이학박사 학위논문

Isolation and genomic characterization of *Escherichia coli* enriched within the tissues of oral lichen planus lesions

구강편평태선 병소 조직 내에서 증가한 *Escherichia coli* 분리 및 유전자 동정

2019년 2월

서울대학교 대학원

치의과학과 면역 및 분자미생물학 전공

백 금 진

ABSTRACT

Isolation and genomic characterization of *Escherichia coli* enriched within the tissues of oral lichen planus lesions

Keumjin Baek

Program in Immunology and Molecular Microbiology in Dentistry

Department of Dental Science

The Graduate School

Seoul National University

Background

Oral lichen planus (OLP) is a T cell-mediated chronic mucocutaneous disease by unknown etiology. Many researchers have suggested that the potential antigens including autoantibodies, dental materials, and infectious agents are

recognized by $CD8^+$ T cells. However, the precise pathogenesis of OLP is not understood.

Many research groups have been focused on the relationship between the bacterial infection and the onset of OLP. Recently, advent of sequencing technologies for analysis of large amount bacterial communities allow researcher to study bacterial profiles using small amount DNA samples. It has been reported that the microbiota of saliva and buccal mucosa showed the differences between healthy controls and OLP patients. However, the bacterial communities within the tissues of OLP lesions have not been studied. The present study aimed to characterize the oral microbiota within the tissues of OLP lesions compared with the mucosal surface and to characterize the genome of the bacterial species enriched within OLP lesions.

Methods

To compare the oral microbiota, the buccal mucosal bacteria (OM) or biopsies (OT) of OLP lesions were obtained from 10 patients with OLP. The mucosal bacteria (HM) from 5 healthy subjects were additionally used. The bacterial genomic DNA was extracted from the mucosal samples and tissues. In case of the tissue samples, the tissues were treated with lysozyme, antibiotics, and DNase I before extraction of genomic DNA to remove the surface bacteria on

the tissues. To analyze the microbiota of mucosal surface and tissues of OLP lesions, DNA fragments including V1-V3 or V3-V4 hypervariable regions of the bacterial 16S rRNA gene were amplified by PCR, and then the PCR product were sequenced using a 454 GS FLX Titanium or an Illumina MiSeq sequencing system. For the analysis, the sequencing data of mucosal samples from 11 healthy individuals and 13 OLP patients in a previous study were additionally used.

Four clinical isolates of *Escherichia coli* (K12-5.1, 7.1, 7.2, and 7.3) were obtained from 2 OLP patients and then subjected to whole genome sequencing.

For the preliminary animal experiment to develop the animal model for OLP, ICR mice were orally inoculated with viable *E. coli* K12-7.2. Then the bacterial invasion and infiltration of immune cells within the tongue tissues were confirmed by *in situ* hybridization and H&E stain.

Results

The bacterial loads of OM (n = 7) communities were significantly increased compared with HM (n = 5) and OT (n = 9) communities. The bacterial richness were different between OM (n = 20) and OT (n = 9). In terms of bacterial composition, 3 phyla, 13 genera, and 90 species/phylotypes showed significant differences in the relative abundance among groups.

In the analysis of microbiota between mucosa and tissues from same sites of 7 OLP patients, the bacterial richness and diversity were not different between 2 groups. However, the relative abundances of 2 phyla, 8 genera, and 52 species/phylotypes were totally different between OM and OT. Among 52 species differently distributed between OM and OT, *E. coli* was highly enriched in the tissue communities.

Whole-genome sequencing analysis revealed that 4 clinical isolates of *E. coli* phylogenetically belonged to *E. coli* K-12 strains. However, the gene contents of the 4 isolates were partially different from those of K-12 reference strain MG1655 isolated from human feces.

In the animal experiment, the bacterial invasion and immune infiltration within the tongue tissues were compared between sham and *E.coli* inoculated groups. However, there were no significant differences in the bacterial invasion between groups. In addition, the immune cell infiltration was not observed in *E. coli* inoculated group.

Conclusion

In conclusion, this is the first study to identify the bacterial communities within the tissues of OLP lesions compared with those of the mucosal surface. Understanding the role of *E. coli* enrichment within the tissues in the

pathogenesis of OLP may provide a new insight into the etiopathogenesis of OLP.

Keywords: Oral lichen planus, Bacterial community, *Escherichia coli*, Comparative genomics

***Student number*: 2016-30620**

CONTENTS

Abstract

Contents

I. Introduction 1

1. Oral lichen planus (OLP) 1

1.1. Oral lichen planus: Definition, epidemiology, histopathologic features, and treatment 1

1.2. Etiology of oral lichen planus 2

1.2.1. Stress 2

1.2.2. Genetic background 2

1.2.3. Microbial infections 3

1.3. Pathogenesis of oral lichen planus 4

1.3.1. Antigen-specific mechanism 4

1.3.2. Non-specific mechanism 5

1.3.3. Humoral immune response 6

1.3.4. Autoimmune response 6

1.4. Bacterial communities in OLP 6

1.5. Bacterial invasion into OLP lesions 7

2. Pilot study: Differences in the microbial communities between the subgingival plaque and intratissue in periodontitis 8

3. Aims of this study 20

II. Materials and Methods 21

1. Study samples 21

2. Preparation of bacterial DNA samples 22

3. Analysis of bacterial communities 22

4. Real-time PCR 23

5. Isolation of *Escherichia coli* from OLP biopsies 23

6. Confirmation of *E. coli* by PCR and sequencing of 16S rRNA gene 24

7. Whole genome sequencing 24

8. Animal experiment 25

9. *In situ* hybridization 26

10. Statistical analysis 28

III. Results 29

1. Study population 29

2. Characteristics of bacterial communities in OLP 30

2.1. Dysbiosis of oral microbiota in the patients with OLP 30

2.2. Altered bacterial composition in OLP 34

3. Features of intratissue bacterial communities in OLP 40

3.1. Alpha and beta diversities of mucosal surface and intratissue communities within the OLP lesions 40

3.2. Differences in oral microbiota composition between mucosa 42

and tissue in OLP

4. Isolation and confirmation of *E. coli* within the tissues from OLP lesions 48

5. Genomic characterization of *E. coli* isolates by whole genome sequencing 49

5.1. General genomic features of *E. coli* isolates 49

5.2. Comparison of gene contents among *E. coli* isolates 50

5.3. Phylogenetic relationships between strains of *E. coli* group 53

5.4. Comparative genomic analysis of *E. coli* isolates 54

6. Development of animal model for OLP 60

IV. Discussion 63

V. Reference 70

국문초록

I. Introduction

1. Oral lichen planus (OLP)

1.1. Oral lichen planus: Definition, epidemiology, histopathologic features, and treatment

Oral lichen planus (OLP) is the oral variant of lichen planus (LP), which is a chronic inflammatory disease affecting the nails, skin, eyes, urinary tract, larynx, and oral mucosa. OLP occurs in 1-4% of the global population (1), and appears more prevalent in middle-aged female than male at a ratio of 1.5:1 (2). However, OLP can also affect children and young adults. OLP shows a bilateral and symmetric distribution with a lace-like network of gray-white lines (reticular pattern). Erosive, papular, bullous, plaque-like, and atrophic lesions can also appear in the presence of reticular lesions. Histological criteria of OLP first described by Dubreuill in 1906 and later by Shklar includes the band-like lymphocytic infiltration confined to the superficial part of the connective tissue, absence of epithelial dysplasia, and liquefaction degeneration in the basal cell layer. For the treatment of OLP, corticosteroid is considered as an effective agent. However, corticosteroid diminish the lesion, but not permanent cure (3).

1.2. Etiology of oral lichen planus

Although the major causing factor of OLP is not fully understood, several factors have been implicated as a predisposing factor.

1.2.1. Stress

Stress is widely accepted as an etiological factor in the pathogenesis of OLP, because the signs of depression and anxiety are commonly found in patients with OLP in comparison with controls. However, few studies have been demonstrated the relationship between stress and OLP (4).

