encompassing quality control, noise reduction, species classification, and microbial community composition analysis at various taxonomic levels. Machine learning algorithms were employed to identify microbial signatures associated with glioma. When compared to healthy controls, our analysis revealed distinct differences in the salivary microbiota of glioma patients. Notably, the genera *Leptotrichia* and *Atopobium* exhibited significant variations in abundance between genders. *Leptotrichia* was prevalent in healthy females but exhibited a reduced abundance in female glioma patients. In contrast, *Atopobium* was more abundant in male glioma patients. These findings suggest that hormonal influences might play a role in shaping the salivary microbiome and its association with glioma. We utilized a combination of LASSO-logistic regression and random forest models for feature selection, and identified key microbial features that differentiated glioma patients from healthy controls. We developed a diagnostic model with high predictive accuracy and area under the curve and principal component analysis metrics confirmed its robustness. The analysis of microbial markers, including *Atopobium* and *Leptotrichia*, highlighted the potential of the salivary microbiota as a non-invasive biomarker for the diagnosis 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: haoqin@bjmu.edu.cn (H. Qin), yanyifangyanyf@hotmail.com (Y.-F. Yan).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2024.e37284
Received 21 March 2024; Received in revised form 29 August 2024; Accepted 30 August 2024
Available online 31 August 2024

prognosis of glioma. Our findings highlight significant gender-specific disparities in the salivary microbiome of patients with glioma, offering new insights into the pathogenesis of glioma and paving the way for innovative diagnostic and therapeutic strategies. The use of saliva as a diagnostic fluid, given its ease of collection and non-invasive nature, holds immense promise for monitoring systemic health and the trajectory of disease. Future research should focus on investigating the underlying mechanisms by which the salivary microbiome influences the development of glioma and identifying potential microbiome-targeted therapies to enhance the management of glioma.

1. Introduction

A diverse array of microorganisms, predominantly bacteria, establishes a critical ecological balance in the human oral ecosystem. The integrity of this microbial consortium is intricately linked to the etiology of multiple pathologies. Therefore, delineating the symbiotic interactions between the host and its resident oral microbiota is imperative if we are to unravel the complexities underlying the pathogenesis of oral disease [1,2]. Daily salivary output, which fluctuates between 0.75 and 1.5 L, can exert significant impact on the gastrointestinal microbiota via the ingestion of saliva and accompanying nutrients. Perturbations in the equilibrium of the oral microbiota have been implicated in a broad spectrum of systemic diseases, thereby enhancing individual susceptibility to these disorders [3]. Emerging evidence has also identified a bidirectional pathway between salivary components and the brain, indicating that substances in the saliva may reach the brain through sublingual routes. This fascinating discovery opens new avenues for utilizing saliva in the diagnosis of brain diseases [4]. Given its ease of collection and non-invasive nature, saliva is posited as a paramount diagnostic fluid with immense potential and benefits for monitoring systemic health and the trajectory of diseases. Consequently, saliva is heralded as an indispensable biomarker for the diagnosis, prognostication, and therapeutic guidance of various illnesses, including neurological disorders [5]. Glioblastoma multiforme stands as the most common primary malignant tumor of the central nervous system in adults, and accounts for approximately 58.4 % of all gliomas [6]. This disease is associated with a poor five-year survival rate and high rates of mortality, with a median survival time of 14–15 months despite the application of various treatment modalities, including surgery, chemotherapy, and radiotherapy [7]. Consequently, the discovery of effective therapeutic targets for glioblastoma is imperative to ameliorate treatment outcomes for patients. Compounding the complexity of glioblastoma multiforme treatment is the inherent heterogeneity observed between different patients and within tumors themselves, in addition to the presence of multiple types of tumor stem cells and the variability of the tumor microenvironment. Collectively, these factors constitute significant obstacles to the diagnosis and treatment of glioblastoma multiforme, severely hindering therapeutic advancements [8]. Recent studies have highlighted the promising role of salivary biomarkers for monitoring therapeutic responses and disease progression in patients with glioma. For instance, advancements in understanding the potential utility of saliva-derived exosomes before and after treatment have been reported for patients with glioma. These findings suggest that salivary biomarkers could play a pivotal role in the management of glioma [9]. Furthermore, the complex interplay between the salivary microbiota and severity of glioma has gained significant attention. For example, research by Wen and coworkers identified a profound association between the composition of the salivary microbiota and the grade of glioma. Specifically, five distinct microbial taxa were identified as markers that possessed the ability to differentiate patients with high-grade gliomas from healthy controls. Notably, the abundance of *Patescibacteria* was inversely correlated with the malignancy of gliomas, thus serving as a negative risk factor [10]. This discovery highlighted the potential of the salivary microbiota for the diagnosis and prognosis of glioma, thus emphasizing the critical role of microbial research in understanding the pathogenesis of glioma.

Recent studies have indicated significant gender differences in the incidence and prognosis of glioma, with males exhibiting a higher incidence rate and females experiencing better outcomes; these differences are probably related to differences in treatment responses that are medicated by sex hormones [11]. However, there is a significant lack of research focusing on differences in the salivary microbiome between male and female patients with glioma. This gap in knowledge was the prime motivation for our investigation into the salivary microbial communities of glioma patients, which utilized bioinformatics analysis to uncover potential sex-specific microbial signatures that could influence the pathology and treatment outcomes of glioma. Our bioinformatics approach included quality control, noise reduction, and species classification of sequencing data, followed by a comprehensive analysis of microbial community composition at various taxonomic levels. In addition, we employed machine learning algorithms for feature selection within the salivary microbiome in an attempt to identify signature microbial communities associated with glioma. Our findings, including the identification of salivary microbial taxa that exhibited significant differences in abundance between genders, highlight the complexity of the interaction between the salivary microbiome and glioma. These insights could pave the way for innovative diagnostic and therapeutic strategies by exploiting the non-invasive nature of saliva collection to enhance the management of glioma.

2. Materials and methods

2.1. Bioinformatics analysis of salivary microbial communities

Salivary microbiota profiling was conducted using 16S rRNA sequencing with BioProject accession number PRJNA750937, as

described in the previous study by Wen et al. [10]. The raw sequencing sequences were subjected to quality control using Fastp software [12]. Following this, noise reduction and other processing steps were performed on the sequences using the DADA2 plugin within the Qiime2 workflow [13]. The sequences processed with DADA2 were named amplicon sequence variants (ASVs) and stored in qza format. To obtain species classification information for each ASV, we utilized the Greengenes2 reference database (version 13.8) [14]. Based on the principles of machine learning, the ASVs were taxonomically annotated using the classify-sklearn method within the q2-feature-classifier plugin. Furthermore, the community composition of each sample was statistically analyzed at various taxonomic levels, including kingdom, phylum, class, order, family, genus, and species, thereby yielding taxonomic annotation. A dilution curve can directly reflect the appropriateness of sequencing data volume and indirectly indicates the species richness within samples. By utilizing the "qiime2R″ and "microeco" tools in R software for graphical representation, we constructed a dilution curve by simulating the resampling process. This allowed for the observation of trends in species variation and the assessment of species richness within the environment. Next, the curve was used to compare the richness of species across samples with varying volumes of sequencing data, thus serving to determine the adequacy of the sequencing data volume for each sample.

