

Doctoral Thesis

Microbiota in Human Diseases

Jaewoong Lee

Department of Biomedical Engineering

Ulsan National Institute of Science and Technology

2024

Microbiota in Human Diseases

Jaewoong Lee

Department of Biomedical Engineering

Ulsan National Institute of Science and Technology



CHURCH OF THE FLYING SPAGHETTI MONSTER

February 09, 2021

Letter of Good Standing

Dear Sir or Madam:

I am pleased to verify that _____

JAEWOONG LEE

is an ordained minister of the Church of the Flying Spaghetti Monster and recognized
within our organization as a member in good standing.

We hereby consent to this minister performing ceremonies and request that they are
granted all privileges and respect appropriate to a spiritual leader.

Any questions can be directed to the undersigned.

Representative,
Church of the Flying Spaghetti Monster
Bobby Henderson



CHURCH OF THE FLYING SPAGHETTI MONSTER

February 09, 2021

Letter of Good Standing

Dear Sir or Madam:

I am pleased to verify that _____

JAEWOONG LEE

is an ordained minister of the Church of the Flying Spaghetti Monster and recognized
within our organization as a member in good standing.

We hereby consent to this minister performing ceremonies and request that they are
granted all privileges and respect appropriate to a spiritual leader.

Any questions can be directed to the undersigned.

Representative,
Church of the Flying Spaghetti Monster
Bobby Henderson

Abstract

(Microbiome)

(PTB) Section 2 introduces...

(Periodontitis) Section 3 describes...

(Lung)

(Conclusion)

This doctoral dissertation is an addition based on the following papers that the author has already published:

- Hong, Y. M., Lee, J., Cho, D. H., Jeon, J. H., Kang, J., Kim, M. G., ... & Kim, J. K. (2023). Predicting preterm birth using machine learning techniques in oral microbiome. *Scientific Reports*, 13(1), 21105.

Contents

1	Introduction	2
2	Predicting preterm birth using random forest classifier in salivary microbiome	3
2.1	Introduction	3
2.2	Materials and methods	5
2.2.1	Study design and study participants	5
2.2.2	Clinical data collection and grouping	5
2.2.3	Salivary microbiome sample collection	5
2.2.4	16s rRNA gene sequencing	5
2.2.5	Bioinformatics analysis	5
2.3	Results	7
2.3.1	Overview of clinical information	7
2.3.2	Comparison of salivary microbiomes composition	7
2.3.3	Random forest classification to predict PTB risk	7
2.4	Discussion	15
3	Random forest prediction model for periodontitis statuses based on the salivary microbiomes	17
3.1	Introduction	17
3.2	Materials and methods	18
3.2.1	Study participants enrollment	18
3.2.2	Periodontal clinical parameter diagnosis	18
3.2.3	Saliva sampling and DNA extraction procedure	18
3.2.4	Bioinformatics analysis	18
3.3	Results	19
3.3.1	Summary of study subjects and sequencing data	19

3.3.2	Diversity indices reveal differences in diversity among the periodontal statuses	19
3.3.3	Differentially abundant taxa among healthy and different periodontitis stages and their correlation	19
3.3.4	Classification of periodontitis stages by random forest models	19
3.4	Discussion	40
4	Lung microbiome	41
4.1	Introduction	41
4.2	Materials and methods	42
4.3	Results	43
4.3.1	Discussion	44
5	Conclusion	45
	References	46
	Acknowledgments	51

List of Figures

1	DAT volcano plot	9
2	Salivary microbiome compositions over DAT	10
3	Random forest-based PTB prediction model	11
4	Diversity indices	12
5	PROM-related DAT	13
6	Validation of random forest-based PTB prediction model	14
7	Diversity indices	26
8	Differentially abundant taxa (DAT)	27
9	Correlation heatmap	28
10	Random forest classification metrics	29
11	Random forest classification metrics from external datasets	30
12	Rarefaction curves for alpha-diversity indices	31
13	Salivary microbiome compositions in the different periodontal statuses	32
14	Correlation plots for differentially abundant taxa	33
15	Clinical measurements by the periodontitis statuses	34
16	Number of read counts by the periodontitis statuses	35
17	Proportion of DAT	36

18	Random forest classification metrics with the full microbiome compositions and ANCOM-selected DAT compositions	37
19	Alpha-diversity indices account for evenness	38
20	Gradient Boosting classification metrics	39

List of Tables

1	Standard clinical information of study participants	8
2	Clinical characteristics of the study subjects	20
3	Feature combinations and their evaluations	21
4	Proportion of taxa according to periodontal statutes	22
5	List of DAT among the periodontally healthy and periodontitis stages	23
6	Feature the importance of taxa in the classification of different periodontal statuses.	24
7	Beta-diversity pairwise comparisons on the periodontitis statuses	25

List of Abbreviations

ACC Accuracy

ASV Amplicon sequence variant

AUC Area-under-curve

BA Balanced accuracy

C-section Cesarean section

DAT Differentially abundant taxa

F1 F1 score

Faith PD Faith's phylogenetic diversity

FTB Full-term birth

GA Gestational age

MWU test Mann-Whitney U-test

PRE Precision

PROM Prelabor rupture of membrane

PTB Preterm birth

ROC curve Receiver-operating characteristics curve

rRNA Ribosomal RNA

SD Standard deviation

SEN Sensitivity

SPE Specificity

t-SNE t-distributed stochastic neighbor embedding

1 Introduction

The microbiome refers to the complex community of microorganisms, including bacteria, viruses, fungi, and other microbes, that inhabit various environments within living organisms (Ursell, Metcalf, Parfrey, & Knight, 2012; Gilbert et al., 2018). In humans, the microbiome plays a crucial role in maintaining health (Lloyd-Price, Abu-Ali, & Huttenhower, 2016), influencing processes such as digestion (Lim, Park, Tong, & Yu, 2020), immune response (Thaiss, Zmora, Levy, & Elinav, 2016; Kogut, Lee, & Santin, 2020; Kim, 2018), and even mental health (Mayer, Tillisch, Gupta, et al., 2015; Zhu et al., 2017; X. Chen, D'Souza, & Hong, 2013). These microbial communities are not static nor constant, but rather dynamic ecosystem that interacts with their host and respond to environmental changes. Recent studies have revealed that imbalances in the microbiome, known as dysbiosis, can contribute to a wide range of diseases, including obesity (John & Mullin, 2016; Tilg, Kaser, et al., 2011; Castaner et al., 2018), diabetes (Barlow, Yu, & Mathur, 2015; Hartstra, Bouter, Bäckhed, & Nieuwdorp, 2015; Sharma & Tripathi, 2019), infections (Whiteside, Razvi, Dave, Reid, & Burton, 2015; Alverdy, Hyoju, Weigerinck, & Gilbert, 2017), inflammatory conditions (Francescone, Hou, & Grivennikov, 2014; Peirce & Alviña, 2019; Honda & Littman, 2012), and cancers (Helmink, Khan, Hermann, Gopalakrishnan, & Wargo, 2019; Cullin, Antunes, Straussman, Stein-Thoeringer, & Elinav, 2021; Sepich-Poore et al., 2021; Schwabe & Jobin, 2013). Thus, understanding the composition of the human microbiomes is essential for developing new therapeutic approaches that target these microbial populations to promote health and prevent diseases.

(Diversity indices)

(Classification evaluation)

2 Predicting preterm birth using random forest classifier in salivary microbiome

This section includes the published contents:

Hong, Y. M., Lee, J., Cho, D. H., Jeon, J. H., Kang, J., Kim, M. G., ... & Kim, J. K. (2023). Predicting preterm birth using machine learning techniques in oral microbiome. *Scientific Reports*, 13(1), 21105.

2.1 Introduction

Preterm birth (PTB), characterized by the delivery of neonates prior to 37 weeks of gestation, is one of the major cause to neonatal mortality and morbidity (Blencowe et al., 2012). Multiple pregnancies including twins, short cervical length, and infection on genitourinary tract are known risk factor for PTB (Goldenberg, Culhane, Iams, & Romero, 2008). Nevertheless, the extent to which these aspects affect birth outcomes is still up for debate. Henceforth, strategies to boost gestation and enhance delivery outcomes can be more conveniently implemented when pregnant women at high risk of PTB are identified early (Iams & Berghella, 2010).

Prediction models that can be utilized as a foundation for intervention methods still have an unacceptable amount of classification evaluations, including accuracy, sensitivity, and specificity, despite a great awareness of the risk factors that trigger PTB (Sotiriadis, Papatheodorou, Kavvadias, & Makrydima, 2010). Several attempts have been made to predict PTB through integrating data such as human microbiome composition, inflammatory markers, and prior clinical data with predictive machine learning methods (Berghella, 2012). Because it is affordable and straightforward to use, fetal fibronectin is commonly used in medical applications. However, with a sensitivity of only 56% that merely similar to random prediction, it has a low classification evaluation (Honest et al., 2009). Due to the difficulty and imprecision of the method in general, as well as the requirement for a qualified specialist cervical length measuring is also restricted (Leitich & Kaider, 2003).

Preterm prelabor rupture of membranes (PROM) brought on by gestational inflammation and infection contribute to about 70% of PTB cases (Romero, Dey, & Fisher, 2014). Nevertheless, as antibiotics and anti-inflammatory therapeutic strategies were ineffective to decrease PTB occurrence rates, the pathology of PTB has not been entirely elucidated by inflammatory and infectious pathways (Romero, Hassan, et al., 2014). Recent researches on maternal microbiomes were beginning to examine unidentified connections of PTB as a consequence of developmental processes in molecular biological technology (Fettweis et al., 2019).

However, as anti-inflammatory and antibiotic therapies were insufficient to lower PTB occurrence rates, infectious and inflammatory processes are insufficient to exhaustively clarify the pathogenesis and pathophysiology of PTB. It has been hypothesized that the microbiota linked to PTB originate from either a hematogenous pathway or the female genitourinary tract increasing through the vagina and/or cervix. (Han & Wang, 2013). Vaginal microbiome compositions have been found in women who eventually acquire PTB, and recent studies have tried to predict PTB risk using cervicovaginal fluid (Kindinger et

al., 2017). Even though previous investigation have confirmed the potential relationships between the vaginal microbiome compositions and PTB, these studies are only able to clarify an upward trajectory.

Multiple unfavorable birth outcomes, including PROM and PTB, have been linked to periodontitis as an independence risk factor, according to numerous epidemiological researches (Offenbacher et al., 1996). It is expected that the oral microbiome will be able to explain additional hematogenous pathways in light of these precedents; however, the oral microbiome composition of fetuses is limited understood.

Hence, in order to identify the salivary microbiome linked to PTB and to establish a machine learning prediction model of PTB determined by oral microbiome compositions, this study examined the salivary microbiome compositions of PTB study participants with a full-term birth (FTB) study participants.

2.2 Materials and methods

2.2.1 Study design and study participants

Between 2019 and 2021, singleton pregnant women who received treatment to Jeonbuk National University Hospital for childbirth were the participants of this study. This study was conducted according to the Declaration of Helsinki (Goodyear, Krleza-Jeric, & Lemmens, 2007). The Institutional Review Board authorized this study (IRB file No. 2019-01-024). Participants who were admitted for elective cesarean sections (C-sections) or induction births, as well as those who had written informed consent obtained with premature labor or PROM, were eligible.

2.2.2 Clinical data collection and grouping

Questionnaires and electronic medical records were implemented to gather information on both previous and current pregnancy outcomes. These clinical data included known risk factors, namely maternal age at delivery, diabetes mellitus, hypertension, overweight, C-section, PROM, and history of PTB, along with demographic neonatal factors, such as gestational week on birth, weight, and sex.

2.2.3 Salivary microbiome sample collection

Salivary microbiome samples were collected 24 hours before to delivery using mouthwash. The standard methods of sterilizing were performed. Medical experts oversaw each stage of the sample collecting procedure. Participants received instruction not to eat, drink, or brush their teeth for 30 minutes before sampling salivary microbiome. Saliva samples were gathered by washing the mouth for 30 seconds with 12 mL of a mouthwash solution (E-zen Gargle, JN Pharm, Pyeongtaek, Korea). The samples were tagged with the anonymous ID for each participant and kept at 4 °C until they underwent further processing. Genomic DNA was extracted using an ExgeneTM Clinic SV kit (GeneAll Biotechnology, Seoul, Korea) following with the manufacturer instructions and store at -20 °C.

2.2.4 16s rRNA gene sequencing

Salivary microbiome samples were transported to the Department of Biomedical Engineering of the Ulsan National Institute of Science and Technology . 16S rRNA sequencing was then carried out using a commissioned Illumina MiSeq Reagent Kit v3 (Illumina, San Diego, CA, USA). Library methods were utilized to amplify the V3-V4 areas. 300 base-pair paired end reads were produced by sequencing the pooled library using a v3 600 cycle chemistry after the samples had been diluted to a final concentration of 6 pM with a 20% PhiX control.

2.2.5 Bioinformatics analysis

The independent *t*-test was utilized to evaluate the differences of continuous values between from the PTB participants than the FTB participants; χ^2 -square test was applied to decide statistical differences of

categorical values. Clinical measurement comparisons were conducted using SPSS (version 20.0) (Spss et al., 2011). At $p < 0.05$, statistical significance was taken into consideration.