1.2.2. Genetic background

Studies about the relationship between genetic background and the pathogenesis of OLP have been reported. In this context, genetic polymorphisms of some cytokines may play a role in the pathogenesis of OLP. It has been reported that genetic polymorphisms in the first intron of promoter gene of interferon-gamma (IFN-γ) and tumor necrosis factor-alpha (TNF-α) gene may contribute to OLP pathogenesis (5). In addition, the relationship between human leukocyte antigen (HLA) and OLP has been reported (6-10).

1.2.3. Microbial infections

Many studies have been suggested that infectious agents including virus, fungi, and bacteria are associated with OLP. Despite these efforts, a specific microbial agent involved in the pathogenesis of OLP has not been identified. Human papillomavirus (HPV), Hepatitis C virus (HCV), and Epstein-Barr virus (EBV) have been widely studied to verify association with onset of OLP. HPV can infect oral epithelia and may provide an antigens to T cells, whereas HCV increase pro-inflammatory cytokine TNF-α in serum that contribute to the development of OLP (11). However, EBV have been shown no relationship with OLP (12-14).

About 100 fungal species are identified in the oral cavity of healthy individuals (15). Among them, *Candida* species are frequently found in healthy subjects as an normal flora (16-18). A relationship between *Candida* and OLP has been suggested since 1980 (19). However, it is difficult to conclude that fungal infections are associated with OLP because the high prevalence of *Candida* in OLP patients is controversial (20, 21).

After detection of *Helicobacter pylori* in saliva (22), two groups showed the presence of *H. pylori* DNA in OLP biopsies by nested PCR. However, most studies showed no association (23-25). Bornstein et al. reported that increased colonization of several species, including *Capnocytophaga sputigena, Mobiluncus curtisii, Eikenella corrodens*, and *Prevotella intermedia* on OLP

lesions using the checkerboard DNA-DNA hybridization method (26). Ertugrul et al. also reported increased prevalence of periodontopathogens, including *Treponema denticola, P. intermedia, Porphyromonas gingivalis, Aggregatibacter actinomycetemcomitans* in patients with OLP by a PCR-based method (27).

Recently, the high-throughput sequencing technology is used to study the profile of the microbiome associated with OLP. Four research groups have studied oral microbiota associated with OLP by sequencing of 16S rRNA gene (1, 28-30). These results showed a relationship between dysbiosis of oral bacterial communities and OLP.

1.3. Pathogenesis of oral lichen planus

Because the precise etiology of OLP is not clear, the pathogenesis of OLP is also not fully understood. From accumulating evidence, many researchers have suggested a various mechanism involved in the pathogenesis of OLP. Generally, OLP is considered a T-cell mediated inflammatory disease.

1.3.1. Antigen-specific mechanism

Even though antigen related with OLP is still unidentified, self- peptides or heat

shock proteins expressed by keratinocyte have been considered as unknown antigens for OLP. Antigens presented by MHC class II and MHC class I to T cells can activate $CD4^+$ helper and $CD8^+$ cytotoxic T lymphocyte, respectively (31). The activated helper T cells can secrete interleukin (IL)-12 and IFN-γ, which can activate $CD8^+$ cytotoxic T cells and promote apoptosis of basal keratinocytes (32). This results in the liquefaction degeneration of basal epithelial cells observed in OLP lesions.

1.3.2. Non-specific mechanism

Non-specific mechanisms include matrix metalloproteinase (MMP), chemokines, mast cell degranulation, and chymase. The main function of MMPs is degradation of matrix proteins in connective tissue by proteolytic activity. Activating MMP-9 can disrupt the basement membrane (33). Chymase released from degranulating mast cell can promote the secretion of RANTES (regulated on activation, normal T cell expressed and secreted, also called as CCL5) from OLP lesional T cells and activate the MMP-9 (34). Chemokines are a family of small cytokine and have chemotactic activity. Among various chemokines, RANTES may lead to attract the mast cells into the developing OLP lesion and stimulate the degranulation of mast cells (35).

1.3.3. Humoral immune response

Several autoantibodies have been identified in serum of patients with OLP. A high frequencies of autoantibodies, including anti-nuclear (ANA), anti-thyroglobulin (TGA), anti-gastric parietal cell (GPCA), anti-thyroid microsomal (TMA), anti-smooth muscle (SMA), anti-mitochondrial (AMA), anti-desmogleins 1 and 3 autoantibodies, appear in OLP patients than those in control subjects (36-38).

1.3.4. Autoimmune response

OLP is considered as an autoimmune disease, because many autoimmune features, including female preference, disease chronicity, and the diminished immunosuppressive activity appear in patients with OLP (39).

1.4. Bacterial communities in OLP

As mentioned above, many studies have shown the relationship between microbial infection and OLP, however specific causative agent is not clarified. To date, four groups have studied oral microbiota associated with OLP by sequencing of the 16S rRNA gene (1, 28-30). In spite of the differences in sampled sites, sequencing method, populations, and targeted sequence of 16S

rRNA gene among four groups, several common findings were emerged. First, the bacterial diversity was increased in OLP. Second, the altered bacterial communities were observed in OLP patients. A principal coordinates analysis showed a segregation of bacterial communities by anatomical site (*i.e*., saliva vs buccal mucosa) and OLP status. In addition, a decreased relative abundance of *Streptococcus* genus and an increased relative abundance of *Leptotrichia* were commonly observed in patients with OLP (40).

Although some oral bacteria can invade into the tissues and have a role in the infiltration of immune cells within OLP lesions, differences in bacterial communities between the mucosal surface and intratissue of OLP lesions have not been studied.

1.5. Bacterial invasion into OLP lesions

Accumulating evidences have shown the presence of bacteria within the OLP lesions by nested PCR, DNA-DNA hybridization method, and immunohistochemistry (26, 41). Recently, the bacterial invasion within the OLP lesions was confirmed by *in situ* hybridization using a universal probe targeting the bacterial 16S rRNA gene. The bacterial invasion into lamina propria was highly increased in OLP compared with control tissues, whereas it was not different in the epithelia. The infiltration of T cells in OLP tissues was associated with an increased bacterial invasion within the lamina propria.

Furthermore, bacteria were detected within the infiltrated T cells in OLP tissues, and induced T cell chemokines. Therefore, the intracellular bacteria within the epithelial cells and T cells may be provided to the infiltrated T cells as target antigens. These results suggest that the intracellular bacteria within the OLP tissues may have a role in the pathogenesis of OLP (28).

2. Pilot study: Differences in the microbial communities between the subgingival plaque and intratissue in periodontitis

The human oral cavity contains a different habitats, including the subgingival sulcus, teeth, buccal mucosa, palate, and tongue, which provides a perfect niche for colonizing of bacteria. It is known that more than 1,000 bacterial species are found in the oral cavity (42). Colonizing bacteria on one area of the oral cavity have a chance to spread into neighboring sites. A number of bacteria can cause various oral diseases such as caries, and periodontitis (43, 44). In addition, it is known that periodontal disease is related to OLP (27, 45).

Among these diseases, periodontitis is a common chronic inflammatory disease resulted from the dysbiosis of subgingival microbiome. Generally, periodontitis is found in 5-15% of the world populations and is the 6th most prevalent disease (46). In addition, the overall prevalence of this disease

increases with age. It leads to the alveolar bone loss, destruction of tooth-supporting tissue, and tooth loss (47). It can also affects the systemic disease such as rheumatoid arthritis, diabetes, and osteoporosis (48, 49). Among the subgingival bacteria, *P. gingivalis*, *T. denticola*, and *Tannerella forsythia* included in so called the 'red-complex group' are highly associated with chronic periodontitis (47, 50). Moreover, theses bacteria have shown the ability to invade various host cells such as gingival epithelial cells, gingival fibroblasts, and endothelial cells (51-57).

In the murine model of periodontitis, the oral inoculation of *P. gingivalis* have shown the increased bacteria within gingival tissues (58). In addition, higher levels of bacteria were detected in the gingival tissues of periodontal lesions than in healthy sites (59). Furthermore, the bacterial invasion within the gingival tissues and alveolar bone loss were positively correlated with T cell infiltration in the murine models of periodontitis (57). Therefore, the bacterial invasion into the gingival tissues is a key event in the pathogenesis of periodontitis.

The general features of subgingival microbiota have been established well. However, the characteristics of bacterial communities within the gingival tissues of periodontal lesion have not been studied.