2.2. Screening for salivary microbiome signature communities, model validation and evaluation

At the genus level, original ASVs were randomly divided into training and validation sets in a 1:1 ratio. Feature selection within the training set was conducted by adopting two approaches. First, the "microeco" package (version 1.2.0) was employed to select characteristic microbial communities via a random forest model algorithm [15]. Second, the "glmnet" package (version 4.1-7) in R, utilizing a 10-fold cross-validation with the family set to binomial, was used to identify model microbial communities and coefficients based on differential microbes [16]. The feature microbes identified by the two algorithms were then compared and the intersecting components were retained as optimized feature communities for further analysis.

2.3. Identification and analysis of salivary microbial diversity by community bar charts, linear discriminant analysis (LDA) effect size (LEfSe), and network analysis

Data analysis was conducted using the Majorbio Cloud Platform (https://www.majorbio.com). Microbiological sequencing data from saliva, sourced from the BioProject database under the accession number PRJNA750937, were acquired as reported previously by Wen et al. [10]. The present study included a total of 47 saliva microbiome samples: 10 healthy males, 14 healthy females, 13 males with high-grade gliomas, and 10 females with high-grade gliomas. Profiling of the microbiota was conducted by 16S rRNA sequencing, as described by Wen et al. [10]. By utilizing LEfSe analysis (http://huttenhower.sph.harvard.edu/LEfSe) [17], we subsequently identified salivary microbial taxa that exhibited significant differential abundances across groups, from domain to the genus level ($P < 0.05$, LDA score >2.0). We also used the Kruskal-Wallis rank-sum test, a non-parametric factorial approach, to identify microbial features that exhibited significant differences in abundance. Next, we used the LDA method to evaluate the magnitude of the effect size relating to each of the identified features. Inter-species correlations were calculated by adopting a network analysis approach, which allowed us to construct a correlation network diagram. This analysis was performed at the genus taxonomic level and used Spearman's rank correlation coefficient to quantify the relationships between species at the genus level. Only correlations with an absolute value of ≥ 0.5 and a significance level of $P < 0.05$ were considered significant. In the resultant network diagrams, the thickness of the lines is proportional to the strength of the correlation between species, whereas the abundance of lines associated with a species represented the level of connectivity to other species within the network. In addition, we performed community bar chart analysis in R software (version 3.3.1) to identify dominant species across various taxonomic levels (phylum, class, order, family, genus, and species). This analysis was performed by representing each species based on its relative abundance (percentage contribution).

3. Results

3.1. PLS-DA and LEfSe analysis and the detection of variations in the salivary microbiota of patients with glioma and healthy controls

First, we employed the PLS-DA method to group samples so that we could intuitively investigate the distribution of salivary microbiota in patients with glioma compared to healthy controls. This analysis revealed the differences and similarities in the composition of the salivary microbiome (Fig. 1A). Subsequently, we conducted multi-level differential analysis using the LEfSe method (LDA threshold = 2.0) to investigate differences in the salivary microbiota between patients with glioma and healthy controls across various taxonomic levels, from domain to genus (Fig. 1B). Moreover, community composition analysis and bar graph presentations were used to illustrate trends in microbial variation between different groups, in which the y-axis represents the relative proportion of microbial genera in both groups, and the length of the bars indicates the abundance of species (Fig. 1C). These bar graphs also display LDA scores for different species, with higher LDA scores signifying a greater impact of species abundance on the inter-group differences. The results indicate that salivary bacteria, such as f_*Porphyromonadaceae* and g_*Porphyromonas*, play a key role in healthy individuals, while bacteria such as f_*Lactobacillaceae* and g_*Atopobium* are more abundant in patients with glioma (Fig. 1D). Next, we used network analysis to detect correlations between microbial species in healthy individuals and patients with glioma, and to analyze interactions among dominant species. Our analysis showed that the dominant species in healthy individuals were g_*Prevotella_7*, g_*Prevotella*, g_*Streptococcus*, g_*Porphyromonas*, and g_*Veillonella*. In patients with glioma, the dominant species were g_*Haemophilus*, g_*Gemella*, g_*Neisseria*, g_*Prevotella_7*, g_*Granulicatella*, g_*Porphyromonas*, g_*Actinomyces*, g_*Veillonella*, and g_*Prevotella* (Fig. 1E). Subsequently, we analyzed rarefaction curves by simulating the resampling process; this analysis revealed that curves became more level

as the volume of sequencing data increased. This trend suggested that the sequencing data was approaching saturation, thus providing a more accurate reflection of species richness within samples (Supplementary Figure 1).

3.2. Gender-specific differences in the salivary microbiome between patients with glioma and healthy controls, with unique changes in microbial abundance