QIIME2 (version 2022.2) was implemented to import 16S rRNA gene sequences from salivary microbiome samples of study participants for additional bioinformatics processing (Bolyen et al., 2019). DADA2 was used to verify the qualities of raw sequences (Callahan et al., 2016). The remain sequences were clustered into amplicon sequence variants (ASVs). Diversity indices, namely Faith PD for alpha diversity index (Faith, 1992) and Hamming distance for beta diversity index (Hamming, 1950), were calculated. Mann–Whitney–Wilcoxon U-test (Mann & Whitney, 1947), and PERMANOVA multivariate test were evaluated for measuring statistical significance (Anderson, 2014; Kelly et al., 2015).

Taxonomic assignment were implemented with HOMD (version 15.22) (T. Chen et al., 2010). Afterward, DESeq2 was implemented to identify differentially abundant taxa (DAT) that could distinguish between salivary microbiome from PTB and FTB participants (Love, Huber, & Anders, 2014). Taxa with $|\log_2 \text{FoldChange}| > 1$ and $p < 0.05$ were considered as statistically significant.

The taxa for predicting PTB using salivary microbiome data were determined using a random forest classifier (Breiman, 2001). Through stratified k -fold cross-validation ($k = 5$) that preserves the existence rate of PTB and FTB participants, consistency and trust worthy classification were ensured (Wong & Yeh, 2019).

2.3 Results

2.3.1 Overview of clinical information

In the beginning, 69 volunteer mothers were recruited for this study. However, due to insufficient clinical information or twin pregnancies, 10 participants were excluded from the study participants. Demographic and clinical information of the study participants are displayed in table 1. Because PROM is one of the leading factors of PTB, it was prevalent in the PTB group than the FTB group. Other maternal clinical factors did not significantly differ between the FTB and PTB groups. There were no cases in both groups that had a history of simultaneous periodontal disease or cigarette smoking.

2.3.2 Comparison of salivary microbiomes composition

The salivary microbiome composition was composed of 13953804 sequences from 59 study participants, with 102305.95 ± 19095.60 and 64823.41 ± 15841.65 (mean \pm SD) reads/sample before and following the quality-check stage, accordingly. There was not a significant distinction between the PTB and FTB groups with regard to on alpha diversity nor beta diversity metrics (Figure 4).

DESeq2 was used to select 32 DAT that distinguish between the PTB and FTB groups out of the 465 species that were examined (Love et al., 2014): 26 FTB-enriched DAT and six PTB-enriched DAT. Seven PROM-related DAT were removed from these 32 PTB-related DAT to lessen the confounding effect of PROM (Figure 5). Therefore, there were a total of 25 PTB-related DAT: 22 FTB-enriched DAT and three PTB-enriched DAT (Figure 1).

A significant negative correlation was found using Pearson correlation analysis between GW and differences between PTB-enriched DAT and FTB-enriched DAT ($r = -0.542$ and $p = 7.8e - 6$; Figure 5).

2.3.3 Random forest classification to predict PTB risk

To classify PTB according to DAT, random forest classifiers were constructed. The nine most significant DAT were used to obtain the best BA (0.765 ± 0.071 ; Figure 3a). Moreover, random forest classification model determined each DAT's importance (Figure 3b). We conducted a validation procedure on nine twin pregnancies that were excluded in the initial study design in order to confirm the reliability and dependability of our random forest-based PTB prediction model (Figure 6). Comparable to the PTB prediction model on the 59 initial singleton study participants, the validation classification on PTB risk of these twin participants have an accuracy of 87.5%.

Table 1: **Standard clinical information of study participants.**

Continuous variable for independent *t*-test. Categorical variable for Pearson's χ^2 -square test. Continuous variable: mean \pm SD. Categorical variable: count (proportion)

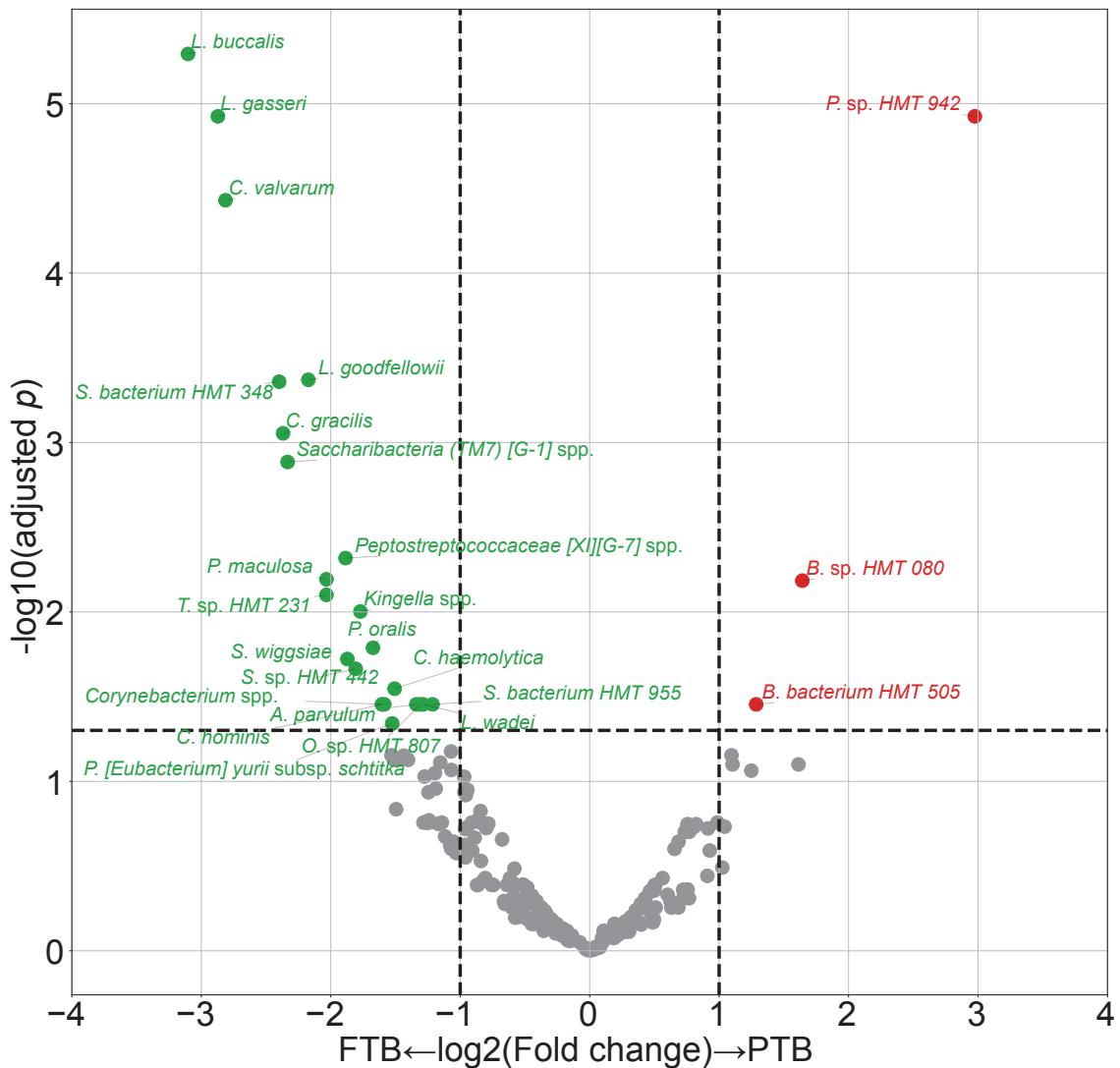


Figure 1: DAT volcano plot.

Red dots represent PTB-enriched DAT, while green dots represent FTB-enriched DAT.

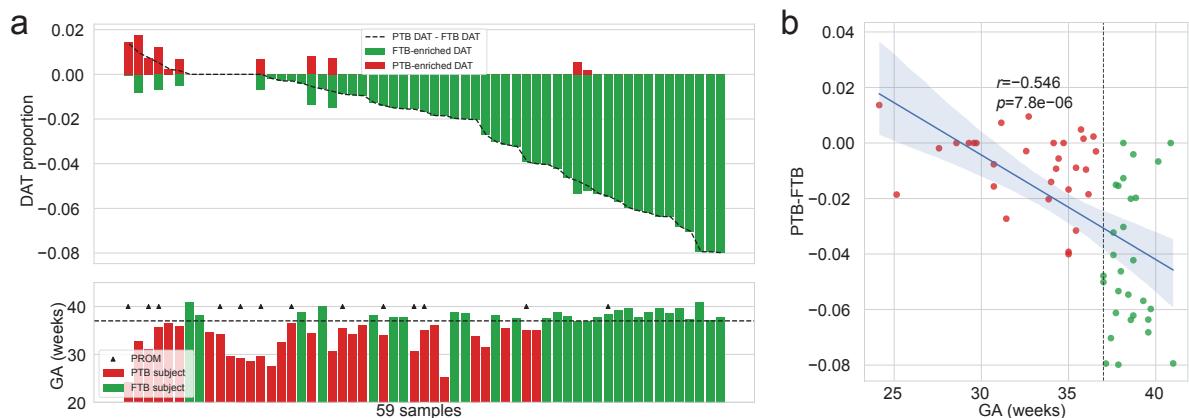


Figure 2: **Salivary microbiome compositions over DAT.**

(a) Frequencies of DAT of study subjects. The study participants are arranged in respect of (PTB-enriched DAT – FTB-enriched DAT). The study participants' GA is displayed in accordance with the upper panel's order (PTB: red bar, FTB: green bar. PROM: arrow head.) **(b)** Correlation plot with GA and (PTB-enriched DAT – FTB-enriched DAT). Strong negative correlation is found with Pearson correlation.

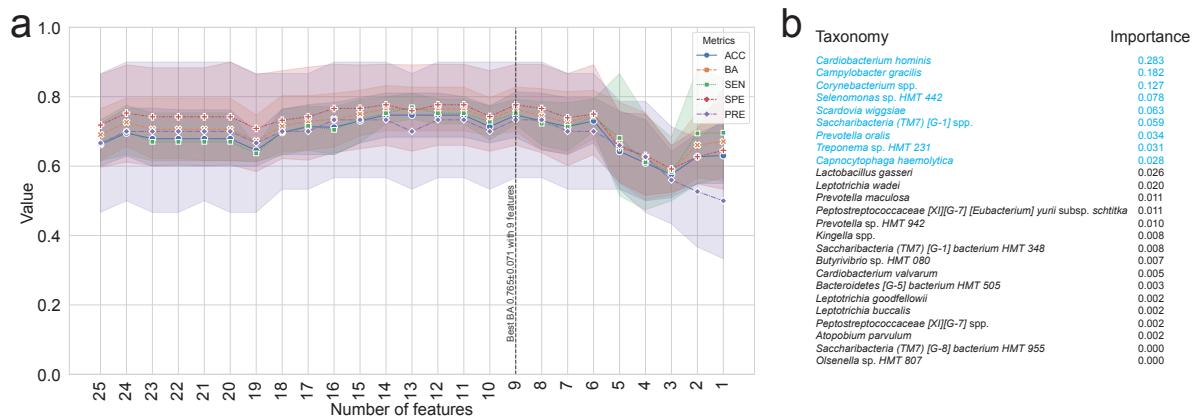


Figure 3: **Random forest-based PTB prediction model.**

(a) Machine learning evaluations upon number of features (DAT). Random Forest classifier has the best BA (0.765 ± 0.071 ; Mean \pm SD) with the nine most important DAT. **(b)** Importance of DAT.