In the previous study, subgingival plaque and gingival tissue were obtained from seven patients with chronic periodontitis to characterize the bacterial

communities within the gingival tissue and compare them with subgingival plaque communities (60). Total bacterial loads within the gingival tissue were much lower than those of subgingival plaque estimated by quantitative real-time PCR of the 16S rRNA gene (Figure 1A).

A sequencing analysis of the 16S rRNA gene revealed that species richness determined by Chao 1 index and Shannon diversity index were not significantly different between the two groups (Figure 1B). In the UniFrac-based principal coordinates analysis, bacterial communities of tissue clustered in a small area, showing a significantly decreased UniFrac distance, and partially separated from the plaque communities (Figure 1C, D).

When the bacterial composition of the two groups compared, they had different bacterial communities. At the phylum level, the relative abundance of Fusobacteria and Chloroflexi was increased but the abundance of Firmicutes was decreased (Figure 2A). At the genus level, the relative abundance of *Fusobacteria, Porphyromonas, Actinobaculum*, and *GG703879_g* (Actinomycetaceae family) was increased in tissue, whereas the abundance of *Bulleidia*, *GQ422727_g* (Peptococcus family), and *Coribacteriaceae_uc* was decreased in tissue (Figure 2B). At the species level, the relative abundance of 24 species/phylotypes were differently distributed between subgingival plaque and tissues. Among these species, *Fusobacterium nucleatum* and *P. gingivalis* were highly enriched in the tissue (Figure 2C). These results shown that the bacterial communities within the gingival tissues are as complex as subgingival

plaque and distinct bacterial communities could contribute in the pathogenesis of periodontitis.

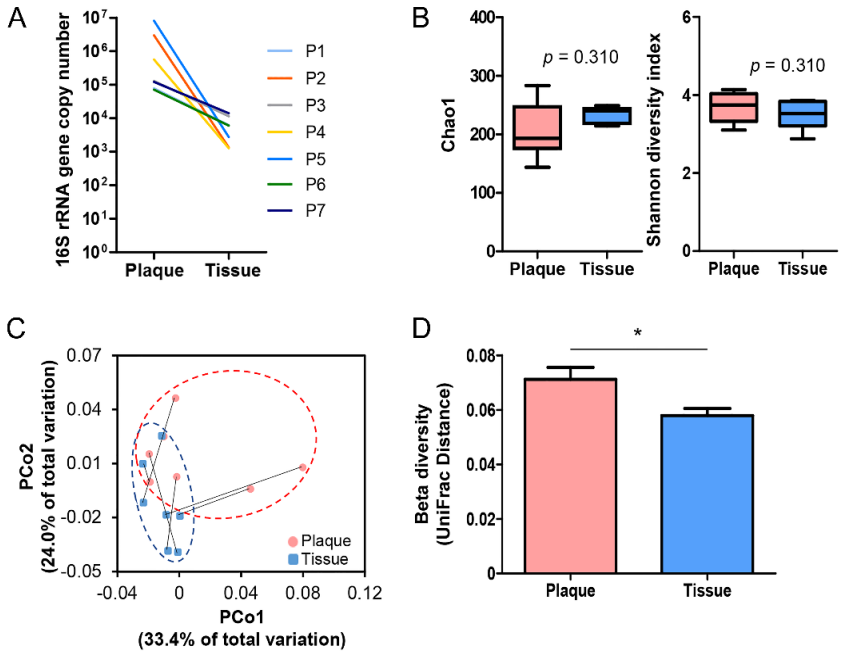


Figure 1. Alpha and beta diversities of subgingival plaque and intratissue communities within periodontal lesions (40). Subgingival plaque and gingival tissue samples were obtained from seven patients with chronic periodontitis. Bacterial genomic DNA was prepared and subjected to analysis by sequencing of the 16S rRNA gene. (A) Total bacterial load in each sample was estimated by real-time PCR using universal primers targeting the bacterial 16S rRNA gene. (B) The Chao1 and Shannon index are expressed using box and whisker plots (*p* by two-tailed Wilcoxon signed-rank test). (C) PCoA plot was generated using weighted UniFrac metric. Samples from the same subject

are connected with a solid line. (D) The inter-subject UniFrac distances of the subgingival plaque and intratissue communities were obtained using a weighted metric (*$p < 0.05$ by two tailed Wilcoxon signed-rank test).

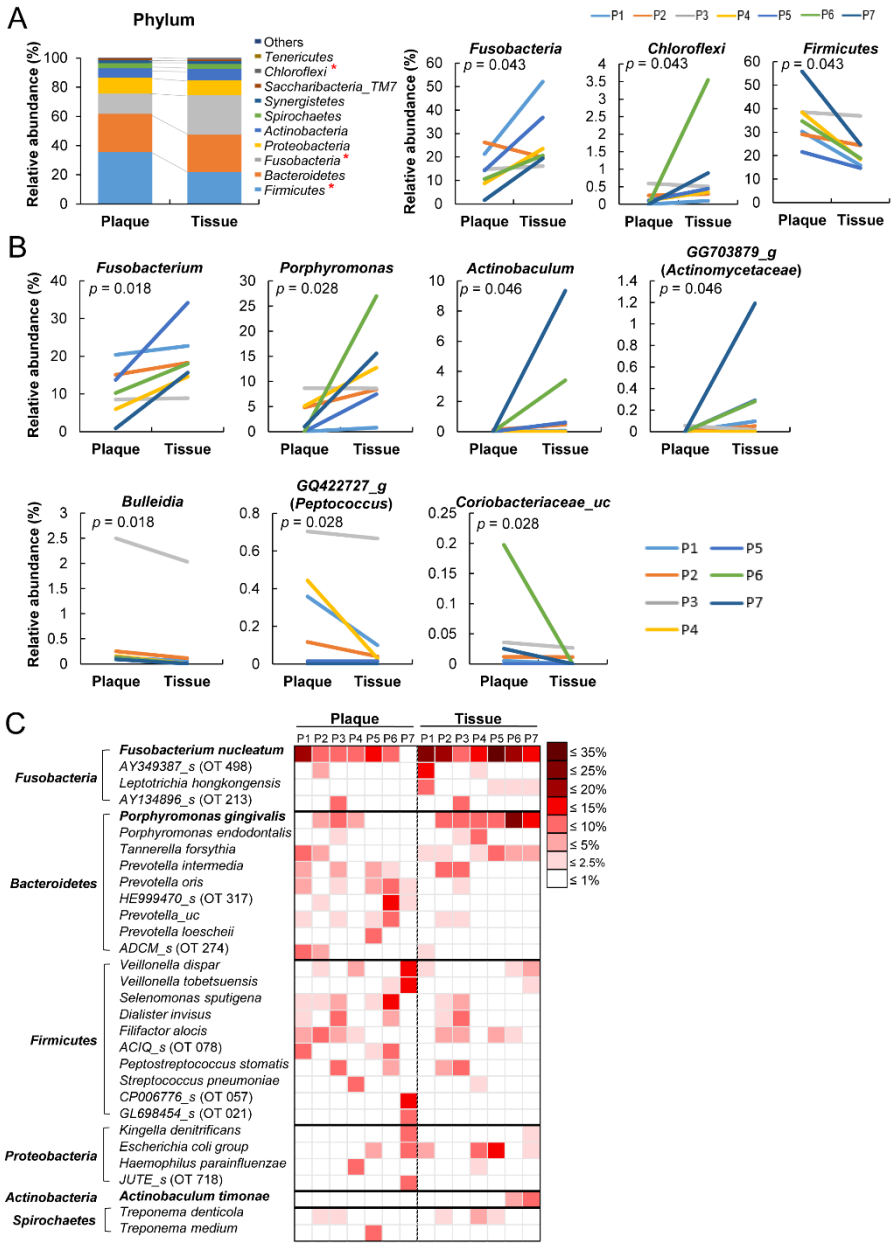


Figure 2. Differences in bacterial composition between the subgingival plaque and intratissue communities (40). The relative abundance of each taxon between the subgingival plaque and intratissue communities was

compared. (A) The members of top ten phyla are shown (left panel). Three phyla were differently distributed between the two communities. (B) Seven genera were differently distributed between the two communities (p by two-tailed Wilcoxon signed-rank test). (C) A heat map was generated for the species/phylotypes whose relative abundance was greater than 2.5% in any sample.