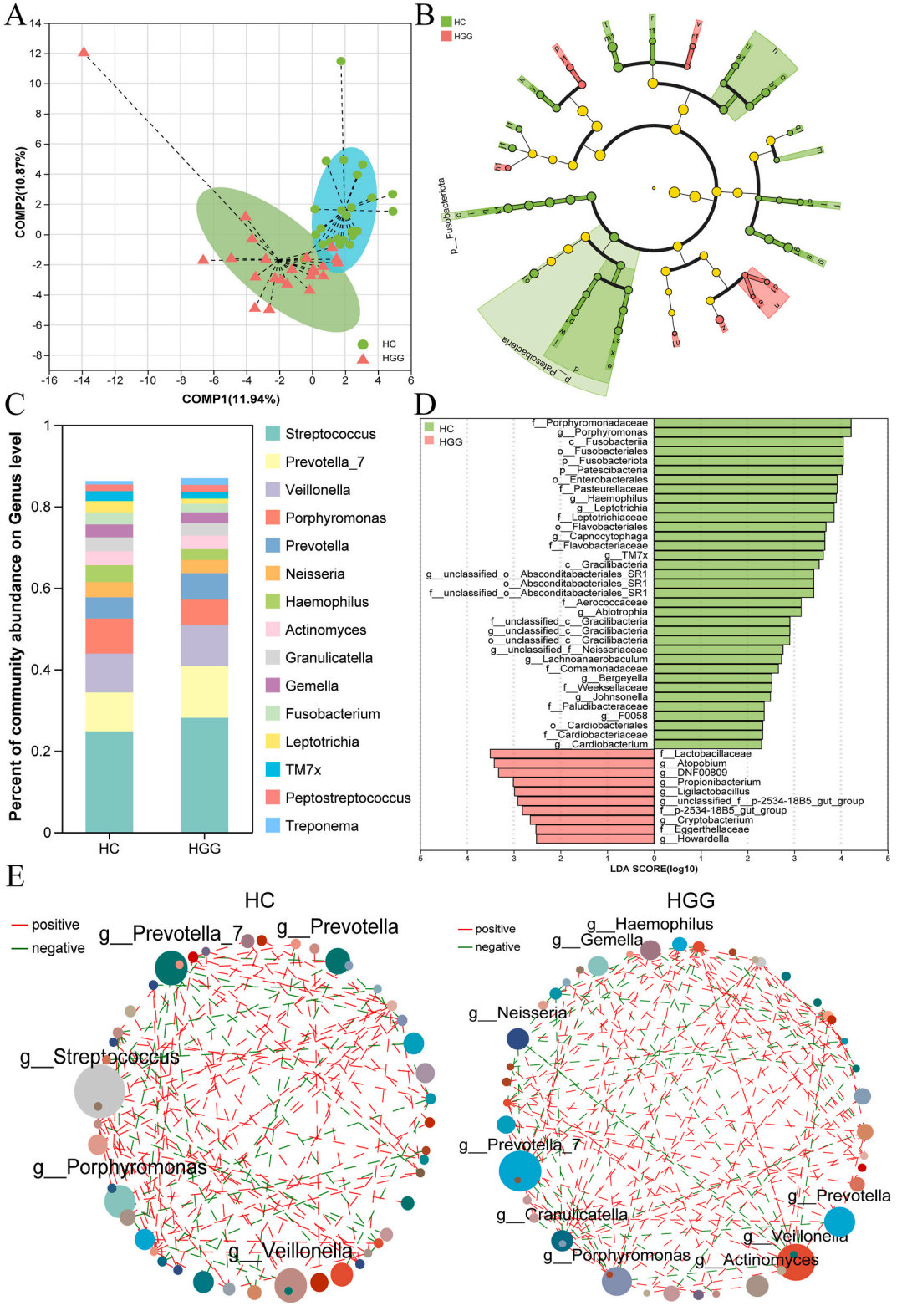
By applying a PLS-DA regression model, we detected notable differences in compositions of the salivary microbiome between patients with glioma (for both genders) and healthy controls (Fig. 2A and B). These differences were highlighted by community bar plots which depicted the top 30 abundant genera in terms of their relative abundance, with "others" representing species with lower abundance. These results indicated that the *Streptococcus*, *Prevotella_7*, and *Veillonella* genera predominated in the salivary microbiome. Specifically, the proportion of *Streptococcus* was higher in female glioma patients when compared to healthy females, whereas the *Prevotella_7* genus was more abundant in male glioma patients than in healthy males, thus indicating distinct microbial compositions in the saliva of patients with glioma when compared to healthy individuals, with unique changes in abundance detected between male and female patients (Fig. 2C and D). In addition, we investigated microbial changes at the phylum, class, order, family, and species levels. Results indicated that no significant change in the abundance of *Gammaproteobacteria* was detected between female glioma patients and healthy females, while the abundance of this microorganism was reduced in male glioma patients when compared to healthy males. In contrast, there was an increased abundance of the *Prevotellaceae* family in male glioma patients when compared to healthy males; in contrast, no significant difference was detected between female glioma patients and healthy females (Supplementary Figure 2). In addition, Circos software (version 0.67-7, http://circos.ca/) was used to generate Circos plots to illustrate the distribution of salivary microbial compositions between different genders of glioma patients and healthy individuals. These plots confirmed our previous bar plots, showing similar compositions of dominant microbes across different groups. *Streptococcus* was the dominant genus, followed by *Prevotella_7* as the second most dominant. In the salivary microbiome of healthy males and male glioma patients, the abundance of *Prevotella_7* increased from 36 % to 64 %, but remained consistent at 48 % and 52 % in healthy females and female glioma patients, respectively (Fig. 2E and F). Similar patterns were detected across various taxonomic levels, including phylum, class, order, family, and species. Taking the family level as an example, the relative abundance of *Prevotellaceae* in the salivary microbiota of healthy women was 48 %, compared to 52 % in women with glioma. In contrast, the relative abundance of *Prevotellaceae* was 39 % in healthy men but was more abundant (61 %) in men with glioma (Supplementary Figure 3). In addition, our analyses identified gender-specific variations in the salivary microbiota of healthy individuals. These differences involved specific bacterial groups, including *Eubacterium nodatum*, *Solobacterium*, and certain unclassified members within the family (Supplementary Figure 4). These findings demonstrated significant differences in the compositions of the salivary microbiome between male and female glioma patients when compared to healthy individuals, thus indicating gender-specific changes in microbial abundance.

3.3. The impact of gender on disparities in the salivary microbiota of patients with glioma, as determined by intergroup significance testing

Next, we investigated how gender influenced disparities in the salivary microbiota between glioma patients and healthy individuals by performing rigorous intergroup significance testing. This analysis aimed to identify bacteria at the genus level that exhibited significant variances, thereby revealing the specific mechanisms by which gender exerts impact on microbial composition. Our findings indicated a significant reduction in the abundance of *Leptotrichia*, *Selenomonas*, *Abiotrophia*, and *Lachnoanaerobaculum* genera in female patients with glioma when compared to the healthy control group, a phenomenon that may relate to the unique physiological or immune responses specific to female patients with glioma (Fig. 3A and B). Conversely, in male patients with glioma, we detected a reduction in the abundance of *Haemophilus*, *Granulicatella*, and *Capnocytophaga* genera, with a significant increase in *Atopobium*, potentially reflecting distinct adjustments in the structure of the salivary microbiota in male glioma patients (Fig. 3C and D). Bar chart analysis further revealed differences in LDA scores among species within the salivary microbiomes (LDA threshold = 2.0). These findings indicated a marked predominance of bacterial species such as g_*Cryptobacterium* in female patients with glioma. In contrast, species such as p_*Patescibacteria* and c_*Fusobacteriia* were more prevalent in the salivary microbiomes of healthy females. We also found that c_*Coriobacteriia* and o_*Coriobacteriales* played a pivotal role in the salivary microbiomes of male patients with glioma. Furthermore, the presence of o_*Enterobacterales* and f_*Pasteurellaceae* was more significant in the salivary microbiomes of healthy males (Supplementary Figure 5). Next, network analysis was used to elucidate the complex correlations between the salivary microbiota of healthy individuals of different genders and those of patients with glioma, thus providing comprehensive insights into microbial interactions in the context of glioma (Supplementary Figure 6). Significant differences in microbial populations could have potential associations with the onset and progression of glioma, thus suggesting that gender plays a crucial role in the variation of salivary microbiome composition in patients with glioma.

3.4. Selection and analysis of key microbial features for the construction and evaluation of a diagnostic model

Next, we performed a systematic feature selection process using two sequential algorithms: LASSO-logistic regression and random forest analysis. First, we used the "glmnet" package in R with 10-fold cross-validation under a binomial family setting to determine model microbial communities and coefficients based on differential microbes. The LASSO-logistic regression algorithm successfully identified eight significant microbial features (Fig. 4A and B). Next, we used the "microeco" package to select characteristic microbial communities using a random forest model; this identified twelve additional key microbial features. The left panel of Fig. 4C shows a MeanDecreaseGini plot, which was used to analyze the contribution of each species to model classification, with higher values



(caption on next page)