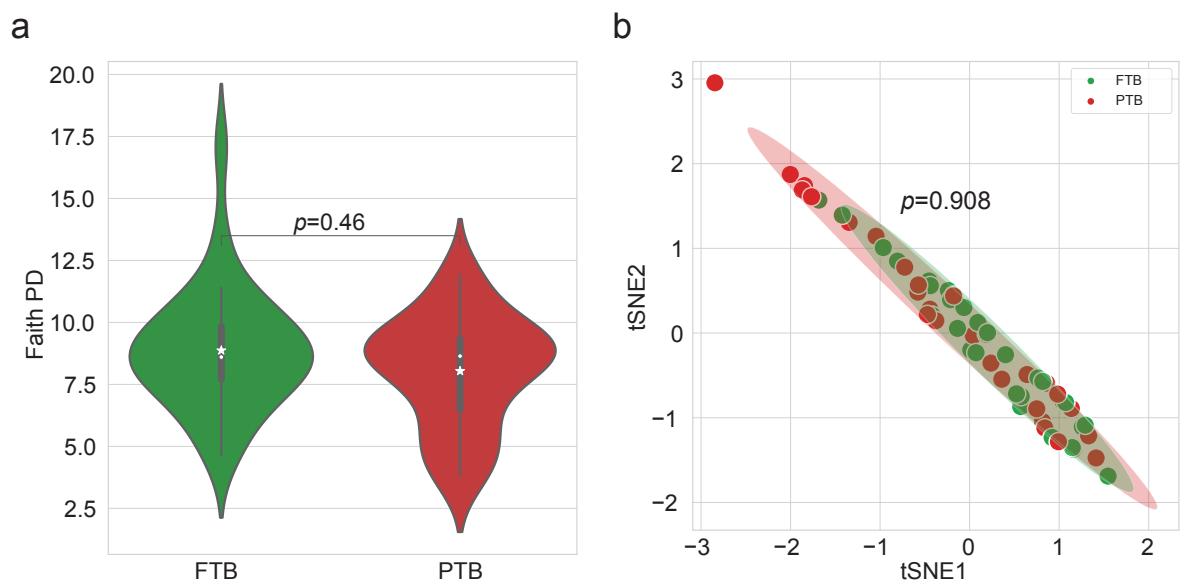


Figure 4: **Diversity indices.**

(a) Alpha diversity index (Faith PD). There is no statistically significant difference between the PTB and FTB group (MWU test $p = 0.46$). **(b)** t-SNE plot with beta diversity index (Hamming distance). There is no statistically significant difference between the PTB and FTB group (PERMANOVA test $p = 0.908$)

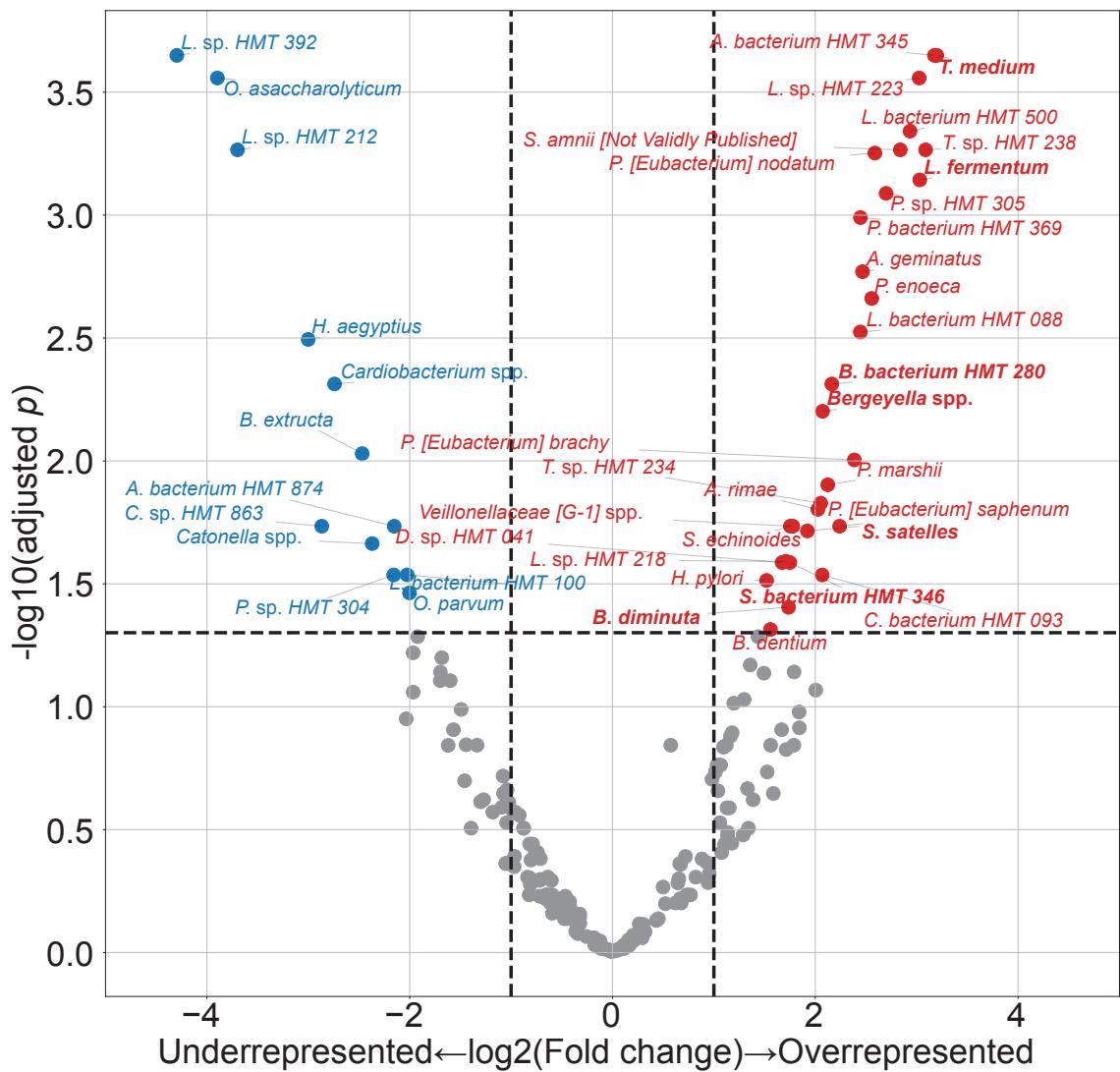


Figure 5: **PROM-related DAT**.

Only seven of these 42 PROM-related DAT overlapped with PTB-related DAT (bold text). Blue dots represented PROM-underrepresented DAT, while red dots represented PROM-overrepresented DAT.

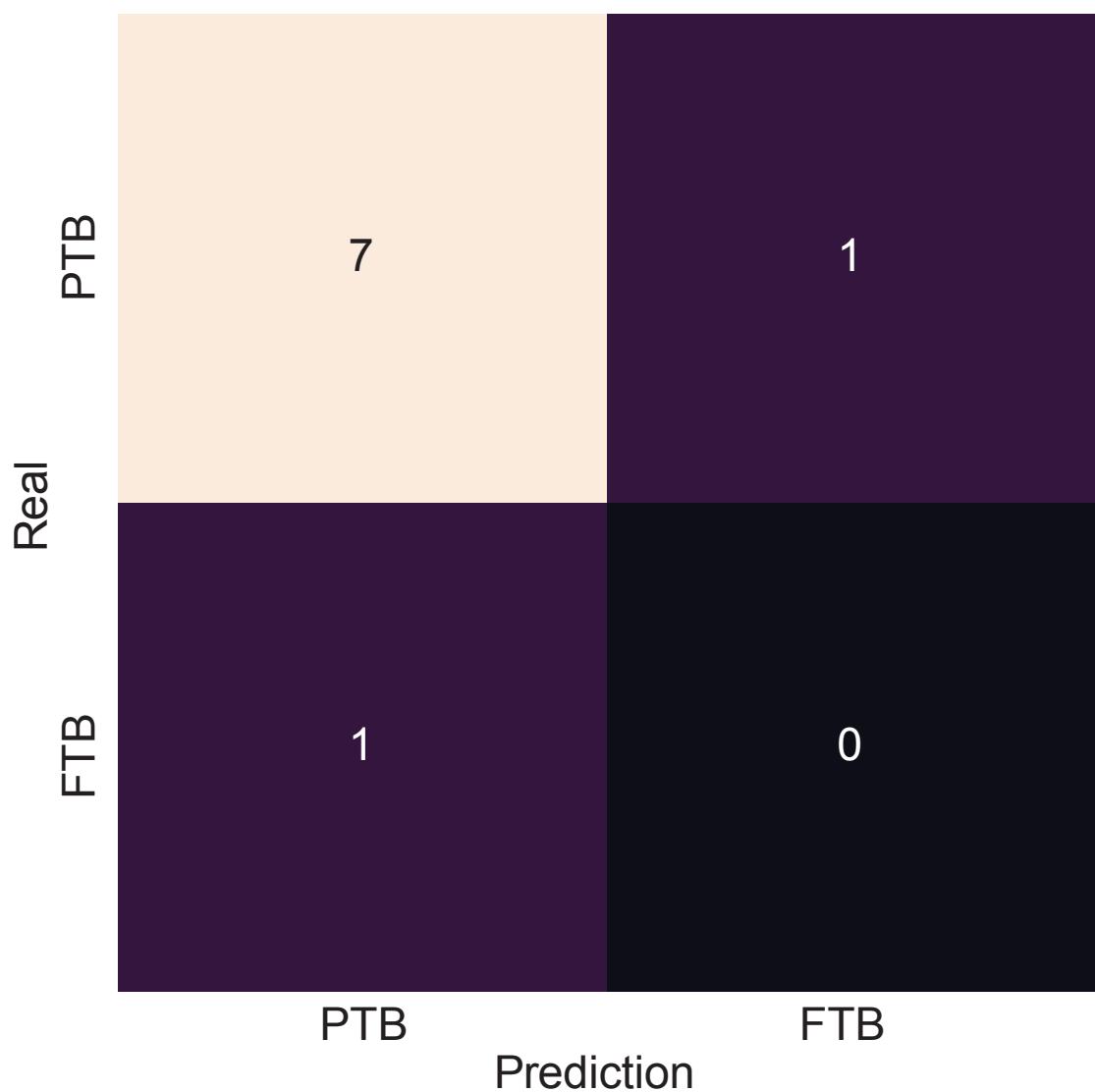


Figure 6: Validation of random forest-based PTB prediction model.

Nine twin pregnancies (eight PTB subjects and a FTB subject) that were excluded in the initial study subjects were subjected to a validation procedure. The random forest-based PTB prediction model shows 87.5% accuracy, comparable to the PTB classification evaluations on the singleton study subjects (0.714 ± 0.061 . Mean \pm SD)

2.4 Discussion

In this study, we employed salivary microbiome compositions to develop the random forest-based PTB prediction models to estimate PTB risks. Previous reports have indicated bidirectional associations between pregnancy outcomes and salivary microbiome compositions (Han & Wang, 2013). Nevertheless, the salivary microbiome composition is not yet elucidated. Salivary microbial dysbiosis, including gingival inflammation and periodontitis, have been connected to unfavorable pregnancy outcomes, such as PTB (Ide & Papapanou, 2013). However, the techniques utilized in recent research that primarily focus on recognized infections have led to inconsistent outcomes.

One of the most common salivary taxa that has been examined is *Fusobacterium nucleatum* (Han, 2015; Brennan & Garrett, 2019; Bolstad, Jensen, & Bakken, 1996), that is a Gram-negative, anaerobic, and filamentous bacteria. *Fusobacterium nucleatum* can be separated from not only the salivary microbiome but also the vaginal microbiome (Vander Haar, So, Gyamfi-Bannerman, & Han, 2018; Witkin, 2019). In both animal and human investigation, *Fusobacterium nucleatum* infection has been linked to risk of PTB (Doyle et al., 2014). According to recent researches, the placenta women who give birth prematurely may include additional salivary microbiome dysbiosis, such as *Bergeyella* spp. and *Porphyromonas gingivalis* (León et al., 2007; Katz, Chegini, Shiverick, & Lamont, 2009). Although *Bergeyella* spp. were one of the PROM-overrepresented DAT (Figure 5), it was excluded in the final 25 PTB-related DAT. Furthermore, *Porphyromonas gingivalis* and *Campylobacter gracilis* were pathogens of periodontitis in subgingival microbiome (Yang et al., 2022). *Lactobacillus gasseri* was also one of the FTB-enriched DAT (Figure 1), and it is well established that early PTB risk can be reduced by *Lactobacillus gasseri* in the vaginal microbiome (Basavaprabhu, Sonu, & Prabha, 2020; Payne et al., 2021).

With DAT comprising 22 FTB-enriched DAT and three PTB-enriched DAT (Figure 1), we discovered that the FTB study participants had the majority of the essential DAT that distinguished between the PTB and FTB groups. Thus, we hypothesize that the pathogenesis and pathophysiology of PTB may have been triggered by an absence of species with protective characteristics. The association between unfavorable pregnancy outcomes and a dysfunctional microbiome has been explained through two distinct processes. According to the first hypothesis, periodontal pathogens originating in the gingival biofilm might spread from the infected salivary microbiome over the placenta microbiome, invade the intra-amniotic fluid and fetal circulation, and then have a direct impact on the fetoplacental unit, leading to bacteremia (Hajishengallis, 2015). Based on the second hypothesis, inflammatory mediators and endotoxins that generated by the sub-gingival inflammation and derived from dental plaque of periodontitis may spread throughout the body and reach the fetoplacental unit (Stout et al., 2013; Aagaard et al., 2014). Despite belonging to the same species, some subgroups of the salivary microbiome may influence pregnancy outcomes in both favorable and adverse manners. Following this line of argumentation, the salivary microbiome composition or their dysbiosis are more significant than the existence of particular bacteria.

Notably, microbial alteration that take place throughout pregnancy may be expected results of a healthy pregnancy. Those pregnancy-related vulnerabilities to dental problem like periodontitis can be explained by three factors. Because of hormone-driven gingival hyper-reactivity to the salivary microbiome in the

oral biofilm including sub-gingival biofilm, these conditions are prevalent in pregnant women. For insight at the relationship between the salivary microbiome composition and PTB, further studies with pathway analysis are warranted.

Our study confirmed that salivary microbiome composition could provide potential biomarkers for predicting pregnancy complications including PTB risks using random forest-based classification models, despite a limited number of study participants and a tiny validation sample size. Another limitation of our study was 16S rRNA sequencing. In other words, unlike the shotgun sequencing, 16S rRNA sequencing only focused on bacteria, not viruses nor fungi. We did not delve into other variables like nutrition status and socioeconomic statuses of study participants that might effect the salivary microbiome composition.

Notwithstanding these limitations, this prospective examination showed the promise of the random forest-based PTB prediction models based on mouthwash-derived salivary microbiome composition. Before applying the methods developed in this study in a clinical context, more multi-center and extensive research is warranted to validate our findings.