The distribution of the two highly abundant species *F. nucleatum* and *P. gingivalis* in gingival tissues was examined by *in situ* hybridization with specific probes. Both species were detected throughout the section from the pocket epithelium, connective tissue, and oral epithelium of all gingival biopsies but were particularly abundant at the area of inflammatory infiltration (Figure 3A-C).

Because the bacteria were often observed as aggregates of different sizes, this result provided a possibility that they form biofilms. To confirm this possibility, alcian blue staining was performed to visualize polysaccharides in the extracellular polymeric substances (EPSs) of biofilms. At low magnification, weak alcian blue staining was observed in almost half the dense connective tissue along the pocket epithelium, where abundant white spaces were observed in the hematoxylin and eosin-stained section. At high magnification, cobweb-like structures with bead-like bacterial clusters (arrows in Figure 3D) were readily observed in free spaces formed by degradation of connective tissue fibers but were rarely observed in the areas where fibers were relatively intact (asterisk in Figure 3D).

To determine associations between the biofilm and *F. nucleatum* or *P. gingivalis*, 3 areas with varying degrees of alcian blue staining were chosen from each sample, and the intensities of alcian blue staining and the signals of *F. nucleatum* and *P. gingivalis* were measured. Although the amounts of *F. nucleatum* and *P. gingivalis* tended to be positively correlated, suggesting

coexistence of the 2 species, the amount of neither species was associated with the amount of biofilm (Figure 3E).

The presence of biofilms within gingival tissues was further verified by atomic force microscopy (AFM). First, a piece of plaque biofilm co-embedded with tissues was examined. The central area of it was not stained with alcian blue, and it revealed tightly packed bacterial cells under AFM. The periphery of the plaque biofilm was stained with alcian blue, and it presented the cobweb-like structures with bacterial clusters under AFM (Figure 4A-C). Similar cobweb-like structures were also observed within tissues where collagen fibers were severely degraded (Figure 4D, E). In the areas where the fibers were intact, scattered bacterial cells adhering to collagen fibers were observed (Figure 4F).

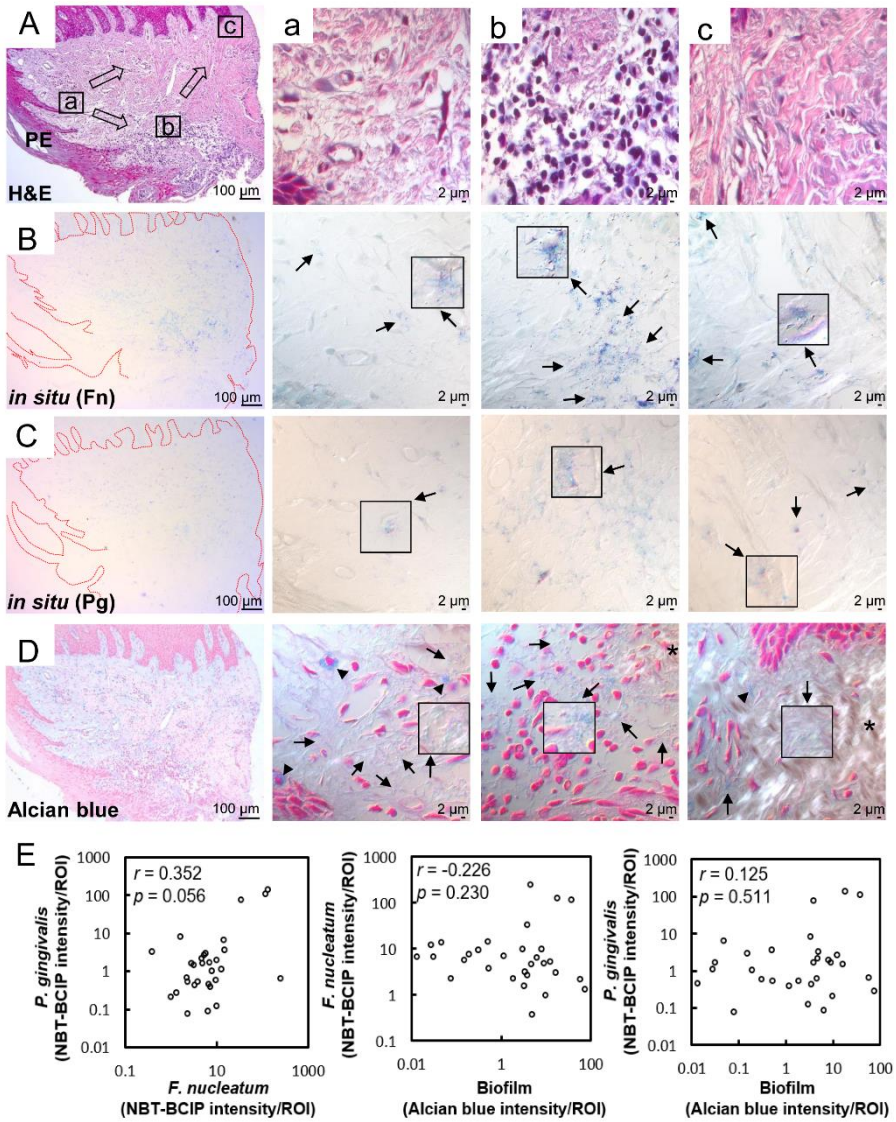


Figure 3. Distribution of *F. nucleatum*, *P. gingivalis*, and biofilm within the gingival tissue (40). (A) Tissues sections were stained with hematoxylin and eosin (H&E). Three areas (a, b, c) were examined under high magnification (x1000) with differential interference contrast microscopy. Arrows indicate the potential directions of infection spread. PE, pocket epithelium. (B, C) Tissues

sections were *in situ* hybridized with *F. nucleatum*- and *P. gingivalis*-specific probes, respectively. Arrows, bacterial signals; insets, areas with biofilm-like structure are magnified. (D) Tissues sections were stained with 1% alcian blue for acid mucopolysaccharide and counterstained with nuclear fast red. Arrows, biofilm-like structures; arrowheads, mast cells; asterisk, area with intact connective tissue fibers. (E) Correlation plots between bacterial signals and biofilm formation (*r* and *p* values by Spearman's rank correlation test). ROI, region of interest.