Fig. 1. Comprehensive analysis and visualization of the salivary microbiota in patients with glioma compared to healthy controls, illustrating genus-level proportions, phylogenetic relationships, and interaction networks. (A) PLS-DA analysis of salivary microbiota characteristics in samples grouped by category (HC, healthy control; HGG, high-grade gliomas). (B) Evolutionary relationships among the salivary microbiota of patients with glioma and healthy controls, as determined by LEfSe analysis. The selected microbial representatives, comprising d: class *Gracilibacteria*; n: family *Eggerthellaceae*; h: order *Flavobacteriales*; k: family *Aerococcaceae*; p: family *Lactobacillaceae*; t: family *Porphyromonadaceae*; and r: family *Paludibacteraceae*. (C) Proportional distribution of genus-level salivary microbiota within samples. (D) Statistical depiction of salivary microbiota LDA scores featured in the results arising from LEfSe analysis. (E) Interaction networks among different genera, with positive and negative associations represented in red and green, respectively. The size of each node reflects the abundance of the species, and the thickness of the lines indicates the strength of the correlation, with thicker lines denoting higher levels of interconnectivity among species.)

indicating better classification performance. The right panel shows a box plot illustrating differences in relative abundance. The features identified by both algorithms were compared, and overlapping features were selected as optimal microbial communities for further analysis. Subsequently, we performed rigorous Venn diagram analysis to extract intersections of the microbial features selected by the two algorithms. This meticulous process ultimately identified five key microbial features that were deemed to play pivotal roles in the construction of a diagnostic model. The formula used to calculate the Risk Score within the diagnostic model was derived as follows: Risk Score = g_*Atopobium* × 0.61733369 + g_*Bacteroides* × 2.93668876 + g_*Capnocytophaga* × (−1.79325740) + g_*Leptotrichia* × (−0.07011431) + g_*Cardiobacterium* × (−0.64509608). This formula represents a critical component of the diagnostic model and provides a quantitative assessment of risk based on the presence and abundance of specific microbial features. To rigorously assess the reliability and stability of our diagnostic model, we conducted several key analyses. First, we calculated the area under the curve (AUC) metric, as depicted by Fig. 4D. The AUC is a critical measure used to evaluate the performance of binary classification models, and provides an indicator of a model's ability to distinguish between different classes. In addition, we used principal component analysis (PCA) to provide a more rigorous analysis of data structure (Fig. 4E). PCA was used to visualize the separation of data points in a multidimensional space, thereby validating the effectiveness of the model. Collectively, these analyses highlight the robustness and predictive power of our diagnostic model, indicating its potential utility for evaluating the risk of glioblastoma.

4. Discussion

The advancing field of microbiome research highlights the intricate interplay between microbial communities and human health. Recent investigations have highlighted the potential impact of the oral microbiota on a variety of central nervous system diseases, including Alzheimer's disease [18] and ischemic stroke [19]. In oncological research, dysbiosis of the oral microbiome has been closely associated with the development of oral squamous cell carcinoma [20], head and neck cancer [21], and glioma [10]. Oral microbes can influence the gut microbiota of a host via the oral-gut microbiome axis, thus leading to dysbiosis and the onset of disease [22]. A previous study identified imbalances in the gut microbiome of a mouse model of glioma and patients with glioma when compared to healthy controls, including alterations in β-diversity [23]. Our comprehensive analysis of the saliva microbiome of patients with glioma, focusing particularly on gender-specific differences, is critically important if we are to enhance our understanding of the pathogenesis of glioma and its potential diagnostic and therapeutic implications. In our study, we elucidated the impact of the salivary microbiome on the risk and progression of glioma, highlighting the intricate interplay between microbial dysbiosis and alterations in host-microbiome interactions. Notably, the presence of gender-specific microbial signatures highlighted the potentially pivotal role of biological gender in modulating these interactions. Given the observed disparities in the incidence and prognosis of glioma between genders, we hypothesize that this variation may partly be attributed to differences in the composition of the salivary microbiome between males and females. These findings are in alignment with recent studies indicating that male patients with brain gliomas are more susceptible to this disease and have poorer outcomes when compared to their female counterparts, exhibiting distinct gender disparities [24]. Moreover, the ability of the gut microbiome to regulate hormone levels by altering reabsorption rates, as well as modulate endocrine signaling through the biotransformation of endogenous hormones [25], highlights the significance of this research. Our study accentuates the potential value of gender-dependent modulation by the salivary microbiome on the risk and therapeutic response to brain glioma disease.

Moreover, the findings of this study identified significant differences in the salivary microbiome between male and female patients with gliomas when compared to healthy individuals. Specifically, the predominant differential bacterial genera in male patients with glioma included *Haemophilus*, *Granulicatella*, *Capnocytophaga*, and *Atopobium*. In female patients with glioma, the genera predominantly consisted of *Leptotrichia*, *Selenomonas*, *Abiotrophia*, and *Lachnoanaerobaculum*. Notably, *Leptotrichia* has been reported to be the most abundant bacterial genus in the oral gingival plaques of pregnant women, coinciding with the period of elevated estrogen levels during human pregnancy [26,27]. Our analysis discovered a significant increase in *Leptotrichia* in the healthy female population when compared to those with gliomas, leading us to hypothesize that a positive correlation between *Leptotrichia* and estrogen levels might be associated with the incidence of gliomas in females. The elevation in estrogen levels mediated by *Leptotrichia* may potentially reduce the incidence rate of gliomas in women; however, whether this association is causal remains unclear. Furthermore, we identified a significantly higher abundance of the *Atopobium* genus in the salivary microbiome of male patients with glioma when compared to healthy controls. Intriguingly, the abundance of *Atopobium* has been shown to differ significantly in vaginal samples from patients with endometriosis when compared to healthy individuals, with almost no presence in the former population [28], a condition that is characterized by high estrogen dependency and secretion as its endocrine hallmark [29]. This suggests that the antagonistic relationship between *Atopobium* and estrogen might explain the higher salivary abundance of *Atopobium* in male patients with glioma when compared to healthy individuals. The observed differences in salivary microbial abundance provide us with valuable insights