3 Random forest prediction model for periodontitis statuses based on the salivary microbiomes

This section includes the published contents:

3.1 Introduction

Saliva microbial dysbiosis brought on by the accumulation of plaque results in periodontitis, a chronic inflammatory disease of the tissue that surrounds the tooth (Kinane, Stathopoulou, & Papapanou, 2017). Loss of periodontal attachment is a consequence of periodontitis, which may lead to irreversible bone loss and, eventually, permanent tooth loss if left untreated. A new classification criterion of periodontal diseases was created in 2018, about 20 years after the 1999 statements of the previous one. Even with this evolution, radio-graphic and clinical markers of periodontitis progression remain the primary methods for diagnosing periodontitis (Papapanou et al., 2018). Such tools, nevertheless, frequently demonstrate the prior damage from periodontitis rather than its present condition. Certain individuals have a higher risk of periodontitis, a higher chance of developing severe generalized periodontitis, and a worse response to common salivary bacteria control techniques utilized to prevent and treat periodontitis. As a result, the 2017 framework for diagnosing periodontitis additionally allows for the potential development of biomarkers to enhance diagnosis and treatment of periodontitis (Tonetti, Greenwell, & Kornman, 2018).

3.2 Materials and methods

3.2.1 Study participants enrollment

Between 2018-08 and 2019-03, 250 study participants—100 healthy controls, 50 patients with stage I periodontitis, 50 patients with stage II periodontitis, and 50 patients with stage III periodontitis—visited the Department of Periodontics at Pusan National University Dental Hospital. The Institutional Review Board of the Pusan National University Dental Hospital accepted this study protocol and design (PNUDH-2016-019). Every study participants provided their written informed authorization after being fully informed about this study's objectives and methodologies. Exclusion criteria for the study participants are followings:

1. People who, throughout the previous six months, underwent periodontal therapy, including root planing and scaling.
2. People who struggle with systemic conditions that may affect periodontitis developments, such as diabetes.
3. People who, throughout the previous three months, were prescribed anti-inflammatory medications or antibiotics.
4. Women who were pregnant or breastfeeding.
5. People who have persistent mucosal lesions, e.g. pemphigus or pemphigoid, or acute infection, e.g. herpetic gingivostomatitis.

An skilled periodontist conducted each clinical procedure. Six sites per tooth were used to quantify gingival recession and probing depth: mesiobuccal, midbuccal, distobuccal, mesiolingual, midlingual, and distolingual (Huang et al., 2007).

3.2.2 Periodontal clinical parameter diagnosis

3.2.3 Saliva sampling and DNA extraction procedure

3.2.4 Bioinformatics analysis

3.3 Results

- 3.3.1 Summary of study subjects and sequencing data**
- 3.3.2 Diversity indices reveal differences in diversity among the periodontal statuses**
- 3.3.3 Differentially abundant taxa among healthy and different periodontitis stages and their correlation**
- 3.3.4 Classification of periodontitis stages by random forest models**

Table 2: Clinical characteristics of the study subjects.

Significant differences were assessed using the Kruskal-Wallis test. NA: Not applicable.

Table 3: Feature combinations and their evaluations

Classification performance with the most important taxon, the two most important taxa, and taxa with the best-balanced accuracy. *P.gingivalis* and *Act.* are *Porphyromonas gingivalis* and *Actinomyces* spp., respectively.

Table 4: **Proportion of taxa according to periodontal statutes**

Table 5: List of DAT among healthy status and periodontitis stages

Table 6: Feature the importance of taxa in the classification of different periodontal statuses
Taxa are ranked in descending order of importance; from most important to least important.

Table 7: Beta-diversity pairwise comparisons on the periodontitis statuses

Statistically significant (p-value) was determined by the PERMANOVA test.

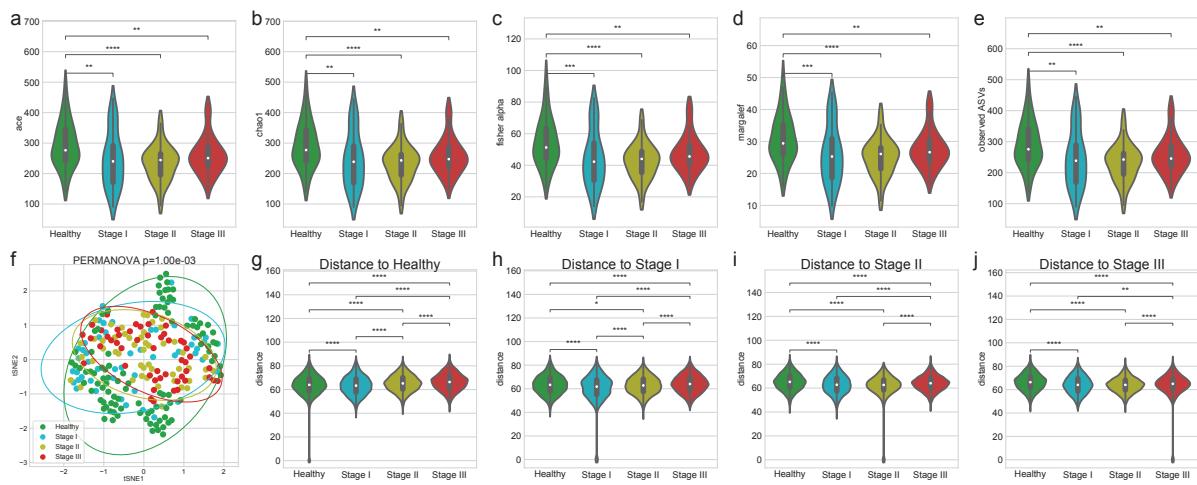


Figure 7: **Diversity indices.**

Alpha-diversity indices (**a-e**) indicate that healthy controls have increased heterogeneity than periodontitis stages as measured by: **(a)** ace **(b)** chao1 **(c)** Fisher alpha **(d)** Margalef, and **(e)** observed ASVs. **(f)** The beta-diversity index (weighted UniFrac) was visualized using a tSNE-transformed plot. The confidence ellipses are shown to display the distribution of each periodontitis stage. The distance to each stage demonstrated that each periodontitis stage was distinguished from the other periodontitis stages: **(g)** distance to Healthy **(h)** distance to Stage I **(i)** distance to Stage II, and **(j)** distance to Stage III. Statistical significance determined by the MWU test and the PERMANOVA test: $p \leq 0.01$ (***) and $p \leq 0.0001$ (****).

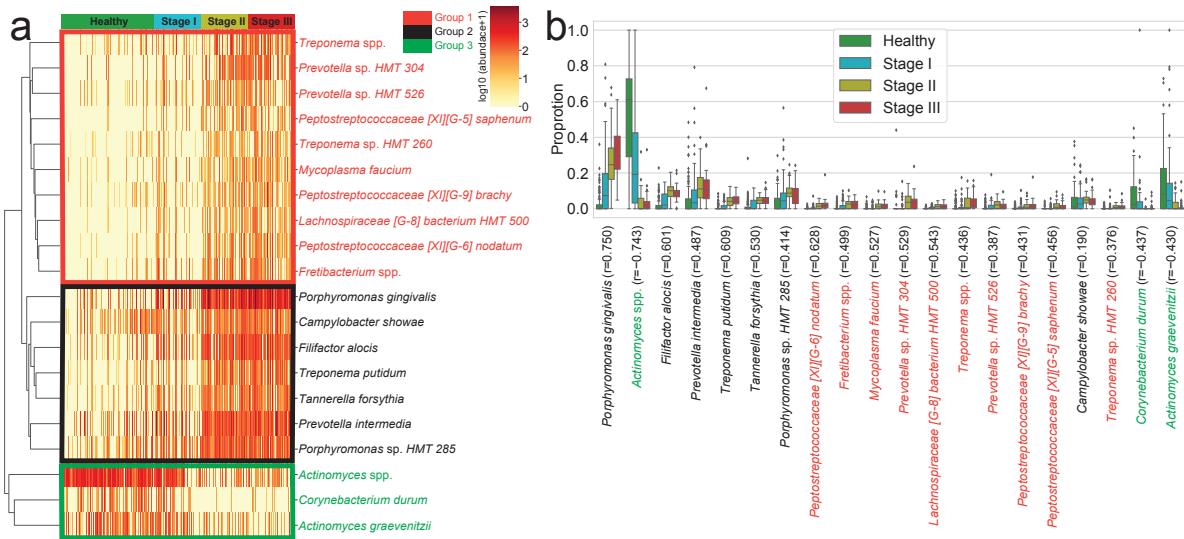


Figure 8: **Differentially abundant taxa (DAT).**

DAT that were identified by ANCOM. **(a)** Heatmap of clustered DAT with similar distribution among subjects. Group 1, Group 2, and Group 3 are marked in red, black, and green, respectively. **(b)** Box plots showing the proportions of DAT. Taxa were sorted by their importance according to ANCOM.

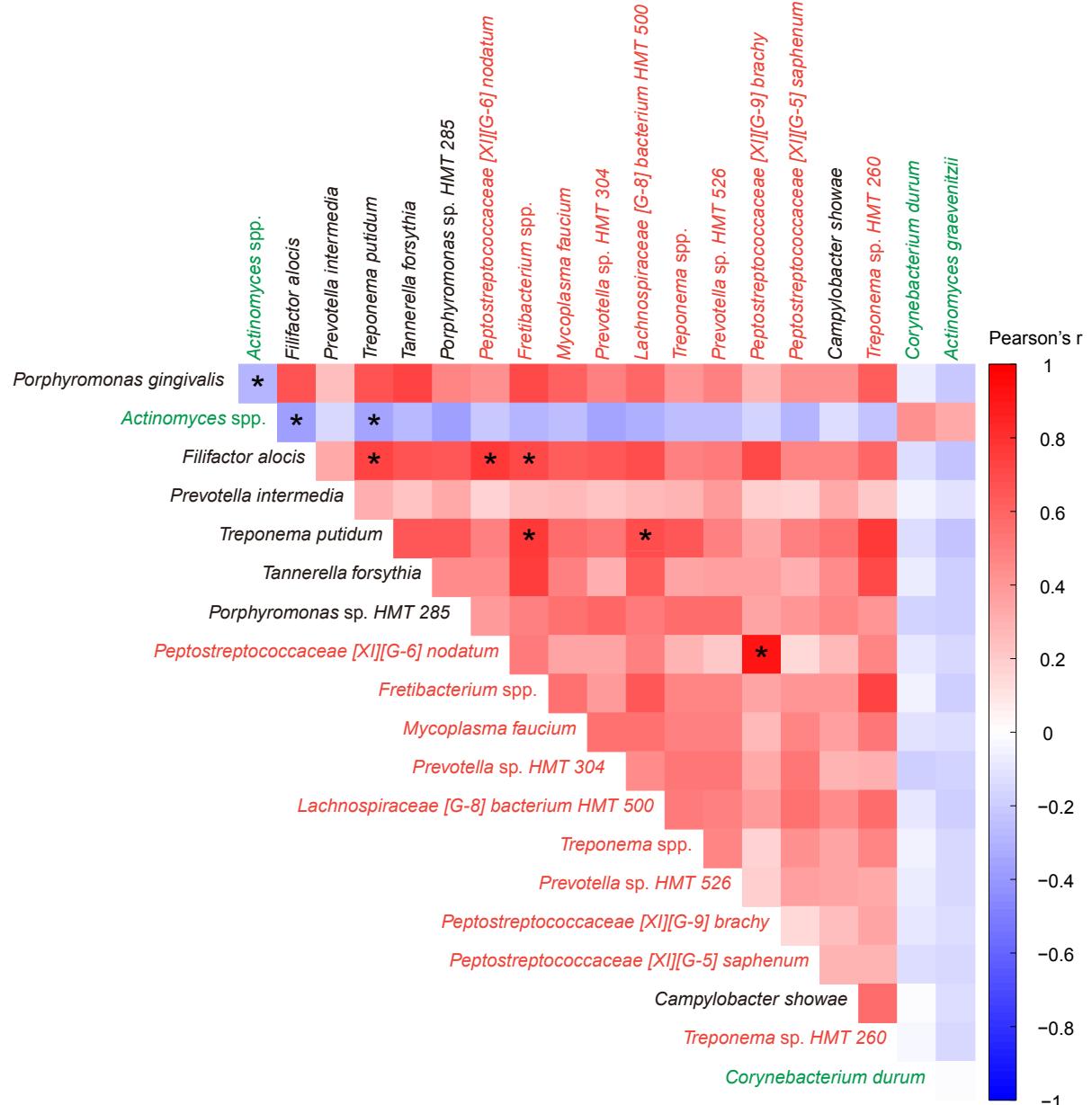


Figure 9: Correlation heatmap.

Pearson's correlations between DAT in healthy status and periodontitis stages. Statistical significance was determined by strong correlation, i.e., $| \text{coefficient} | \geq 0.5$ (*).