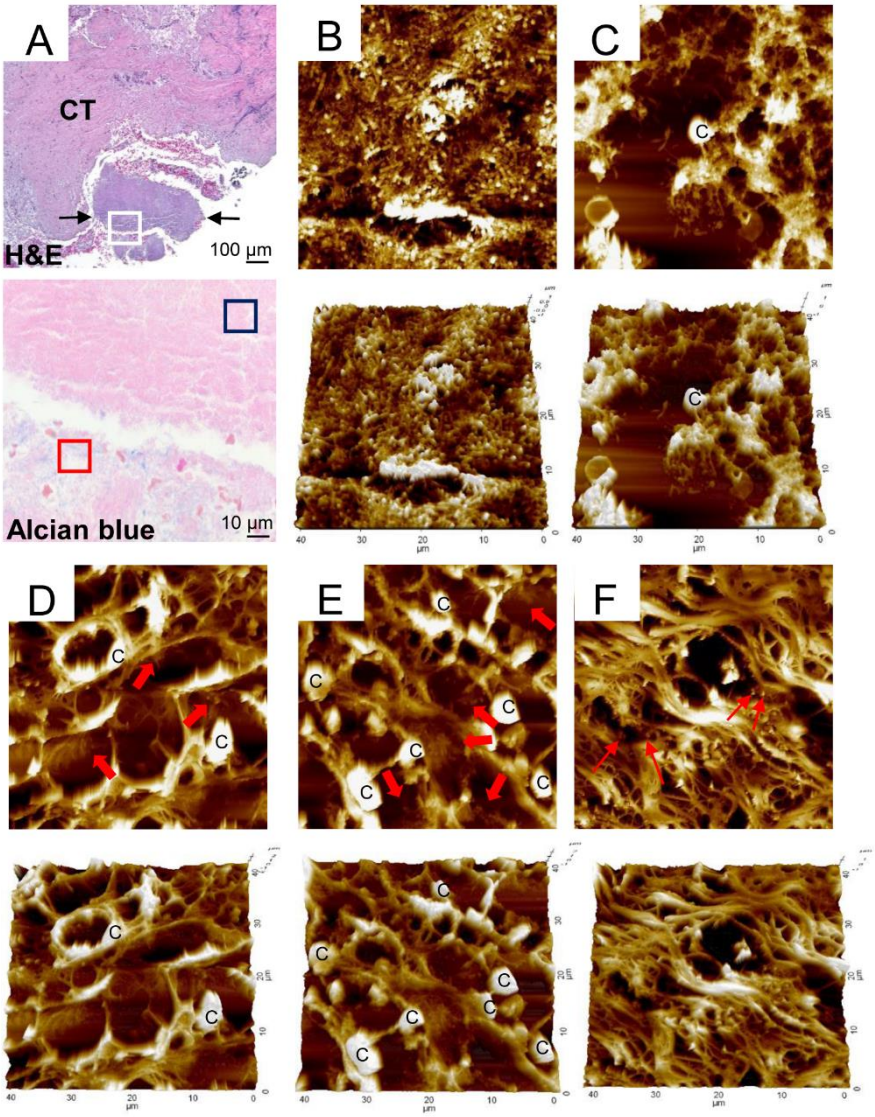


Figure 4. Atomic force microscopy examination of biofilm (40). (A) A piece of plaque biofilm co-embedded with tissue was located in a hematoxylin and eosin (H&E)-stained section (between 2 arrows in top panel). CT, connective tissue. Area corresponding to the white-boxed region in the H&E-stained section was examined in the serial section stained with alcian blue under high

magnification (x1000, bottom panel). Two typical areas were chosen based on alcian blue staining and examined by atomic force microscopy: (B) blue- and (C) red-boxed areas from panel A. Areas a-c (D-F, respectively) from Figure 3A were examined by atomic force microscopy. Thick arrows indicate biofilm-like structures within tissue. Thin arrows indicate scattered bacterial cells. C, eukaryotic cell.

3. Aims of this study

Although periodontitis and OLP do not share the same symptoms and histological features, both conditions occur in the oral cavity, which is exposed to numerous bacteria. In addition, these diseases arise from chronic inflammation within the host tissues. Based on the findings of previous studies, the present study hypothesized that the predominant species within the OLP tissues may have a key role in the pathogenesis of OLP. Therefore, this study aimed to characterize the bacterial communities located within the tissues of OLP lesions and to characterize the genome of the bacterial species enriched within these lesions.

II. Materials and Methods

1. Study samples

This study was performed in accordance with the Helsinki Declaration under approved procedures by the Institutional Review Board at the Seoul National University Dental Hospital (SNUDH) (IRB No. CRI 15023). An informed Consent was obtained from all 10 patients with OLP and 5 control subjects. All enrolled patients not a history of steroid treatment or antibiotics within the last month. Patients with > 0.1 ml/min unstimulated whole salivary flow rate and smokers were excluded. A semiquantitative REU (reticulation/erythema/ulceration) scoring system was adapted to examine the severity of the lesion.

For the bacterial sampling from the mucosa, a sterilized 20 mm x 20 mm polyvinylidene difluoride membrane was placed on the buccal mucosa of subjects for 30 seconds. A biopsy was performed on the reticular lesion with or without erythema but with no ulceration at the buccal mucosa of OLP patients. In addition, the mucosal bacteria and biopsy were obtained two times from OLP 4 patient during 2 years. Among 11 biopsy samples from OLP patients, two samples were not included for the sequencing but used to isolate the clinical *Escherichia coli* isolates.

2. Preparation of bacterial DNA samples

Bacterial genomic DNA was isolated from the buccal mucosal samples or tissues of subjects using a commercial kit for soil bacteria (MO BIO Laboratories). The tissues were washed with PBS and incubated with 1ml PBS containing lysozyme (300 μg/ml) and antibiotics (penicillin, streptomycin, and gentamicin) at 37℃ for 1 hour to remove the surface bacteria on tissues. After incubation, the tissues were washed with PBS and incubated with DNase I to digest bacterial DNA on the surface of tissues. Subsequently, the DNase I was inactivated by heating at 65℃ for 10 min and washed, the tissues were homogenized and subjected to bacterial DNA extraction. Bacterial DNA samples were stored at -80℃ until further analysis.

3. Analysis of bacterial communities

Genomic DNA of buccal mucosa and tissue samples was subjected to analyzing of bacterial communities. The V1-V3 (for OLP1, 2, 4-1st, 6, 7, and control samples) or V3-V4 (for OLP4-2nd, and 9-11 samples) hypervariable regions of the bacterial 16S rRNA gene were amplified by PCR, and then the PCR products were sequenced using a 454 GS FLX Titanium (Roche, Branford, CT, USA) or an Illumina MiSeq (Illumina, San Diego, CA, USA) sequencing system at ChunLab Inc. (Seoul, Korea), respectively. Processing and analysis

of sequences were performed using the CLcommunity™ software provided by ChunLab Inc.

4. Real-time PCR

Real-time PCR was performed in a 20 μl reaction mix containing 2 μl of bacterial genomic DNA, SYBR *Premix Ex Taq*, ROX Reference Dye II (Takara Bio, Otsu, Japan), and each primer. Universal (forward: 5'-AGTCACTGACGAGTTTGATCMTGGCTCAG-3' and reverse: 5'-CAGTGACTACWTTACCGCGGCTGCTGG-3') and *E. coli*-specific primers (forward: 5'- CCA TGC CGC GTG TAT GAA GA-3' and reverse 5'- AGA TGC AGT TCC CAG GTT GAG-3') targeting the bacterial 16S rRNA gene were used. *E. coli* K12-5.1 genomic DNA was used to generate standard curves.

5. Isolation of *Escherichia coli* from OLP biopsies

To isolate the clinical isolates of *E. coli* within the tissues from OLP lesions of two patients, bacteria on the surface of tissues were treated with lysozyme and antibiotics (gentamicin, streptomycin, and penicillin) as mentioned above. After washing, the tissues were homogenized and suspended in 5 ml of tryptic soy broth (TSB). Samples in TSB were incubated aerobically at 41.5℃ for 24 hours for pre-enrichment. Following incubation, the enrichments were streaked on eosin methylene blue (EMB) agar and incubated at 37℃ for overnight.

Colonies on agar plate were inoculated into TSB and cultured at 37℃ for overnight.

6. Confirmation of *E. coli* by PCR and sequencing of 16S rRNA gene

After growth of selected colonies in TSB, polymerase chain reaction (PCR) was performed using a 2 μl of bacterial cultures as a template with universal primers (27F: 5'-AGA GTT TGA TCM TGG CTC AG-3'; 1492R: 5'-TAC GGY TAC CTT GTT ACG ACT T-3') for whole 16S rRNA sequence of bacteria. After purification of PCR product using MEGAquick-spin™ Plus Total Fragment DNA Purification Kit (iNtRON, Gyeonggi-do, Korea), the purified PCR products were subjected to sequencing of the whole 16S rRNA gene with 27F primer. The obtained sequences were blasted with the GenBank database and confirmed as *E. coli*.

7. Whole genome sequencing

Four clinical isolates of *E. coli* were grown in Luria-Bertani (LB) medium at 37℃ for 16-18 hours under aerobic condition. Genomic DNA was extracted using the G-spin Genomic DNA extraction kit (iNtRON). The library was

prepared using the TruSeq DNA LT Sample Prep kit (Illumina), and was sequenced using an Illumina MiSeq instrument at ChunLab Inc. (Seoul, Korea). Data was analyzed using the Comparative Genomics (CG) pipeline of BIOiPLUG Apps (http://www.bioiplug.com/apps, ChunLab Inc.).

8. Animal experiment

The mouse experiments were performed and were approved by the Seoul National University Animal Care and Use committee (No. SNU-150901-1-6). Mice were maintained under specific pathogen-free conditions in the Laboratory Animal Facility at the School of Dentistry, Seoul National University.