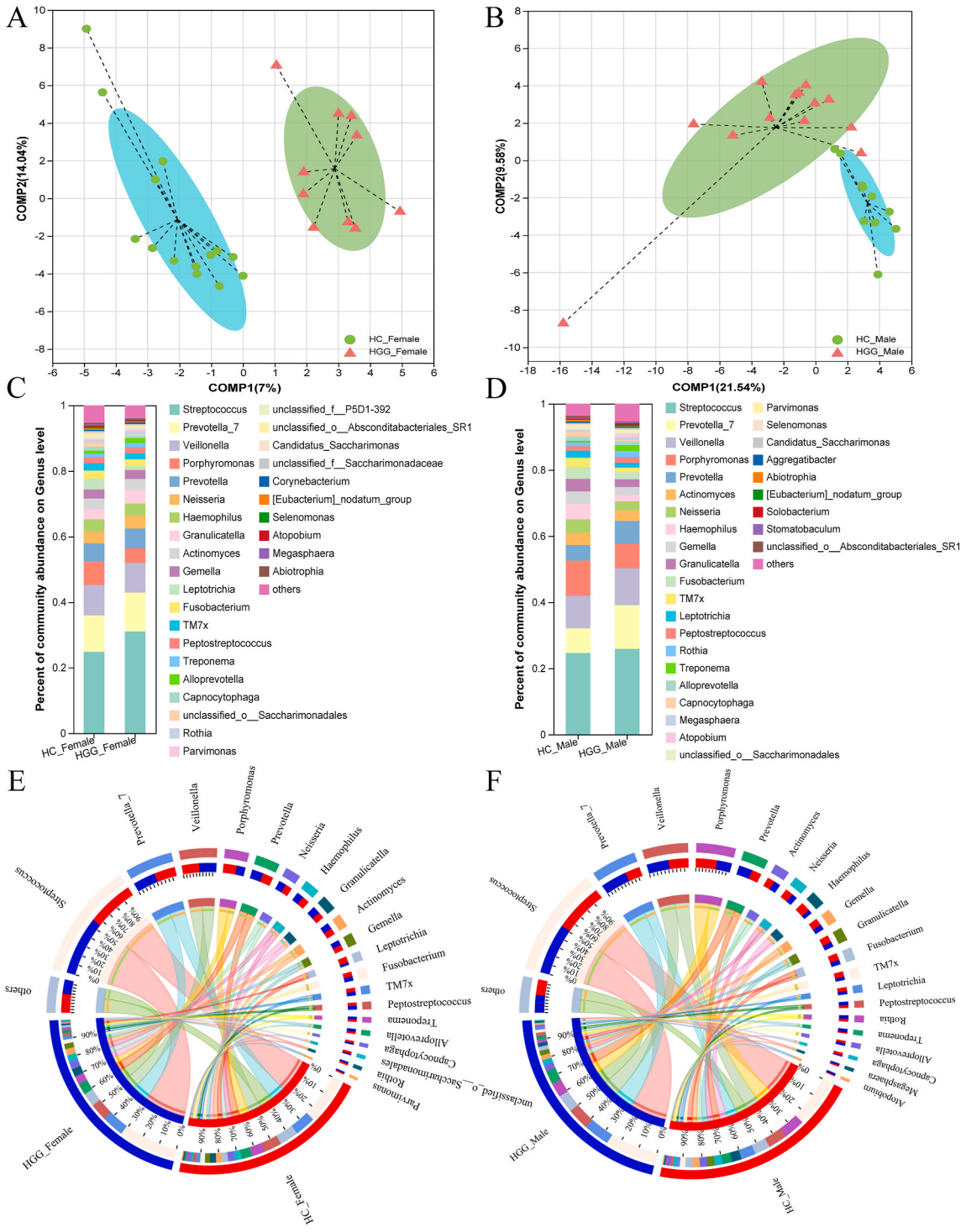
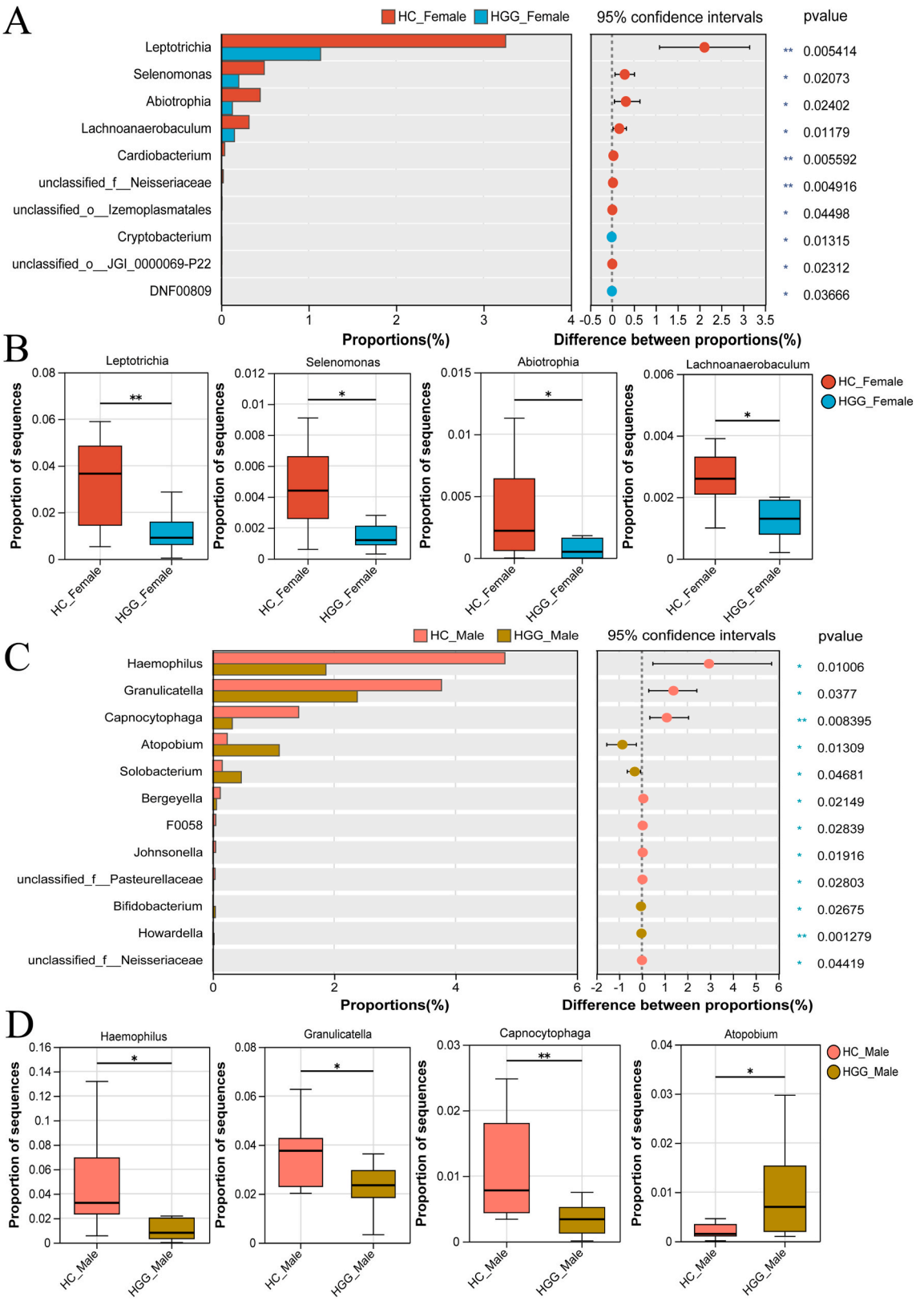


Fig. 2. Gender-specific salivary microbiome profiles in glioma patients in comparison to healthy individuals. PLS-DA analyses contrasting salivary microbiome samples between healthy females and glioma patients (A), and between healthy males and glioma patients (B). Community bar charts depicting the top 30 microbial species by abundance in the saliva of healthy females in comparison to patients with glioma (C), and in healthy males in comparison with patients with glioma, with species abundance depicted along the y-axis against sample names on the x-axis (D). Circos plots illustrating the distribution of microbial species in the saliva across genders in both healthy individuals and glioma patients, in which the inner ribbon colors represent different groups and their lengths indicate the proportion of species distribution within the samples (E–F).