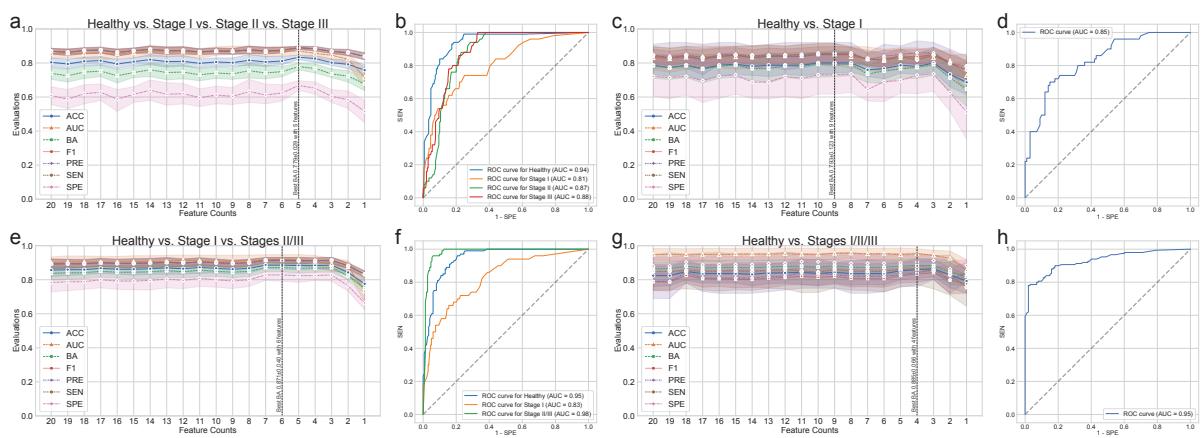


Figure 10: Random forest classification metrics.

The classification metrics in the random forest classifications were as follows: ACC, AUC, BA, F1, PRE, SEN, and SPE. **(a)** Classification performance for healthy vs. stage I vs. stage II vs. stage III. **(b)** ROC curve for the highest BA of (a). **(c)** Classification performance for healthy vs. stage I. **(d)** ROC curve on the highest BA of (c). **(e)** Classification performance for healthy vs. stage I vs. stages II/III. **(f)** ROC curve for the highest BA of (e). **(g)** Classification performance for healthy vs. stages I/II/III. **(h)** ROC curve for the highest BA of (h).

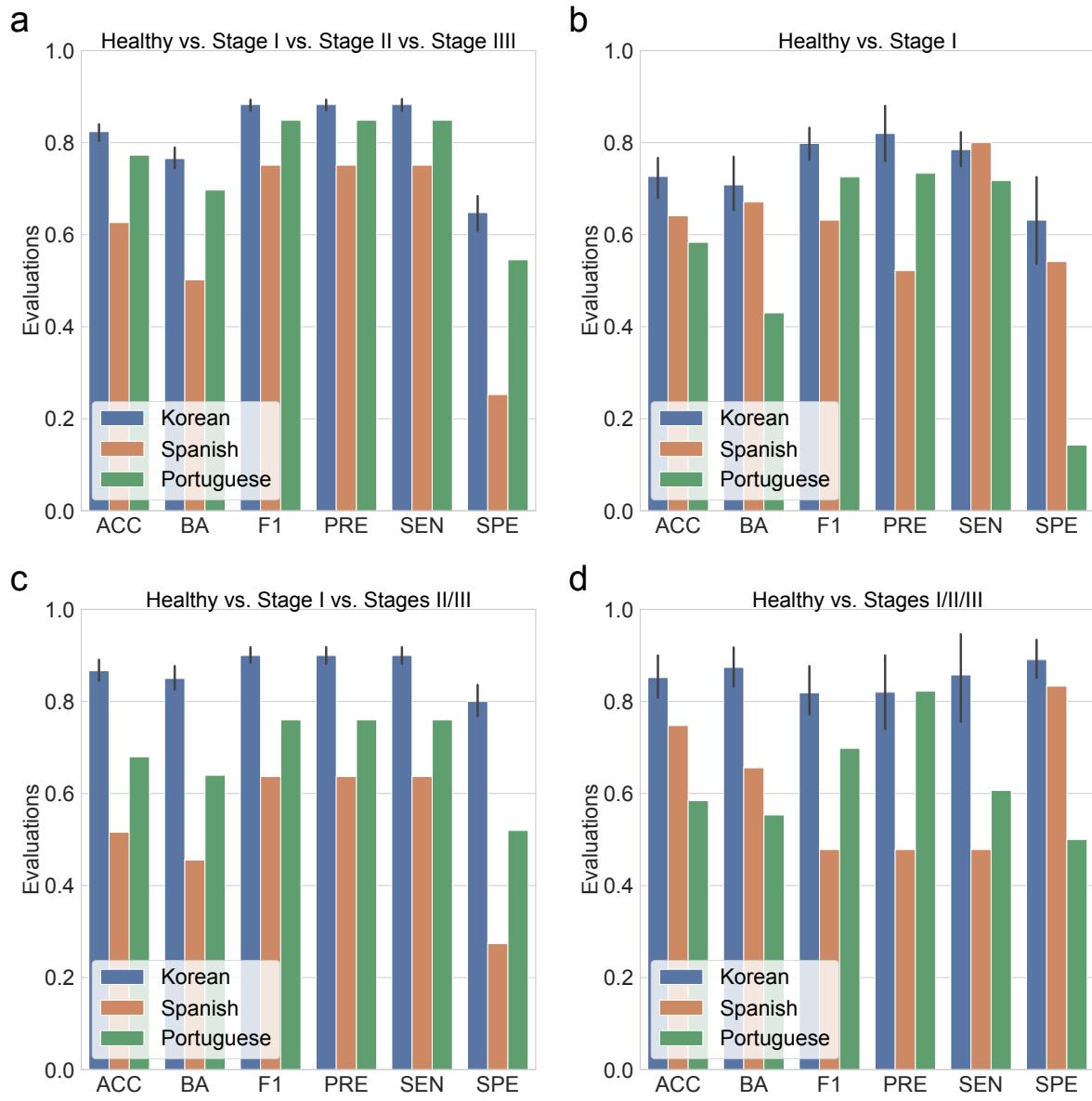


Figure 11: **Random forest classification metrics from external datasets.**

The classification metrics in the random forest classifications were as follows: ACC, AUC, BA, F1, PRE, SEN, and SPE. (a) Classification performance for healthy vs. stage I vs. stage II vs. stage III. (b) Classification performance for healthy vs. stage I. (c) Classification performance for healthy vs. stage I vs. stages II/III. (d) Classification performance for healthy vs. stages I/II/III.

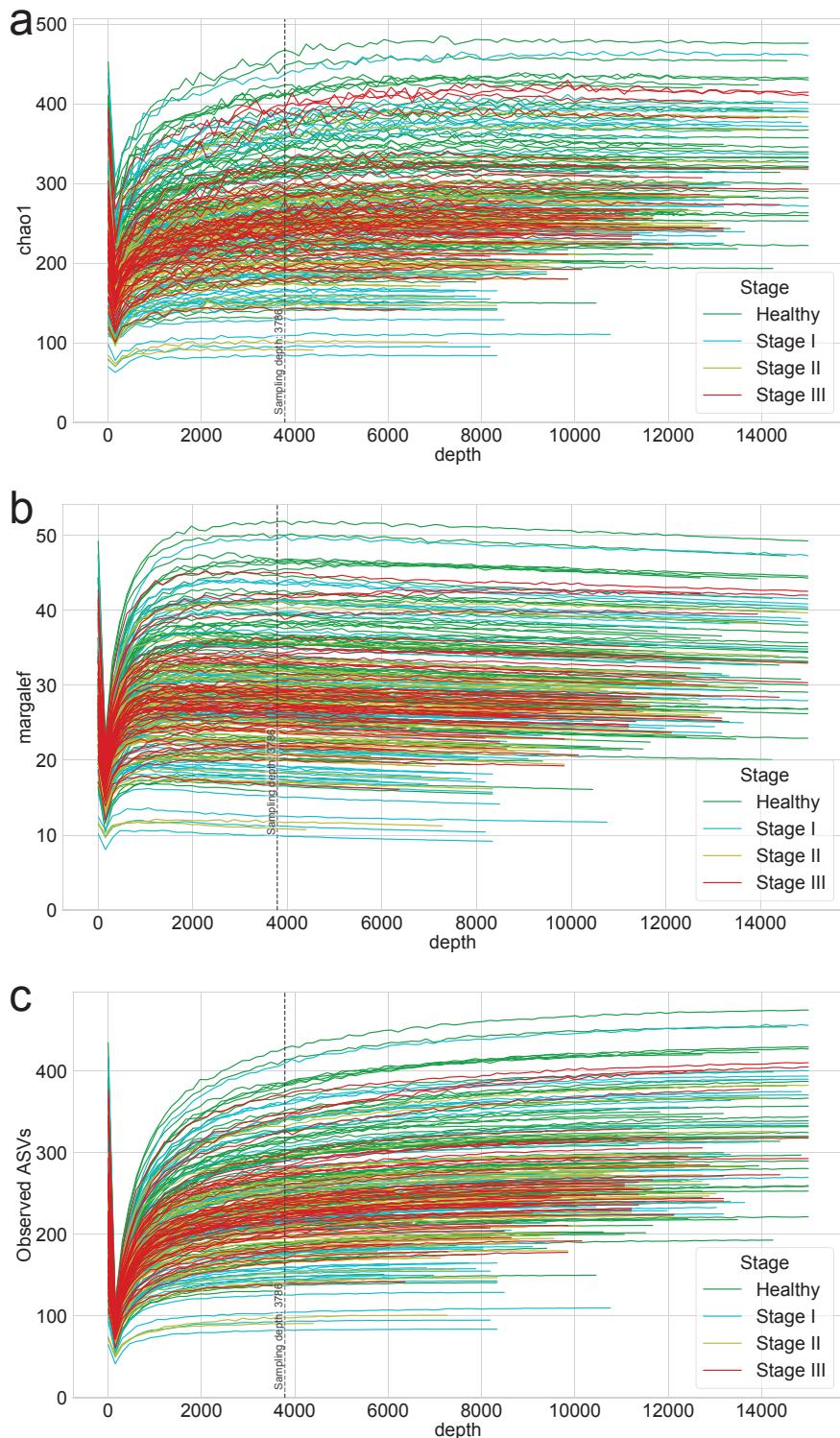


Figure 12: **Rarefaction curves for alpha-diversity indices.**

Rarefaction of (a) chao1 (b) margalef, and (c) observed ASVs were generated to measure species richness and determine the sampling depth of each sample.

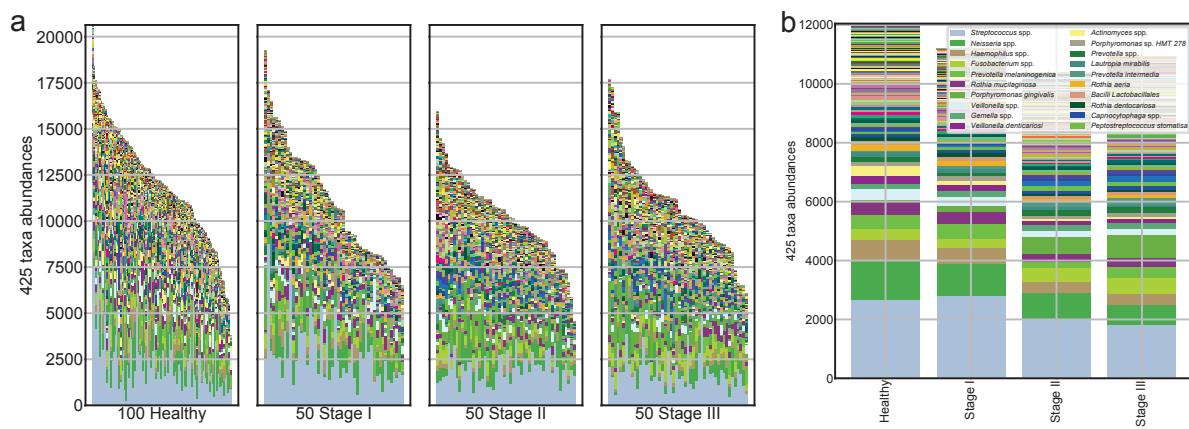


Figure 13: Salivary microbiome compositions in the different periodontal statuses.

Stacked bar plot of the absolute abundance of bacterial species for all samples (**a**) and the mean absolute abundance of bacterial species in the healthy, stage I, stage II, and stage III groups (**b**).