Six-week-old female ICR mice were orally inoculated with 10^9 cells of viable *E. coli* K12-7.2 in 50 μl phosphate-buffered saline (PBS) containing 2% carboxymethylcellulose (CMC; Tokyo Chemical Industry, Tokyo, Japan) once a week (Figure 5), and sham control was inoculated with PBS in 2% CMC. Mice were euthanized at 16h after bacterial inoculation, and the tongue tissues were obtained to confirm the bacterial invasion and immune infiltration.

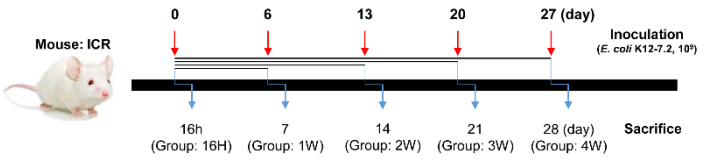


Figure 5. Experimental scheme for animal model.

9. *In situ* hybridization

A 244-bp DNA fragment chosen for the well conserved area located between V3 and V4 of *E. coli* 16S rRNA was amplified by PCR using the following primers: 5'- CCA TGC CGC GTG TAT GAA GA-3' and 5'- AGA TGC AGT TCC CAG GTT GAG-3'. Genomic DNA of *E. coli* SG13009 strain was used as a template. Amplification was performed under the following cycling conditions: 35 cycles at 94℃ for 30 sec, 60℃ for 30 sec, and 72℃ for 80 sec followed by a 5 min extension at 72℃. After precipitation of PCR products, the amplified products were labeled with digoxigenin (DIG)-dUTP by random priming using a commercial DIG DNA labeling and detection kit (Roche).

The specificity and sensitivity of DIG-labeled *E. coli*-specific probe was confirmed by dot blot analysis. To check the specificity of probe, 100 ng of denatured genomic DNA from various bacterial species [*P. gingivalis* (Pg), *T. denticola* (Td), *Fusobacterium nucleatum* Sub.*vincentii* (Fnv), *Capnocytophaga gingivalis* (Cg), *Veillonella dispar* (Vd), *Streptococcus gordonii* (Sg), *Acinetobacter johnsonii* (Aj), *Streptococcus salivarius* (Ss), *Escherichia coli* (Ec)] was loaded on the nylon membrane and dried at room temperature (RT). To immobilize the DNA on the membrane, the membrane was incubated at 80℃ for 1 hour and then blocked with 1X blocking solution of DIG DNA Labeling and Detection Kit (Roche). After blocking, the membrane was hybridized with 1 ng/μl of DIG-labeled probe and then washed with SSC buffer. To visualize hybridized DIG-labeled probe, anti-DIG-AP

antibody (1:2000) and NBT/BCIP substrate were applied to the membrane sequentially. To check the sensitivity of probe, serial dilutions of the DIG-labeled probe and a labeled control DNA provided in the kit (Roche) were loaded on the nylon membrane and dried at RT. The membrane was incubated at 80℃ for 1 hour and then blocked with 1% BSA in maleic acid buffer. The anti-DIG-AP antibody (1:5000) was applied to the membrane. After equilibration of membrane with detection buffer, NBT-BCIP substrate were applied to the membrane.

Serial 4-㎛ paraffin-embedded sections were subjected to de-paraffinization, re-hydration, and sequential treatment with 1 μg/ml of proteinase K, and 0.1 M triethanolamine-HCl. The DIG-labeled probe diluted in hybridization buffer [4X saline sodium citrate (SSC), 46% formamide, 0.1% sodium dodecyl sulfate (SDS), 1X Denhardt's solution, 0.4 mg/ml of salmon sperm DNA, 12% dextran sulfate] was heated at 95℃ for 10 min and chilled on ice. After applying probe on tissue sections, the slides were incubated at 90℃ for 10 min and hybridized at 60℃ for overnight in humidified chamber. As a negative control, hybridization was performed with the labeled probe with a 10-fold excess amount of unlabeled probe. After stringent washing with serial diluted SSC, the tissue sections were blocked and incubated with anti-DIG-alkaline phosphatase (AP) antibody. To block the endogenous AP activity, the tissues were treated with 1 mM of levamisole. The bacterial signals were visualized with nitro blue tetrazolium chloride (NBT)/5-bromo-4-chloro-3-

indolyl phosphate (BCIP), and then the tissues were counterstained with methyl green. Additional sections were stained with hematoxylin and eosin.

10. Statistical analysis

The Kruskal-Wallis test, followed by Mann-Whitney U test, was used to determine differences in relative abundance between healthy and patient with OLP . The Wilcoxon signed-rank test was used to compare various parameters between the mucosa and intratissue communities obtained from the same OLP patient. All statistical analyses were performed with SPSS Statistics 25 software (IBM). Significance was set at $p < 0.05$.

III. Results

1. Study population

For this analysis, the mucosal samples and biopsies of OLP lesions were obtained from 10 new patients (age 61.5 ± 2.6 years) diagnosed with OLP at SNUDH. In case of OLP4 patient, mucosal sample and biopsy were obtained twice because she had a relapse 2 years later after first visit. Among 10 patients, biopsy samples obtained from 2 patients (OLP5 and OLP7) were used to isolate the *E. coli*. Since only tissue samples were obtained from 4 patients (OLP5 – 8), the mucosal microbiota of 4 patients were not included in this analysis. For the control group, the samples of mucosal surface were obtained from 5 healthy subjects (age 42.0 ± 5.2 years, 4 female/1 male). In addition, the sequencing data of mucosal microbiota from 11 healthy subjects (age 52.5 ± 3.7 years, 6 female/ 5 male) and 13 OLP patients (age 56.8 ± 3.3 years, 6 female/ 7 male) in a previous study (28) were additionally used. Detailed clinical information of 10 patients is summarized in Table 1.

Table 1. Clinical information of enrolled OLP patients

No.	Age	Sex	REU[a] scoring		Sequencing		*E. coli* isolation
			REU	Score	Mucosa	Tissue	
OLP1	67	F	R5E3	9.5	G[b]	G	ND[d]
OLP2	68	F	R2E2	5	G	G	ND
OLP4-1st	55	F	R2E4	8	G	G	ND
OLP4-2nd	57	F	R5E3	9.5	M[c]	M	ND
OLP5	74	M	R2E1	3.5	ND	ND	OLPC3
OLP6	61	F	R7E5	14.5	ND	G	ND
OLP7	60	F	R1E3U1	7.5	ND	ND	OLPC73 OLPC76 OLPC84
OLP8	52	F	R2E5	9.5	ND	G	ND
OLP 9	76	M	R3E4U1	11	M	M	ND
OLP10	54	M	R6E6U2	19	M	M	ND
OLP11	59	M	R5E5	12.5	M	M	ND

[a]R: reticulation; E: erythema; U: ulceration, [b]G: Roche 454 GS FLX Titanium system, [c]M: Illumina MiSeq system, [d]ND: not done

2. Characteristics of bacterial communities in OLP

2.1. Dysbiosis of oral microbiota in the patients with OLP

To determine the characteristics of bacterial communities in OLP, the bacterial samples obtained from mucosal surface (OM, n = 20, age 58.3 ± 2.5 years, 10 female/10 male) and tissues (OT, n = 9, age 60.1 ± 2.6 years, 6 female/3 male) of lesion in patients with OLP were compared with mucosal microbiota of healthy control subjects (HM, n = 16, age 49.2 ± 3.1 years, 10 female/6 male)

by sequencing of 16S rRNA gene.

Total bacterial loads were quantified by real-time PCR of the 16S rRNA genes. The gene copy number of OM were significantly increased compared to HM and OT groups (Figure 6A).

From sequencing of 16S rRNA gene, a total of 760,115 filtered sequences with an average length of 444 bp were obtained, which presented ≥99.4% Good's coverage for each sample. The nonparametric richness estimated by Chao1 index (61) [314.9 (128.1) vs. 377.7 (228.0) vs. 115.3 (124.8)] showed a differences among groups, but not the species diversity for an operational taxonomic units (OTUs) definition determined by Simpson's (62) [0.04 (0.04) vs. 0.04 (0.04) vs. 0.07 (0.05)] and Shannon (63) [3.8 (0.76) vs. 4.0 (1.09) vs. 3.5 (0.84)] index (Figure 6B-D).