(caption on next page)

Fig. 3. Bar graphs depicting gender-specific differences in salivary microbiome abundance between patients with glioma and healthy individuals. These bar graphs compare the relative abundance of salivary microbial genera between female patients with glioma and healthy females, represented as percentages of relative abundance (A). Statistical analysis of microbial differences within the female cohort, including g_*Leptotrichia*, g_*Selenomonas*, g_*Abiotrophia*, and g_*Lachnoanaerobaculum* (B). Bar chart comparing the relative abundance of salivary microbial genera between male patients with glioma and healthy male controls, expressed as a percentage (C). Statistical analysis of microbial differences illustrating disparities in the relative abundance of salivary microbial genera between male patients with glioma and healthy males, emphasizing the importance of g_*Haemophilus*, g_*Granulicatella*, g_*Capnocytophaga*, and g_*Atopobium* (D). $^{*}P < 0.05$ and $^{**}P < 0.01$.

into the etiology of glioma. Estrogen-related bacterial genera may directly or indirectly modulate the microenvironment of the brain, thus influencing the development and progression of glioma. Future identification of specific microbial characteristics will further elucidate the potential mechanisms by which the oral microbiome may impact glioma biology.

Recent applications of machine learning algorithms, such as PLS-DA and LEfSe, have facilitated detailed analyses of the salivary microbiome, uncovering significant disparities between patients with glioma and healthy individuals. These computational approaches have been instrumental in identifying key salivary microbial markers for glioma, employing algorithms such as LASSO-logistic regression and random forest analyses. Notably, LASSO-logistic regression identified eight characteristic microbes, while random forest revealed twelve. The diagnostic model was subsequently refined to include only the overlapping microbial markers, thus marking the clear advantages of utilizing algorithms for the selection of diagnostic or prognostic factors in recent studies [30–32]. Moreover, the high predictive accuracy of this diagnostic model, validated by PCA, highlights its potential utility in clinical diagnostics, offering a non-invasive, saliva-based tool for the detection of glioma. This enhances our comprehension of the microbial landscape associated with glioma. Our findings also highlight the potential therapeutic implications of targeting the oral microbiome when treating glioma. Given the observed microbial dysbiosis in patients with glioma, strategies aimed at modulating the oral microbiome could represent novel therapeutic avenues. For instance, the application of probiotics to modulate the microbial equilibrium and influence glioma-associated biological activities has already been investigated. One previous study highlighted the capacity of a probiotic blend to suppress intracranial tumor growth in a mouse model of glioma by altering the gut microbiome, facilitated by the restoration of intestinal barrier integrity [33]. Furthermore, our discovery of gender-specific microbial features in the salivary microbiome indicates significant gender differences in glioma, suggesting that personalized microbiome interventions that consider individual patient characteristics (including gender) could enhance therapeutic outcomes. Future research should aim to elucidate the mechanisms underlying observed gender differences in the impact of the salivary microbiome on glioma and investigate the therapeutic potential of modulating the microbiome in the treatment of glioma.

5. Conclusion

Our comprehensive analysis demonstrates the critical role of the salivary microbiome in the pathogenesis of glioma and potential diagnostic applications, particularly highlighting gender-specific differences. Our research successfully identified distinctive microbial taxa in the saliva of male and female patients with glioma, distinguishing them from healthy controls. Notably, genera such as *Atopobium* and *Leptotrichia* showed significant variations in abundance linked to gender, thus suggesting their influence on glioma susceptibility and hormone interactions. These findings emphasize the potential of saliva as a non-invasive diagnostic medium that is capable of reflecting underlying disease processes via microbial signatures. The integration of machine learning techniques has further refined the identification of key microbial markers, thus enhancing the diagnostic precision for glioma. These findings emphasize the promise of saliva-based diagnostics in clinical settings, offering insights into disease progression and response to treatment. Future research should investigate the specific mechanisms by which the salivary microbiota can influence the development of glioma and investigate targeted therapeutic interventions that could modulate these microbial communities.

Ethics approval and consent to participate

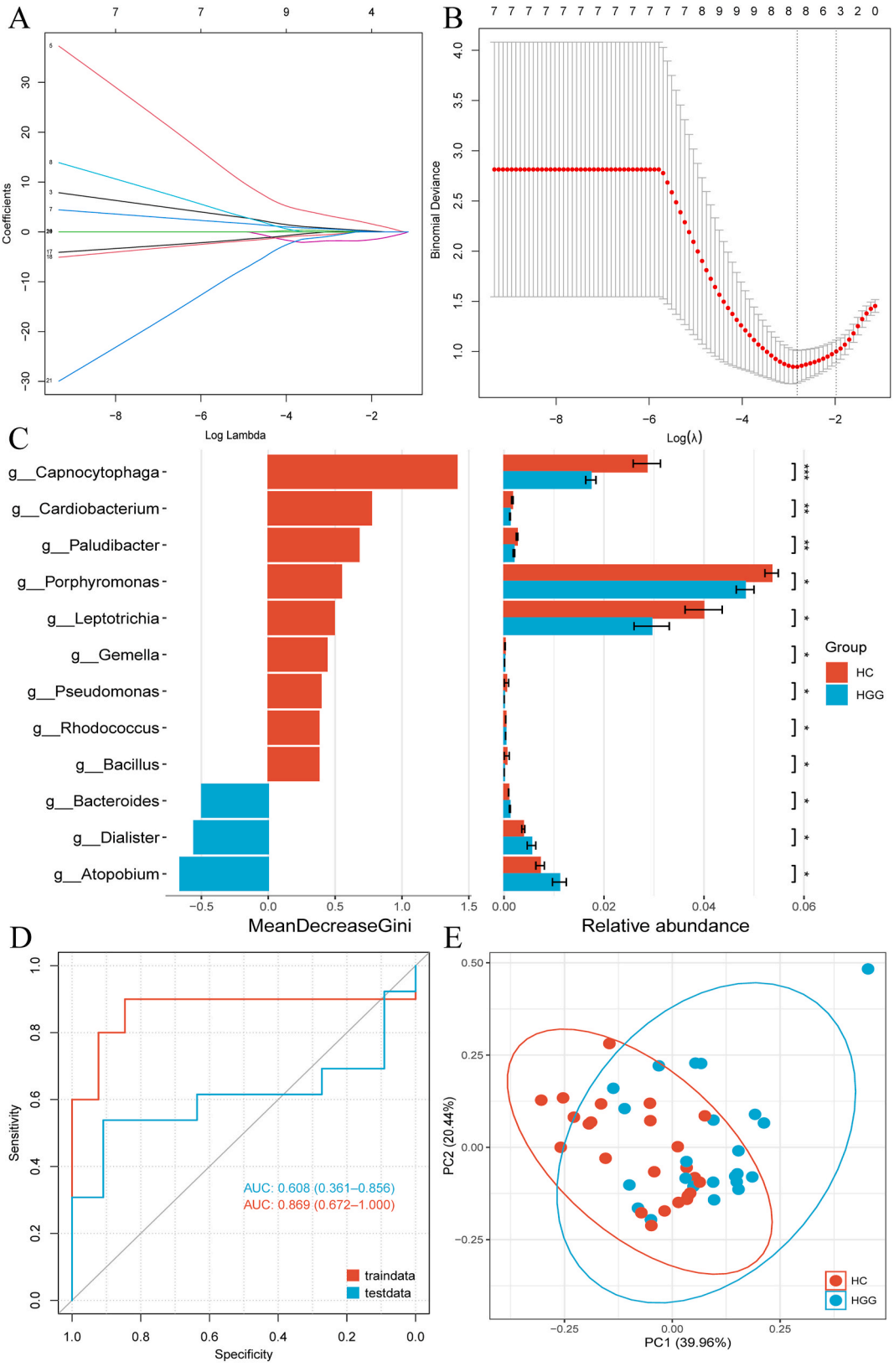
Not applicable.