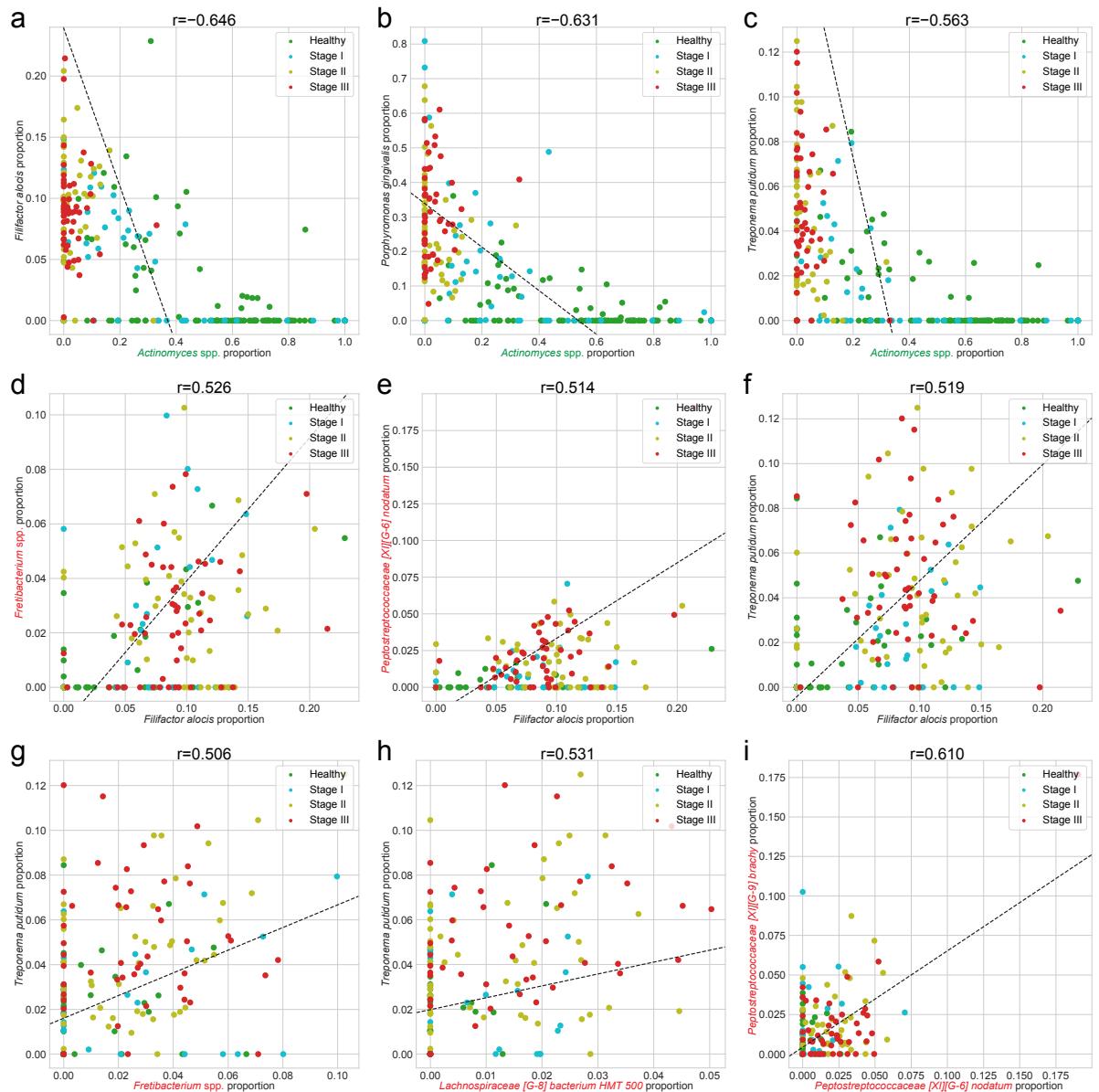


Figure 14: Correlation plots for differentially abundant taxa.

We selected the combinations of DAT with absolute Spearman correlation coefficients greater than 0.5. The color represents periodontal healthy periodontal statuses (green: healthy, cyan: stage I, yellow: stage II, and red: stage III).

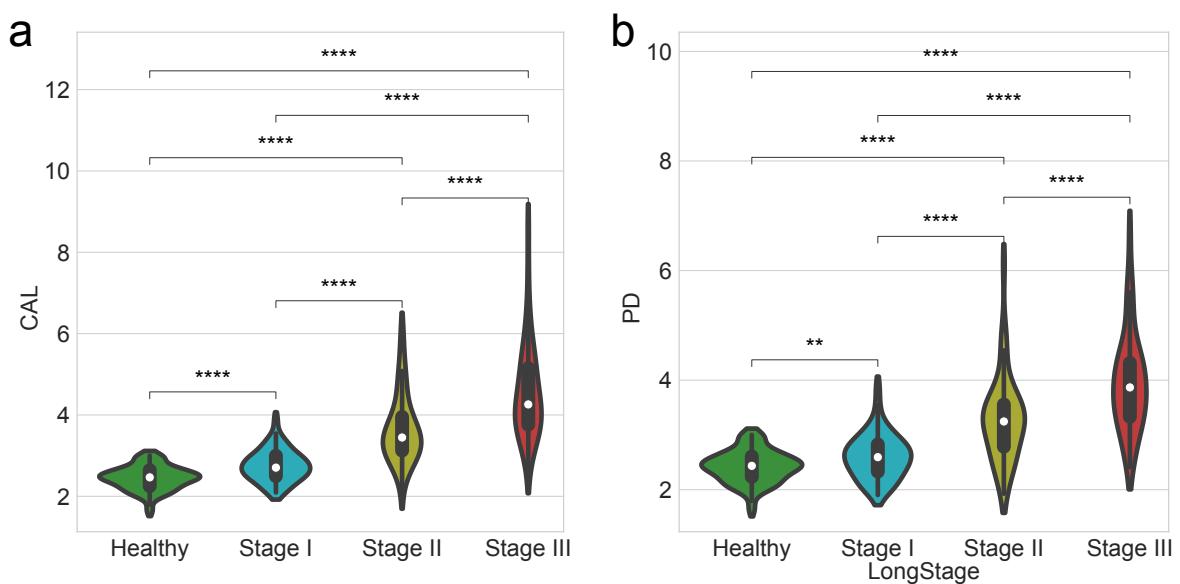


Figure 15: Clinical measurements by the periodontitis statuses.

Comparisons of clinical measurement among healthy controls and patients with various periodontitis stages. (a) Clinical attachment level (b) Probing depth. Statistical significance determined by the Mann-Whitney U-test: $p \leq 0.01$ (**) and $p \leq 0.0001$ (****).

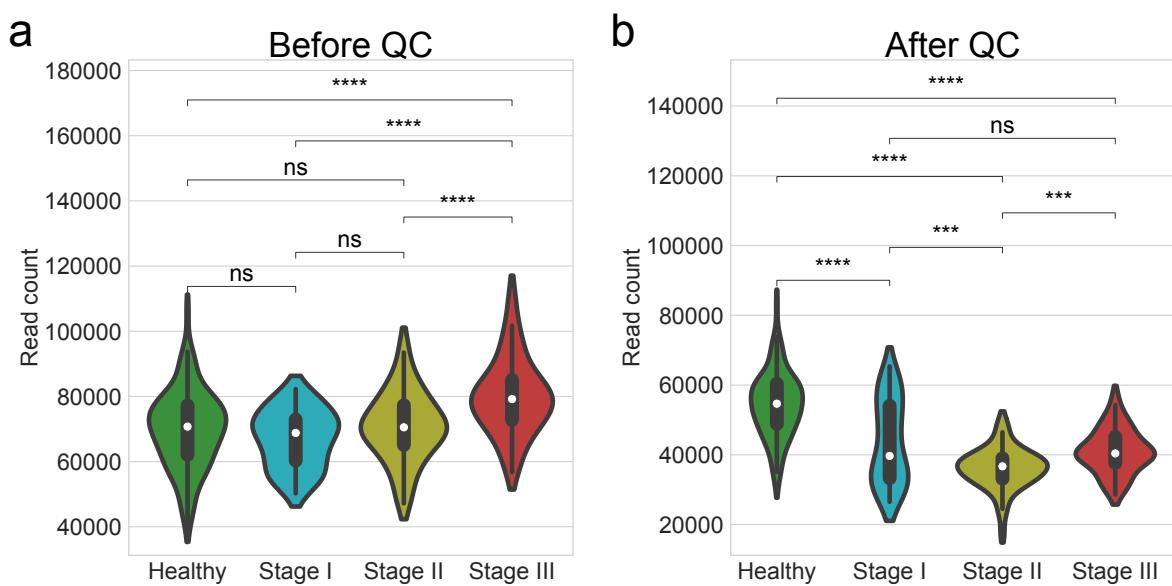


Figure 16: Number of read counts by the periodontitis statuses.

Comparisons of the number of read counts among healthy controls and patients with various periodontitis stages. (a) Before quality check (b) After quality check. Statistical significance determined by the Mann-Whitney U-test: $p > 0.05$ (ns), $p \leq 0.001$ (***) , and $p \leq 0.0001$ (****).

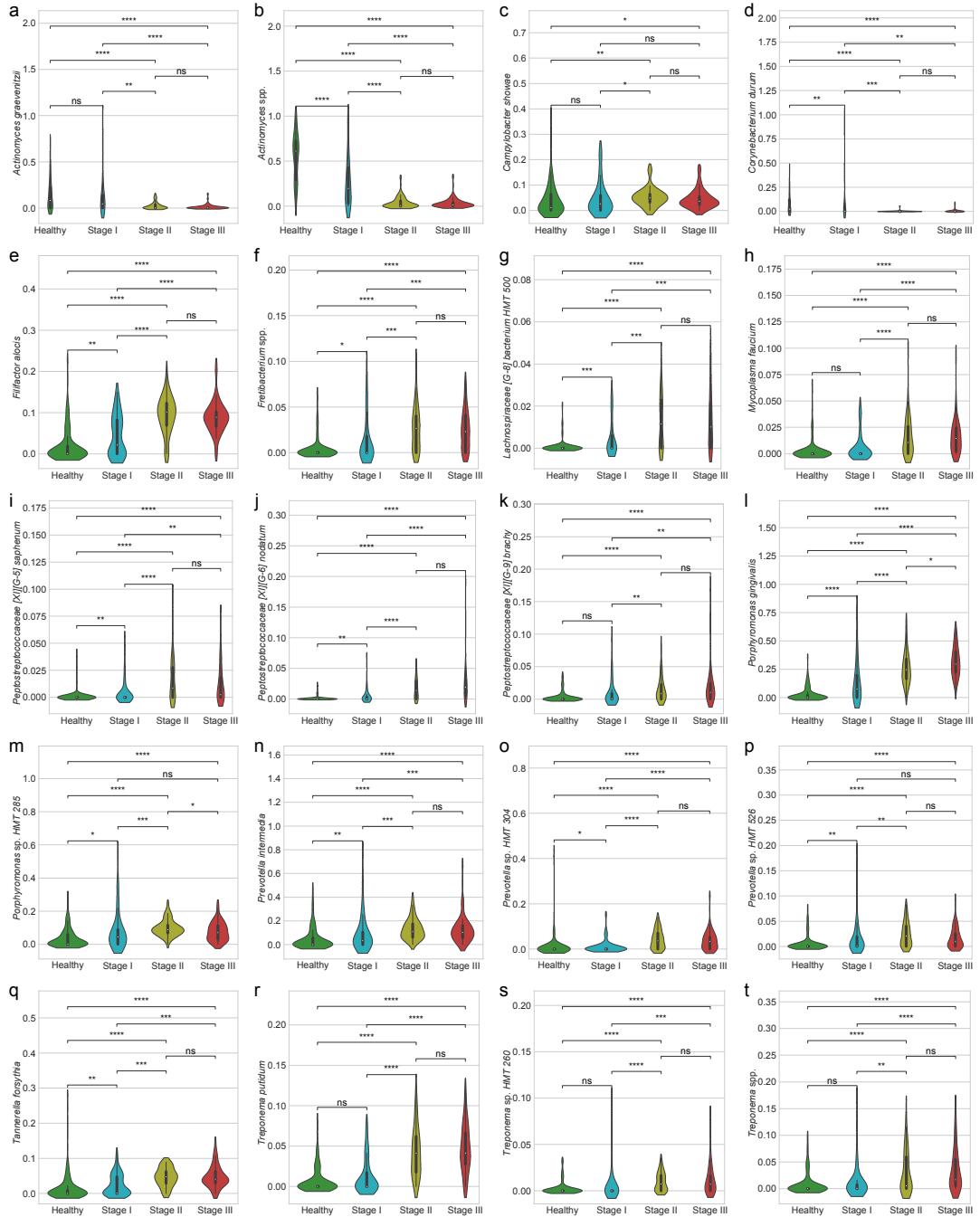


Figure 17: Proportion of DAT.

(a) *Actinomyces graevenitzii* (b) *Actinomyces* spp. (c) *Campylobacter showae* (d) *Corynebacterium durum* (e) *Filifactor alocis* (f) *Fretibacterium* spp. (g) *Lachnospiraceae* [G-8] bacterium HMT 500 (h) *Mycoplasma faecium* (i) *Peptostreptococcaceae* [XI][G-5] saphenum (j) *Peptostreptococcaceae* [XI][G-6] nodatum (k) *Peptostreptococcaceae* [XI][G-9] brachy (l) *Porphyromonas gingivalis* (m) *Porphyromonas* sp. HMT 285 (n) *Prevotella intermedia* (o) *Prevotella* sp. HMT 304 (p) *Prevotella* sp. HMT 526 (q) *Tannerella forsythia* (r) *Treponema putidum* (s) *Treponema* sp. HMT 260 (t) *Treponema* spp. Statistical significance determined by the MWU test: $p > 0.05$ (ns), $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***), and $p \leq 0.0001$ (****).

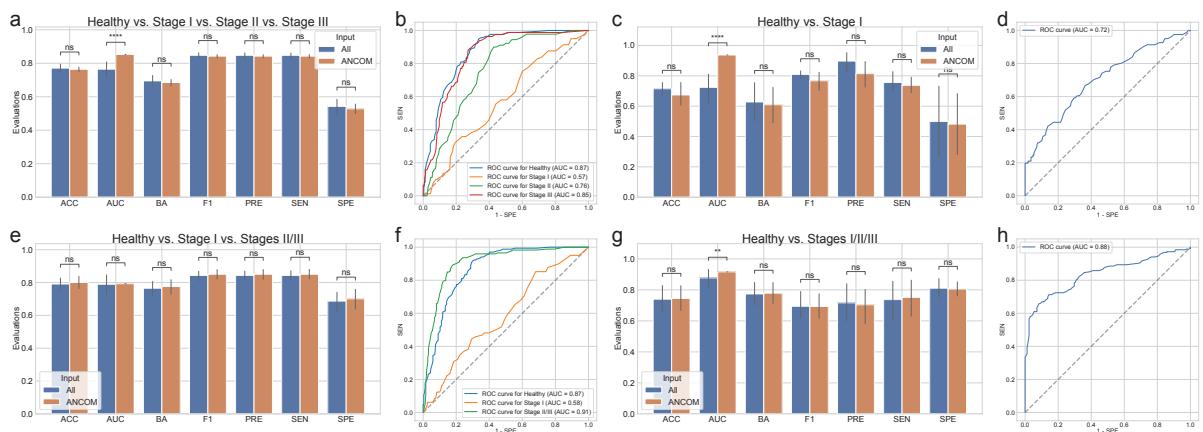


Figure 18: Random forest classification metrics with the full microbiome compositions and ANCOM-selected DAT compositions.