To compare the degree of phylogenetic distances among communities, UniFrac-based principal coordinates analysis (PCoA) was performed. In the PCoA plot, the tissue samples were partially separated from the mucosal surface samples of control and OLP and scattered in a bigger area than the other groups (Figure 7A), presenting a significantly increased UniFrac distance (0.056 ± 0.001 vs. 0.069 ± 0.001 vs. 0.095 ± 0.004), which is an index of intragroup inter-subject variability. In addition, the surface bacterial communities of OLP showed a higher UniFrac distance than that of control (Figure 7B).

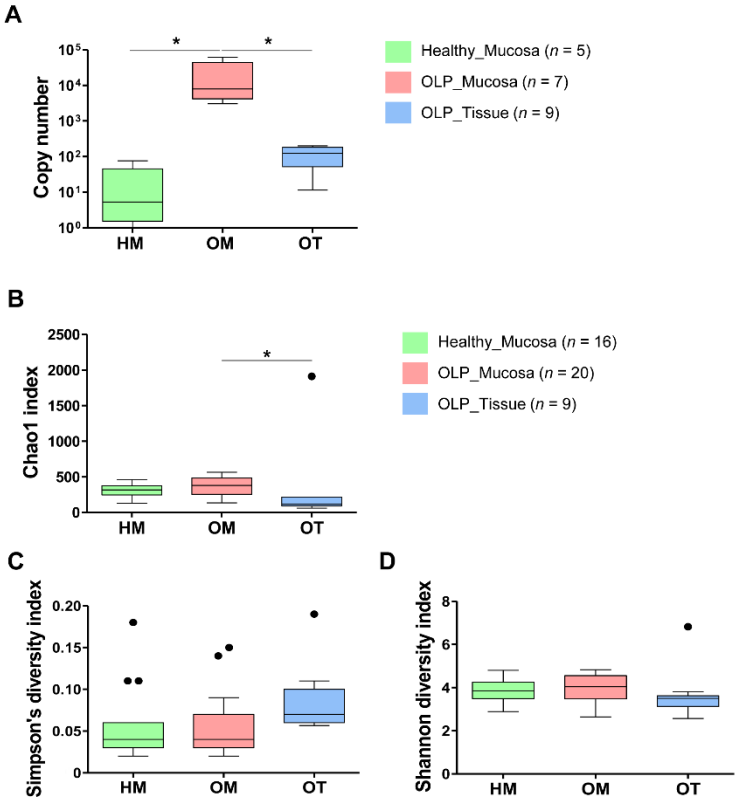


Figure 6. Alpha diversity analysis for healthy mucosa (HM, *n* = 16), mucosal surface (OM, *n* = 20), and intratissue (OT, *n* = 9) communities of OLP lesions. Bacterial communities of oral mucosa and tissue samples from healthy subject and OLP patients were analyzed by sequencing of 16S rRNA gene. (A) Bacterial loads were estimated by real-time PCR. (B) The species richness determined by Chao1 index was expressed as the Box-and-Whisker plots. (C, D) Simpson's and Shannon diversity index was expressed as the Box-and-Whisker plots. **p* < 0.05 by Kruskal-Wallis test, followed by Mann-Whitney U test.

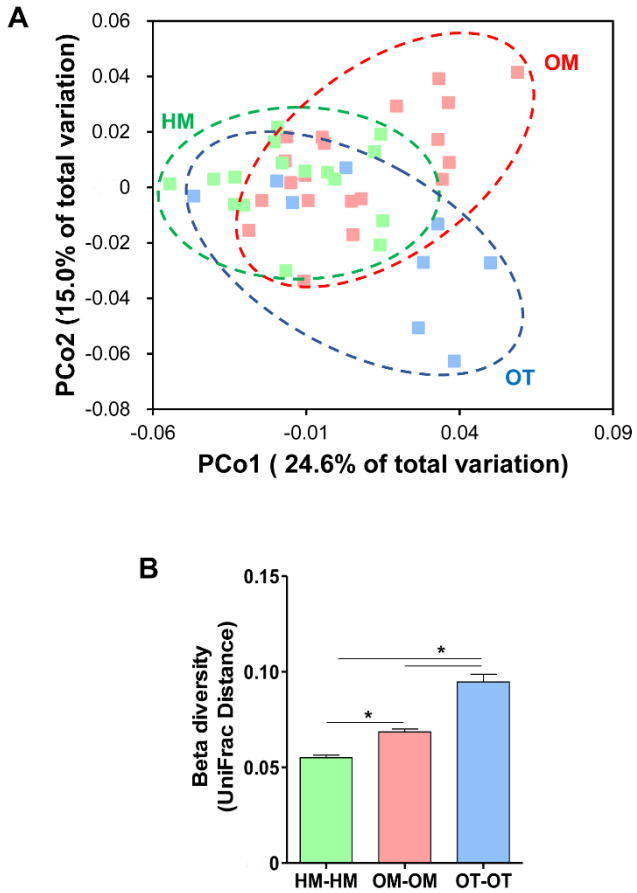


Figure 7. Beta diversity analysis for healthy mucosa (HM, *n* = 16), mucosal surface (OM, *n* = 20), and intratissue (OT, *n* = 9) communities of OLP lesions. (A) PCoA plot was generated using weighted UniFrac metric (green circle: mucosa samples of control, red circle: mucosa samples of OLP, blue circle: tissue samples of OLP). (B) The inter-subject UniFrac distances of each communities were obtained using a weighted metric and presented as the mean ± SEM. *$p < 0.05$ by one-way ANOVA with Tukey's post hoc.

2.2. Altered bacterial composition in OLP

To investigate the differences in the composition of bacterial communities, the relative abundance of each taxon among three groups was compared from phylum to species. From total samples, 51 phylum, 1076 genus, and 2423 species/phylotype were identified. Although a total 19, 20, and 43 phyla were identified from HM, OM, and OT group, respectively, only 5 phyla-Firmicutes, Proteobacteria, Bacteroidetes, Actinobacteria, and Fusobacteria-were detected in all samples. At the phylum level, the relative abundances of Spirochaetes and Synergistetes were clearly decreased in OT group compared to OM group. In addition, the relative abundance of TM7 was significantly decreased in OT group compared to HM and OM groups (Figure 8A, B). The relative abundance of Proteobacteria tended to increase in OT group without a statistical significance. At the genus (≥ 2.5% of relative abundance) level, the relative abundances of *Streptococcus*, *Fusobacterium*, *Leptotrichia*, *Lautropia*, *Actinomyces, Escherichia*, *Corynebacterium*, *Aggregatibacter, Treponema, Granulicatella, Campylobacter, Eikenella, and AM420062_g (Lachnospiraceae)* were significantly different among groups (Figure 9). At the species level (≥ 0.1% of relative abundance), the differences among 3 groups were observed for 90 species/phylotypes. Among the 90 species/phylotypes, differences in the relative abundances of 12 and 62 species/phylotypes were observed HM-OM and OM-OT, respectively. When compared with HM group, increase of *AY134896_s (Leptotrichia)*, *Parvimonas micra*, *Treponema_uc*,

Aggregatibacter segnis, *Eubacterium nodatum*, *Treponema denticola*, *Acinetobacter oryzae, AM420271_s (Lachnospiraceae), Leptotrichia trevisanii,* and *Senenomonas sputigena* and decrease of *E. coli* and *EF016847_s (Alicyclobacillaceae)* were observed in OM group. Interestingly, *E. coli* was highly enriched in OT community compare to OM, but statistical significance was not achieved. In addition, the relative abundances of gingivitis or periodontitis-associated bacteria were significantly decreased in OT compared with OM (Table 2).