Funding

This work was supported by Beijing Natural Science Foundation (L232075), the National Natural Science Foundation of China (Nos. 82304151 and 32100647), Supported by the Fundamental Research Funds for the Central Universities (3332023033), Funding from National Clinical Research Center for Obstetrics and Gynecology (Peking University Third Hospital) (No. BYSYSZKF2023014), Key Clinical Project of Peking University Third Hospital (No. BYSYZD2023036), and the Cancer Hospital of Chinese Academy of Medical Sciences-Shenzhen Hospital Cooperation Fund (CFA202202023).

Data availability statement

The data used to support the findings of this study are available from the corresponding authors on reasonable request.



(caption on next page)

Fig. 4. Enhanced model accuracy derived from LASSO regularization and RandomForests feature selection, as demonstrated by advanced AUC and PCA metrics. The coefficient of variation and distribution under LASSO regularization, highlighting the precision of our methodology (A). The optimized selection of lambda.min based on likelihood deviation within the LASSO coefficient distribution (B). The MeanDecreaseGini chart (left), generated by RandomForests feature selection analysis, was used to quantify the influence of various species on the classification accuracy of the model, in which higher MeanDecreaseGini values correlate with enhanced performance. Conversely, the right section of (C) features a boxplot showcasing relative abundance differences, thus offering insights into species impact. The models exceptional AUC (D) and PCA (E) outcomes, confirmed the predictive reliability and dimensional reduction capability of the model, thus emphasizing the efficacy of the method.

CRediT authorship contribution statement

Hao Qin: Writing – original draft, Supervision, Methodology, Funding acquisition, Conceptualization. **Jie Liu:** Writing – review & editing, Methodology, Investigation, Data curation. **Yang-Yang Li:** Investigation. **Ya-Lan Xu:** Writing – review & editing. **Yi-Fang Yan:** Writing – original draft, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e37284.

References

[1] W.G. Wade, Characterisation of the human oral microbiome, J. Oral Biosci. 55 (2013) 143–148, https://doi.org/10.1016/j.job.2013.06.001.

[2] B.J.F. Keijser, E. Zaura, S.M. Huse, J.M.B.M. van der Vossen, F.H.J. Schuren, R.C. Montijn, J.M. ten Cate, W. Crielaard, Pyrosequencing analysis of the Oral Microflora of healthy adults, J. Dent. Res. 87 (2008) 1016–1020, https://doi.org/10.1177/154405910808701104.

[3] L. Sedghi, V. DiMassa, A. Harrington, S.V. Lynch, Y.L. Kapila, The oral microbiome: role of key organisms and complex networks in oral health and disease, Periodontol. 2000 87 (2021) 107–131, https://doi.org/10.1111/prd.12393.

[4] C. Zürcher, C. Humpel, Saliva: a challenging human fluid to diagnose brain disorders with a focus on Alzheimer's disease, Neural regeneration research 18 (2023) 2606–2610, https://doi.org/10.4103/1673-5374.373675.

[5] P. Kumar, S. Gupta, B.C. Das, Saliva as a potential non-invasive liquid biopsy for early and easy diagnosis/prognosis of head and neck cancer, Translational oncology 40 (2024) 101827, https://doi.org/10.1016/j.tranon.2023.101827.

[6] Q.T. Ostrom, G. Cioffi, K. Waite, C. Kruchko, J.S. Barnholtz-Sloan, CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2014-2018, Neuro Oncol. 23 (2021) iii1–iii105, https://doi.org/10.1093/neuonc/noab200.

[7] J.H. Sampson, M.D. Gunn, P.E. Fecci, D.M. Ashley, Brain immunology and immunotherapy in brain tumours, Nat. Rev. Cancer 20 (2020) 12–25, https://doi.org/10.1038/s41568-019-0224-7.

[8] D. Eisenbarth, Y.A. Wang, Glioblastoma heterogeneity at single cell resolution, Oncogene 42 (2023) 2155–2165, https://doi.org/10.1038/s41388-023-02738-y.

[9] J. Müller Bark, L. Trevisan França de Lima, X. Zhang, D. Broszczak, P.J. Leo, R.L. Jeffree, B. Chua, B.W. Day, C. Punyadeera, Proteome profiling of salivary small extracellular vesicles in glioblastoma patients, Cancer 129 (2023) 2836–2847, https://doi.org/10.1002/cncr.34888.

[10] Y. Wen, L. Feng, H. Wang, H. Zhou, Q. Li, W. Zhang, M. Wang, Y. Li, X. Luan, Z. Jiang, L. Chen, J. Zhou, Association between oral microbiota and human brain glioma grade: a case-control study, Front. Microbiol. 12 (2021) 746568, https://doi.org/10.3389/fmicb.2021.746568.

[11] A. Carrano, J.J. Juarez, D. Incontri, A. Ibarra, H. Guerrero Cazares, Sex-specific differences in glioblastoma, Cells 10 (2021), https://doi.org/10.3390/cells10071783.

[12] S. Chen, Y. Zhou, Y. Chen, J. Gu, fastp: an ultra-fast all-in-one FASTQ preprocessor, Bioinformatics 34 (2018) i884–i890, https://doi.org/10.1093/bioinformatics/bty560.

[13] E. Bolyen, J.R. Rideout, M.R. Dillon, N.A. Bokulich, C.C. Abnet, G.A. Al-Ghalith, H. Alexander, E.J. Alm, M. Arumugam, F. Asnicar, Y. Bai, J.E. Bisanz, K. Bittinger, A. Brejnrod, C.J. Brislawn, C.T. Brown, B.J. Callahan, A.M. Caraballo-Rodríguez, J. Chase, E.K. Cope, R. Da Silva, C. Diener, P.C. Dorrestein, G. M. Douglas, D.M. Durall, C. Duvallet, C.F. Edwardson, M. Ernst, M. Estaki, J. Fouquier, J.M. Gauglitz, S.M. Gibbons, D.L. Gibson, A. Gonzalez, K. Gorlick, J. Guo, B. Hillmann, S. Holmes, H. Holste, C. Huttenhower, G.A. Huttley, S. Janssen, A.K. Jarmusch, L. Jiang, B.D. Kaehler, K.B. Kang, C.R. Keefe, P. Keim, S.T. Kelley, D. Knights, I. Koester, T. Kosciolek, J. Kreps, M.G.I. Langille, J. Lee, R. Ley, Y.X. Liu, E. Loftfield, C. Lozupone, M. Maher, C. Marotz, B.D. Martin, D. McDonald, L. J. McIver, A.V. Melnik, J.L. Metcalf, S.C. Morgan, J.T. Morton, A.T. Naimey, J.A. Navas-Molina, L.F. Nothias, S.B. Orchanian, T. Pearson, S.L. Peoples, D. Petras, M.L. Preuss, E. Pruesse, L.B. Rasmussen, A. Rivers, M.S. Robeson 2nd, P. Rosenthal, N. Segata, M. Shaffer, A. Shiffer, R. Sinha, S.J. Song, J.R. Spear, A. D. Swafford, L.R. Thompson, P.J. Torres, P. Trinh, A. Tripathi, P.J. Turnbaugh, S. Ul-Hasan, J.J.J. van der Hooft, F. Vargas, Y. Vázquez-Baeza, E. Vogtmann, M. von Hippel, W. Walters, Y. Wan, M. Wang, J. Warren, K.C. Weber, C.H.D. Williamson, A.D. Willis, Z.Z. Xu, J.R. Zaneveld, Y. Zhang, Q. Zhu, R. Knight, J. G. Caporaso, Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2, Nat. Biotechnol. 37 (2019) 852–857, https://doi.org/10.1038/s41587-019-0209-9.