The classification metrics in the random forest classifications were as follows: ACC, AUC, BA, F1, PRE, SEN, and SPE. (a) Classification performance for healthy vs. stage I vs. stage II vs. stage III. (b) ROC curve for the highest BA of (a). (c) Classification performance for healthy vs. stage I. (d) ROC curve on the highest BA of (c). (e) Classification performance for healthy vs. stage I vs. stages II/III. (f) ROC curve for the highest BA of (e). (g) Classification performance for healthy vs. stages I/II/III. (h) ROC curve for the highest BA of (g). Statistical significance determined by the MWU est test: $p > 0.05$ (ns), $p \leq 0.01$ (**), and $p \leq 0.0001$ (****).

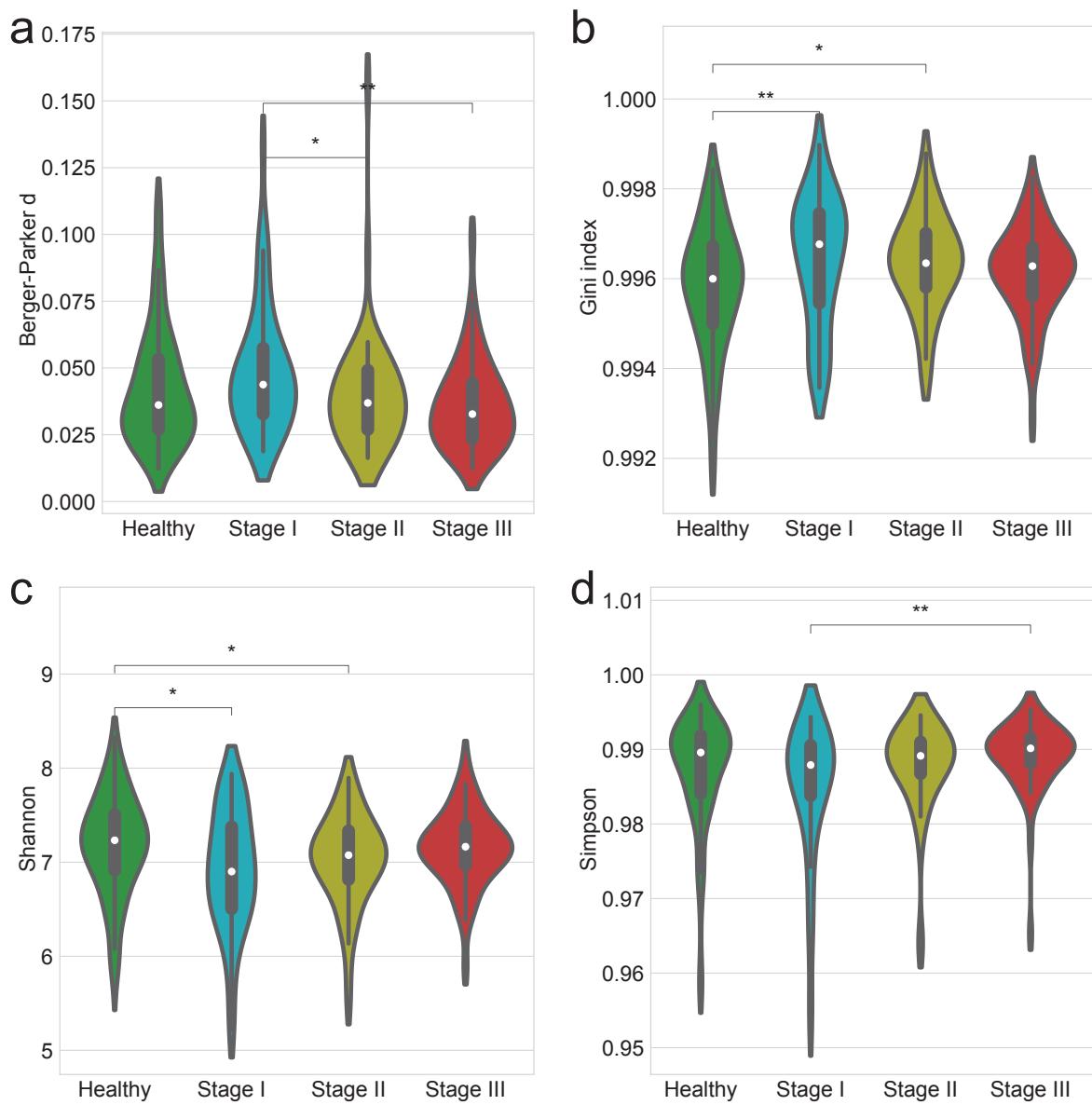


Figure 19: Alpha-diversity indices account for evenness.

Alpha-diversity indices (a-d) indicate that the heterogeneity between the periodontitis stages as measured by: (a) Berger-Parker d (b) Gini (c) Shannon (d) Simpson. Statistical significance determined by the MWU test test: $p \leq 0.05$ (*) and $p \leq 0.01$ (**)

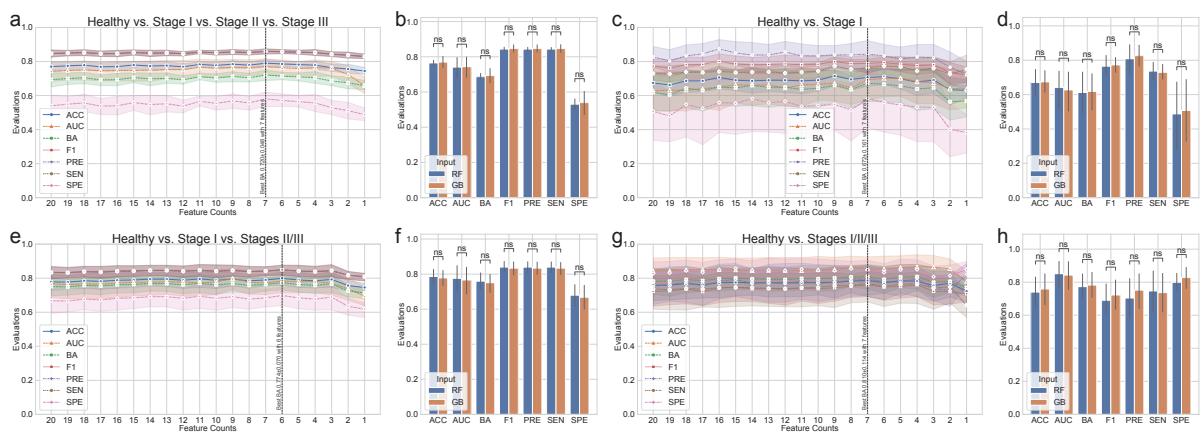


Figure 20: Gradient Boosting classification metrics.

The classification metrics in the random forest classifications were as follows: ACC, AUC, BA, F1, PRE, SEN, and SPE. The feature counts mean that the classification model trained on the most important n features as the table 5. (a) Comparison of Random forest (RF) and Gradient boosting (GB) for healthy vs. stage I vs. stage II vs. stage III. (b) Comparison of RF and GB for the highest BA of (a). (c) Classification performance for healthy vs. stage I. (d) Comparison of RF and GB for healthy vs. stage I vs. stages II/III. (f) Comparison of RF and GB for Healthy vs. Stage I vs. Stages II/III. (g) Classification performance for healthy vs. stages I/II/III. (h) Comparison of RF and GB for Healthy vs. Stages I/II/III.

3.4 Discussion

4 Lung microbiome

4.1 Introduction

4.2 Materials and methods

4.3 Results

4.3.1 Discussion

5 Conclusion

In conclusion, the research described in this doctoral dissertation was conducted to identify significant ...

In the section 2, I show that

References

- Aagaard, K., Ma, J., Antony, K. M., Ganu, R., Petrosino, J., & Versalovic, J. (2014). The placenta harbors a unique microbiome. *Science translational medicine*, 6(237), 237ra65–237ra65.
- Alverdy, J., Hyoju, S., Weigerinck, M., & Gilbert, J. (2017). The gut microbiome and the mechanism of surgical infection. *Journal of British Surgery*, 104(2), e14–e23.
- Anderson, M. J. (2014). Permutational multivariate analysis of variance (permanova). *Wiley statsref: statistics reference online*, 1–15.
- Barlow, G. M., Yu, A., & Mathur, R. (2015). Role of the gut microbiome in obesity and diabetes mellitus. *Nutrition in clinical practice*, 30(6), 787–797.
- Basavaprabhu, H., Sonu, K., & Prabha, R. (2020). Mechanistic insights into the action of probiotics against bacterial vaginosis and its mediated preterm birth: An overview. *Microbial pathogenesis*, 141, 104029.
- Berghella, V. (2012). Universal cervical length screening for prediction and prevention of preterm birth. *Obstetrical & gynecological survey*, 67(10), 653–657.
- Blencowe, H., Cousens, S., Oestergaard, M. Z., Chou, D., Moller, A.-B., Narwal, R., ... others (2012). National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *The lancet*, 379(9832), 2162–2172.
- Bolstad, A., Jensen, H. B., & Bakken, V. (1996). Taxonomy, biology, and periodontal aspects of fusobacterium nucleatum. *Clinical microbiology reviews*, 9(1), 55–71.
- Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., ... others (2019). Reproducible, interactive, scalable and extensible microbiome data science using qiime 2. *Nature biotechnology*, 37(8), 852–857.
- Breiman, L. (2001). Random forests. *Machine learning*, 45, 5–32.
- Brennan, C. A., & Garrett, W. S. (2019). Fusobacterium nucleatum—symbiont, opportunist and oncobacterium. *Nature Reviews Microbiology*, 17(3), 156–166.
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). Dada2: High-resolution sample inference from illumina amplicon data. *Nature methods*, 13(7), 581–583.
- Castaner, O., Goday, A., Park, Y.-M., Lee, S.-H., Magkos, F., Shiow, S.-A. T. E., & Schröder, H. (2018). The gut microbiome profile in obesity: a systematic review. *International journal of endocrinology*,

2018(1), 4095789.

- Chen, T., Yu, W.-H., Izard, J., Baranova, O. V., Lakshmanan, A., & Dewhirst, F. E. (2010). The human oral microbiome database: a web accessible resource for investigating oral microbe taxonomic and genomic information. *Database*, 2010.
- Chen, X., D'Souza, R., & Hong, S.-T. (2013). The role of gut microbiota in the gut-brain axis: current challenges and perspectives. *Protein & cell*, 4, 403–414.
- Cullin, N., Antunes, C. A., Straussman, R., Stein-Thoeringer, C. K., & Elinav, E. (2021). Microbiome and cancer. *Cancer Cell*, 39(10), 1317–1341.
- Doyle, R., Alber, D., Jones, H., Harris, K., Fitzgerald, F., Peebles, D., & Klein, N. (2014). Term and preterm labour are associated with distinct microbial community structures in placental membranes which are independent of mode of delivery. *Placenta*, 35(12), 1099–1101.
- Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological conservation*, 61(1), 1–10.
- Fettweis, J. M., Serrano, M. G., Brooks, J. P., Edwards, D. J., Girerd, P. H., Parikh, H. I., ... others (2019). The vaginal microbiome and preterm birth. *Nature medicine*, 25(6), 1012–1021.
- Francescone, R., Hou, V., & Grivennikov, S. I. (2014). Microbiome, inflammation, and cancer. *The Cancer Journal*, 20(3), 181–189.
- Gilbert, J. A., Blaser, M. J., Caporaso, J. G., Jansson, J. K., Lynch, S. V., & Knight, R. (2018). Current understanding of the human microbiome. *Nature medicine*, 24(4), 392–400.
- Goldenberg, R. L., Culhane, J. F., Iams, J. D., & Romero, R. (2008). Epidemiology and causes of preterm birth. *The lancet*, 371(9606), 75–84.
- Goodyear, M. D., Krleza-Jeric, K., & Lemmens, T. (2007). *The declaration of helsinki* (Vol. 335) (No. 7621). British Medical Journal Publishing Group.
- Hajishengallis, G. (2015). Periodontitis: from microbial immune subversion to systemic inflammation. *Nature reviews immunology*, 15(1), 30–44.
- Hamming, R. W. (1950). Error detecting and error correcting codes. *The Bell system technical journal*, 29(2), 147–160.
- Han, Y. W. (2015). Fusobacterium nucleatum: a commensal-turned pathogen. *Current opinion in microbiology*, 23, 141–147.
- Han, Y. W., & Wang, X. (2013). Mobile microbiome: oral bacteria in extra-oral infections and inflammation. *Journal of dental research*, 92(6), 485–491.
- Hartstra, A. V., Bouter, K. E., Bäckhed, F., & Nieuwdorp, M. (2015). Insights into the role of the microbiome in obesity and type 2 diabetes. *Diabetes care*, 38(1), 159–165.
- Helmink, B. A., Khan, M. W., Hermann, A., Gopalakrishnan, V., & Wargo, J. A. (2019). The microbiome, cancer, and cancer therapy. *Nature medicine*, 25(3), 377–388.
- Honda, K., & Littman, D. R. (2012). The microbiome in infectious disease and inflammation. *Annual review of immunology*, 30(1), 759–795.
- Honest, H., Forbes, C., Durée, K., Norman, G., Duffy, S., Tsourapas, A., ... others (2009). Screening to prevent spontaneous preterm birth: systematic reviews of accuracy and effectiveness literature with