[14] D. McDonald, Y. Jiang, M. Balaban, K. Cantrell, Q. Zhu, A. Gonzalez, J.T. Morton, G. Nicolaou, D.H. Parks, S.M. Karst, M. Albertsen, P. Hugenholtz, T. DeSantis, S.J. Song, A. Bartko, A.S. Havulinna, P. Jousilahti, S. Cheng, M. Inouye, T. Niiranen, M. Jain, V. Salomaa, L. Lahti, S. Mirarab, R. Knight, Greengenes2 unifies microbial data in a single reference tree, Nat. Biotechnol. (2023), https://doi.org/10.1038/s41587-023-01845-1.

[15] C. Liu, Y. Cui, X. Li, M. Yao, microeco: an R package for data mining in microbial community ecology, FEMS Microbiol. Ecol. 97 (2021), https://doi.org/10.1093/femsec/fiaa255.

[16] B. Jiang, P. Sun, B. Luo, GLMNet: graph learning-matching convolutional networks for feature matching, Pattern Recogn. 121 (2022) 108167, https://doi.org/10.1016/j.patcog.2021.108167.

[17] N. Segata, J. Izard, L. Waldron, D. Gevers, L. Miropolsky, W.S. Garrett, C. Huttenhower, Metagenomic biomarker discovery and explanation, Genome Biol. 12 (2011) R60, https://doi.org/10.1186/gb-2011-12-6-r60.

[18] J. Wan, H. Fan, Oral microbiome and Alzheimer's disease, Microorganisms 11 (2023) 2550, https://doi.org/10.3390/microorganisms11102550.

[19] W. Sun, S. Huang, X. Yang, Y. Luo, L. Liu, D. Wu, The oral microbiome of patients with ischemic stroke predicts their severity and prognosis, Front. Immunol. 14 (2023) 1171898, https://doi.org/10.3389/fimmu.2023.1171898.
[20] L. Zhang, Y. Liu, H.J. Zheng, C.P. Zhang, The oral microbiota may have influence on oral cancer, Front. Cell. Infect. Microbiol. 9 (2019) 476, https://doi.org/10.3389/fcimb.2019.00476.
[21] Z. Wu, Y. Han, Y. Wan, X. Hua, S.S. Chill, K. Teshome, W. Zhou, J. Liu, D. Wu, A. Hutchinson, K. Jones, C.L. Dagnall, B.D. Hicks, L. Liao, H. Hallen-Adams, J. Shi, C.C. Abnet, R. Sinha, A. Chaturvedi, E. Vogtmann, Oral microbiome and risk of incident head and neck cancer: a nested case-control study, Oral Oncol. 137 (2023) 106305, https://doi.org/10.1016/j.oraloncology.2022.106305.
[22] Y. Lu, Z. Li, X. Peng, Regulatory effects of oral microbe on intestinal microbiota and the illness, Front. Cell. Infect. Microbiol. 13 (2023) 1093967, https://doi.org/10.3389/fcimb.2023.1093967.
[23] A. Patrizz, A. Dono, S. Zorofchian, G. Hines, T. Takayasu, N. Husein, Y. Otani, O. Arevalo, H.A. Choi, J. Savarraj, N. Tandon, B.P. Ganesh, B. Kaur, L. D. McCullough, L.Y. Ballester, Y. Esquenazi, Glioma and temozolomide induced alterations in gut microbiome, Sci. Rep. 10 (2020) 21002, https://doi.org/10.1038/s41598-020-77919-w.
[24] F. Tournant, Immunological sex differences in glioblastoma, Nat. Can. (Ott.) 4 (2023), https://doi.org/10.1038/s43018-023-00660-6, 1636-1636.
[25] L.M. Cox, H. Abou-El-Hassan, A.H. Maghzi, J. Vincentini, H.L. Weiner, The sex-specific interaction of the microbiome in neurodegenerative diseases, Brain Res. 1724 (2019) 146385, https://doi.org/10.1016/j.brainres.2019.146385.
[26] J.C. Condon, C. Kyathanahalli, P. Anamthathmakula, P. Jeyasuria, Estrogen/estrogen receptor action and the pregnant myometrium, Current Opinion in Physiology 13 (2020) 135–140, https://doi.org/10.1016/j.cophys.2019.10.017.
[27] Y. Zhang, Z. Wu, L. Li, X. Wang, W. Fan, J. Zhao, Characterizing the supragingival microbiome of healthy pregnant women, Front. Cell. Infect. Microbiol. 12 (2022), https://doi.org/10.3389/fcimb.2022.1016523.
[28] B. Ata, S. Yildiz, E. Turkgeldi, V.P. Brocal, E.C. Dinleyici, A. Moya, B. Urman, The endobiota study: comparison of vaginal, cervical and gut microbiota between women with stage 3/4 endometriosis and healthy controls, Sci. Rep. (2019) 2204, https://doi.org/10.1038/s41598-019-39700-6.
[29] E. Chantalat, M.C. Valera, C. Vaysse, E. Noirrit, M. Rusidze, A. Weyl, K. Vergriete, E. Buscail, P. Lluel, C. Fontaine, J.F. Arnal, F. Lenfant, Estrogen receptors and endometriosis, Int. J. Mol. Sci. 21 (2020), https://doi.org/10.3390/ijms21082815.
[30] Z. Wu, H. Chen, S. Ke, L. Mo, M. Qiu, G. Zhu, W. Zhu, L. Liu, Identifying potential biomarkers of idiopathic pulmonary fibrosis through machine learning analysis, Sci. Rep. 13 (2023) 16559, https://doi.org/10.1038/s41598-023-43834-z.
[31] Z. Li, Y. Qin, X. Liu, J. Chen, A. Tang, S. Yan, G. Zhang, Identification of predictors for neurological outcome after cardiac arrest in peripheral blood mononuclear cells through integrated bioinformatics analysis and machine learning, Funct. Integr. Genom. 23 (2023) 83, https://doi.org/10.1007/s10142-023-01016-0.
[32] D. Lin, W. Zhao, J. Yang, H. Wang, H. Zhang, Integrative analysis of biomarkers and mechanisms in adamantinomatous craniopharyngioma, Front. Genet. 13 (2022), https://doi.org/10.3389/fgene.2022.830793.
[33] L. Wang, S. Li, H. Fan, M. Han, J. Xie, J. Du, F. Peng, Bifidobacterium lactis combined with Lactobacillus plantarum inhibit glioma growth in mice through modulating PI3K/AKT pathway and gut microbiota, Front. Microbiol. 13 (2022) 986837, https://doi.org/10.3389/fmicb.2022.986837.