- economic modelling. *Health Technol Assess*, 13(43), 1–627.
- Hong, Y. M., Lee, J., Cho, D. H., Jeon, J. H., Kang, J., Kim, M.-G., ... Kim, J. K. (2023). Predicting preterm birth using machine learning techniques in oral microbiome. *Scientific Reports*, 13(1), 21105.
- Huang, R.-Y., Lin, C.-D., Lee, M.-S., Yeh, C.-L., Shen, E.-C., Chiang, C.-Y., ... Fu, E. (2007). Mandibular disto-lingual root: a consideration in periodontal therapy. *Journal of periodontology*, 78(8), 1485–1490.
- Iams, J. D., & Berghella, V. (2010). Care for women with prior preterm birth. *American journal of obstetrics and gynecology*, 203(2), 89–100.
- Ide, M., & Papapanou, P. N. (2013). Epidemiology of association between maternal periodontal disease and adverse pregnancy outcomes—systematic review. *Journal of clinical periodontology*, 40, S181–S194.
- John, G. K., & Mullin, G. E. (2016). The gut microbiome and obesity. *Current oncology reports*, 18, 1–7.
- Katz, J., Chegini, N., Shiverick, K., & Lamont, R. (2009). Localization of *p. gingivalis* in preterm delivery placenta. *Journal of dental research*, 88(6), 575–578.
- Kelly, B. J., Gross, R., Bittinger, K., Sherrill-Mix, S., Lewis, J. D., Collman, R. G., ... Li, H. (2015). Power and sample-size estimation for microbiome studies using pairwise distances and permanova. *Bioinformatics*, 31(15), 2461–2468.
- Kim, C. H. (2018). Immune regulation by microbiome metabolites. *Immunology*, 154(2), 220–229.
- Kinane, D. F., Stathopoulou, P. G., & Papapanou, P. N. (2017). Periodontal diseases. *Nature reviews Disease primers*, 3(1), 1–14.
- Kindinger, L. M., Bennett, P. R., Lee, Y. S., Marchesi, J. R., Smith, A., Caciato, S., ... MacIntyre, D. A. (2017). The interaction between vaginal microbiota, cervical length, and vaginal progesterone treatment for preterm birth risk. *Microbiome*, 5, 1–14.
- Kogut, M. H., Lee, A., & Santin, E. (2020). Microbiome and pathogen interaction with the immune system. *Poultry science*, 99(4), 1906–1913.
- Leitich, H., & Kaider, A. (2003). Fetal fibronectin—how useful is it in the prediction of preterm birth? *BJOG: An International Journal of Obstetrics & Gynaecology*, 110, 66–70.
- León, R., Silva, N., Ovalle, A., Chaparro, A., Ahumada, A., Gajardo, M., ... Gamonal, J. (2007). Detection of *porphyromonas gingivalis* in the amniotic fluid in pregnant women with a diagnosis of threatened premature labor. *Journal of periodontology*, 78(7), 1249–1255.
- Lim, J. W., Park, T., Tong, Y. W., & Yu, Z. (2020). The microbiome driving anaerobic digestion and microbial analysis. In *Advances in bioenergy* (Vol. 5, pp. 1–61). Elsevier.
- Lloyd-Price, J., Abu-Ali, G., & Huttenhower, C. (2016). The healthy human microbiome. *Genome medicine*, 8, 1–11.
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome biology*, 15, 1–21.
- Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically

- larger than the other. *The annals of mathematical statistics*, 50–60.
- Mayer, E. A., Tillisch, K., Gupta, A., et al. (2015). Gut/brain axis and the microbiota. *The Journal of clinical investigation*, 125(3), 926–938.
- Offenbacher, S., Katz, V., Fertik, G., Collins, J., Boyd, D., Maynor, G., ... Beck, J. (1996). Periodontal infection as a possible risk factor for preterm low birth weight. *Journal of periodontology*, 67, 1103–1113.
- Papapanou, P. N., Sanz, M., Buduneli, N., Dietrich, T., Feres, M., Fine, D. H., ... others (2018). Periodontitis: Consensus report of workgroup 2 of the 2017 world workshop on the classification of periodontal and peri-implant diseases and conditions. *Journal of periodontology*, 89, S173–S182.
- Payne, M. S., Newnham, J. P., Doherty, D. A., Furfaro, L. L., Pendal, N. L., Loh, D. E., & Keelan, J. A. (2021). A specific bacterial dna signature in the vagina of australian women in midpregnancy predicts high risk of spontaneous preterm birth (the predict1000 study). *American journal of obstetrics and gynecology*, 224(2), 206–e1.
- Peirce, J. M., & Alviña, K. (2019). The role of inflammation and the gut microbiome in depression and anxiety. *Journal of neuroscience research*, 97(10), 1223–1241.
- Romero, R., Dey, S. K., & Fisher, S. J. (2014). Preterm labor: one syndrome, many causes. *Science*, 345(6198), 760–765.
- Romero, R., Hassan, S. S., Gajer, P., Tarca, A. L., Fadrosh, D. W., Nikita, L., ... others (2014). The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. *Microbiome*, 2, 1–19.
- Schwabe, R. F., & Jobin, C. (2013). The microbiome and cancer. *Nature Reviews Cancer*, 13(11), 800–812.
- Sepich-Poore, G. D., Zitvogel, L., Straussman, R., Hasty, J., Wargo, J. A., & Knight, R. (2021). The microbiome and human cancer. *Science*, 371(6536), eabc4552.
- Sharma, S., & Tripathi, P. (2019). Gut microbiome and type 2 diabetes: where we are and where to go? *The Journal of nutritional biochemistry*, 63, 101–108.
- Sotiriadis, A., Papatheodorou, S., Kavvadias, A., & Makrydimas, G. (2010). Transvaginal cervical length measurement for prediction of preterm birth in women with threatened preterm labor: a meta-analysis. *Ultrasound in Obstetrics and Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology*, 35(1), 54–64.
- Spss, I., et al. (2011). Ibm spss statistics for windows, version 20.0. *New York: IBM Corp*, 440, 394.
- Stout, M. J., Conlon, B., Landeau, M., Lee, I., Bower, C., Zhao, Q., ... Mysorekar, I. U. (2013). Identification of intracellular bacteria in the basal plate of the human placenta in term and preterm gestations. *American journal of obstetrics and gynecology*, 208(3), 226–e1.
- Thaiss, C. A., Zmora, N., Levy, M., & Elinav, E. (2016). The microbiome and innate immunity. *Nature*, 535(7610), 65–74.
- Tilg, H., Kaser, A., et al. (2011). Gut microbiome, obesity, and metabolic dysfunction. *The Journal of clinical investigation*, 121(6), 2126–2132.
- Tonetti, M. S., Greenwell, H., & Kornman, K. S. (2018). Staging and grading of periodontitis: Framework

- and proposal of a new classification and case definition. *Journal of periodontology*, 89, S159–S172.
- Ursell, L. K., Metcalf, J. L., Parfrey, L. W., & Knight, R. (2012). Defining the human microbiome. *Nutrition reviews*, 70(suppl_1), S38–S44.
- Vander Haar, E. L., So, J., Gyamfi-Bannerman, C., & Han, Y. W. (2018). Fusobacterium nucleatum and adverse pregnancy outcomes: epidemiological and mechanistic evidence. *Anaerobe*, 50, 55–59.
- Whiteside, S. A., Razvi, H., Dave, S., Reid, G., & Burton, J. P. (2015). The microbiome of the urinary tract—a role beyond infection. *Nature Reviews Urology*, 12(2), 81–90.
- Witkin, S. (2019). Vaginal microbiome studies in pregnancy must also analyse host factors. *BJOG: An International Journal of Obstetrics & Gynaecology*, 126(3), 359–359.
- Wong, T.-T., & Yeh, P.-Y. (2019). Reliable accuracy estimates from k-fold cross validation. *IEEE Transactions on Knowledge and Data Engineering*, 32(8), 1586–1594.
- Yang, I., Claussen, H., Arthur, R. A., Hertzberg, V. S., Geurs, N., Corwin, E. J., & Dunlop, A. L. (2022). Subgingival microbiome in pregnancy and a potential relationship to early term birth. *Frontiers in cellular and infection microbiology*, 12, 873683.
- Zhu, X., Han, Y., Du, J., Liu, R., Jin, K., & Yi, W. (2017). Microbiota-gut-brain axis and the central nervous system. *Oncotarget*, 8(32), 53829.

Acknowledgments

I would like to disclose my earnest appreciation for my advisor, Professor Semin Lee, who provided solicitous supervision and cherished opportunities throughout the course of my research. His advice and consultation encouraged me to become as a researcher and to receive all humility and gentleness. I am also grateful to all of my committee members, Professor AAA, Professor BBB, Professor CCC, and Professor DDD, for their critical and meaningful mentions and suggestions.

I extend my deepest gratitude to my Lord, *the Flying Spaghetti Monster*, His Noodly Appendage has guided me through the twist and turns of this academic journey. His presence, ever comforting and mysterious, has been a source of strength and humor during both highs and lows. In moments of doubt, I found solace in the belief that you were there, gently reminding me to keep faith in the process. His Holy Noodle has nourished my mind, and for that, I am truly overwhelmed. May His Holy Noodle continue to guide me in all my future endeavors. R’Amen.

(Professors)

I would like to extend my heartfelt gratitude to my colleagues of the Computational Biology Lab @ UNIST, whose collaboration, friendship, brotherhood, and support have been an invaluable part of my journey. Your willingness to share insights, engage in thoughtful discussions, and offer encouragement during the challenging moments of research has significantly shaped my academic experience. The camaraderie in Computational Biology Lab made even the most demanding days more enjoyable, and I am deeply grateful for the collaborative environment we created together. I appreciate you for standing by my side throughout this Ph.D. journey.

I would like to express my heartfelt gratitude to my family, whose unwavering support has been the foundation of everything I have achieved. Your love, encouragement, and belief in me have sustained me through every challenge, and I could not have come this far without you. From your words of wisdom to your patience and understanding, each of you has played a vital role in helping me navigate this journey. The strength and comfort I have drawn from our family bond have been my greatest source of resilience. Your presence, both near and far, has filled my life with warmth and motivation. I am deeply grateful for your unconditional love and for always being there when I needed you the most. Thank you for being my constant source of strength and inspiration.

I am incredibly pleased to my friends, especially my GSHS alumni (○망특), for their unwavering support and encouragement throughout this journey. The bonds we formed back in our school days have only grown stronger over the years, and I am fortunate to have had such loyal and understanding friends by my side. Your constant words of motivation, and even moments of levity during stressful times have helped keep me grounded. Whether it was a late-night conversations, a shared laugh, or a simple message of reassurance, you all have played a vital role in keeping me focused and motivated. I am relieved for the ways you celebrated each small achievement with me and how you patiently listened to my worries. The memories of our shared past provided me with comfort and a sense of stability when the road ahead felt uncertain. I could not have reached this point without the love and friendship that you all have generously given. Each of your, in your unique way, has contributed to this dissertation, even if indirectly, and for

that, I am forever beholden. I look forward to continuing our friendship as we all grow in our individual paths, knowing that the support we share is something truly special.

I would like to express my sincere gratitude to the amazing members of my animal protection groups, DRDR (두루두루) and UNIMALS (유니멀스), whose dedication and compassion have been a constant source of motivation. Your unwavering commitment to improving the lives of animals has inspired me throughout this journey. I am also thankful for the beautiful cats we have cared for, whose presence brought both joy and purpose to our allegiance. Their playful spirits and gentle companionship served as daily reminders of why we continue to fight for animal rights. The bond we share, both with each other and with the animals we protect, has enriched my life in countless ways. I appreciate you all again for your support, dedication, and for being part of this meaningful cause.

I would like to express my deepest gratitude to everyone I have had the honor of meeting throughout this journey. Your kindness, encouragement, and support have carried me through both the challenging and rewarding moments of my life. Whether through a kind word, thoughtful advice, or simply being there when I needed it most, your presence has made all the difference. I am incredibly fortunate to have received such generosity and warmth from those around me, and I do not take it for granted. Every act of kindness, no matter how big or small, has been a source of strength and motivation for me. To all my friends, colleagues, mentors, and beloved ones, thank you for your unwavering support. I am truly grateful for each of you, and your kindness has left an indelible mark on my journey.

My Lord, *the Flying Spaghetti Monster*,
give us grace to accept with serenity the things that cannot be changed,
courage to change the things that should be changed,
and the wisdom to distinguish the one from the other.

Glory be to *the Meatball*, to *the Sauce*, and to *the Holy Noodle*.

As it was in the beginning, is now, and ever shall be.

R'Amen.