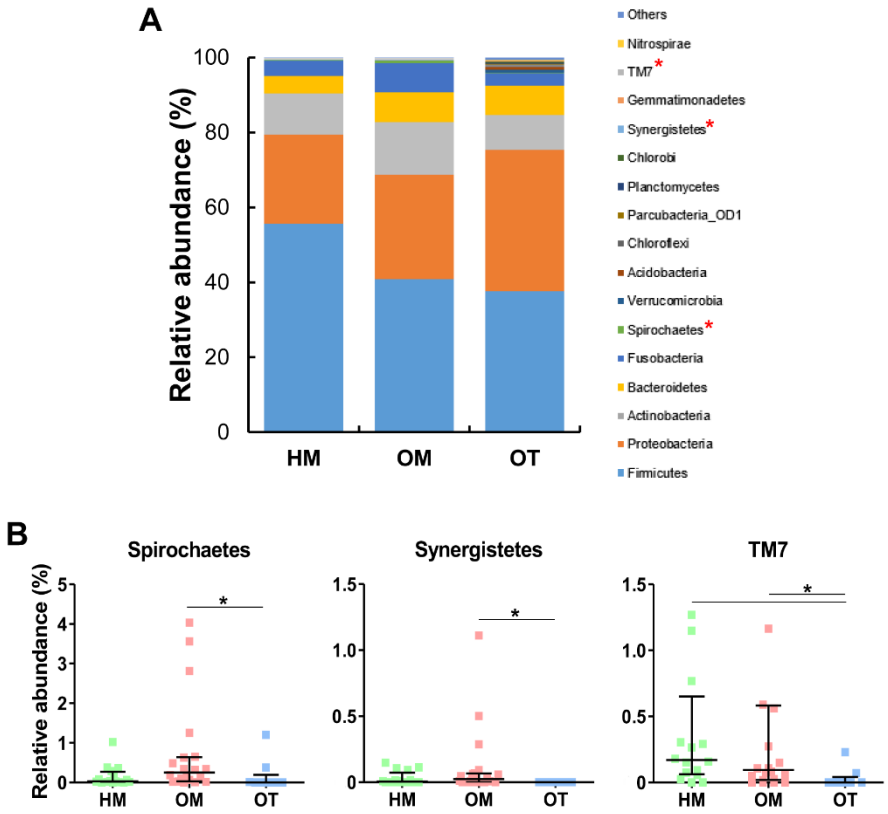


Figure 8. Differences in the relative abundance at phylum level. The relative abundance at the phylum level was compared among 3 communities. (A) The members of top 16 phyla are shown (> 1% of relative abundance) and presented as the mean. (B) Three phyla were differently distributed among the 3 communities and present the median with interquartile range (IQR). *$p < 0.05$ by Kruskal-Wallis test, followed by Mann-Whitney U test.

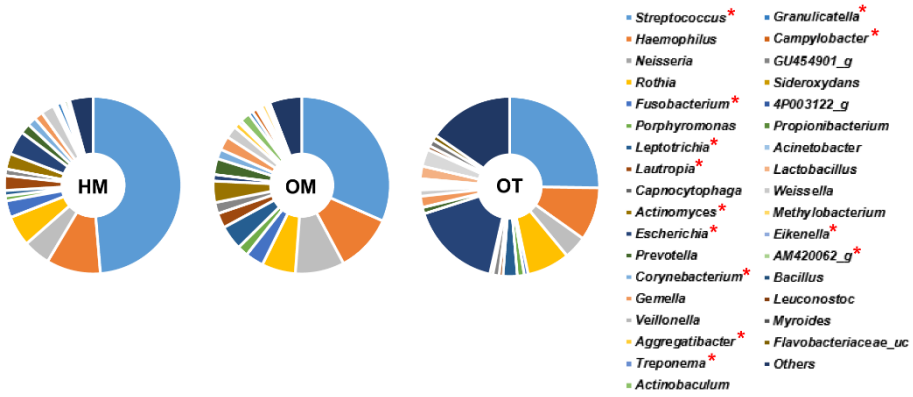


Figure 9. Differences in the relative abundance at genus level. The relative abundance at the genus level was compared among 3 communities. Pie charts present the mean relative abundance of dominant genera (≥ 2.5% of relative abundance). *$p < 0.05$ by Kruskal-Wallis test, followed by Mann-Whitney U test.

Table 2. Relative abundance of species/phylotypes differently distributed in HM, OM, and OT group

Species/phylotype	Relative abundance (%)[a]			*p* value[b]		
	HM (*n* = 16)	OM (*n* = 20)	OT (*n* = 9)	HM-OM	HM-OT	OM-OT
Streptococcus pseudopneumoniae	7.30 (53.28)	5.46 (25.31)	0.17 (44.40)	0.217	**0.005**[c]	0.232
Lautropia mirabilis	1.35 (13.97)	1.53 (8.96)	0.15 (3.34)	1.000	0.207	**0.027**
Fusobacterium nucleatum[d]	0.72 (6.27)	1.24 (9.39)	0.11 (1.39)	0.261	0.242	**0.004**
Streptococcus mitis	3.04 (19.87)	2.75 (19.29)	0.06 (12.60)	0.438	**0.014**	0.250
Escherichia coli	1.07 (13.06)	0.05 (18.24)	2.62 (64.19)	**0.003**	1.000	0.169
Rothia aeria	0.94 (17.99)	1.58 (10.99)	0 (0.29)	1.000	**0.001**	**0.000**
Streptococcus sanguinis	2.07 (16.90)	1.11 (8.73)	0.45 (0.91)	1.000	**0.002**	**0.009**
HQ767899_s (Streptococcus)	0.61 (42.91)	0 (6.52)	0 (0)	0.316	**0.005**	0.170
Streptococcus cristatus	0.96 (6.27)	0.98 (17.32)	0 (0.16)	0.802	**0.001**	**0.018**
Neisseria oralis	0 (7.16)	0.14 (33.77)	0 (0)	0.317	0.145	**0.002**
Porphyromonas endodontalis	0.16 (0.50)	0.11 (2.66)	0 (0.54)	0.568	0.541	**0.039**
Corynebacterium matruchotii	0.20 (6.84)	0.18 (16.06)	0 (0.05)	1.000	**0.003**	**0.001**
Fusobacterium periodonticum	0.82 (8.24)	0.58 (2.81)	0 (0.85)	0.713	**0.010**	0.117
AY134896_s (Leptotrichia)	0 (0)	0 (4.94)	0 (0.10)	**0.002**	1.000	0.092
Streptococcus oralis	0.41 (1.89)	0.67 (4.42)	0 (0.45)	1.000	0.058	**0.037**
Parvimonas micra	0.02 (0.35)	0.12 (1.22)	0 (0.50)	**0.014**	0.973	**0.002**
Granulicatella adiacens	0.71 (2.67)	0.53 (3.63)	0.05 (1.45)	0.651	**0.006**	0.086
Treponema_uc	0 (0.14)	0 (1.29)	0 (0.08)	**0.038**	1.000	0.216
DQ016843_s (Streptococcus)	0.33 (3.07)	0.11 (3.94)	0 (0)	0.975	**0.007**	0.058
Actinomyces viscosus	0.18 (2.11)	0.26 (4.56)	0 (0.22)	1.000	**0.015**	**0.006**
Corynebacterium durum	0.22 (2.52)	0.14 (1.38)	0 (0.12)	1.000	**0.015**	**0.018**
Actinomyces oris	0.35 (3.16)	0.16 (3.67)	0 (0.22)	1.000	**0.007**	**0.015**
Streptococcus australis	0.53 (1.66)	0.17 (1.65)	0 (0.16)	0.362	**0.001**	0.056
Streptococcus_uc	0.26 (12.21)	0.13 (0.66)	0.03 (0.27)	0.098	**0.001**	0.196
Aggregatibacter segnis	0 (0.23)	0.09 (2.57)	0 (0.06)	**0.010**	0.819	**0.001**
Actinomyces_uc	0.03 (0.13)	0.09 (1.09)	0 (0.33)	0.520	0.333	**0.016**
Actinomyces naeslundii	0.02 (0.67)	0.06 (2.87)	0 (0.32)	0.324	0.258	**0.005**
Prevotella nigrescens	0.02 (0.68)	0.07 (4.66)	0 (0.08)	0.316	0.113	**0.001**
Capnocytophaga gingivalis	0.13 (0.61)	0.13 (2.51)	0 (0.07)	1.000	**0.010**	**0.001**
Eubacterium brachy	0.01 (0.26)	0.06 (0.47)	0 (0.04)	0.116	0.240	**0.001**
Neisseria elongata	0.29 (0.82)	0.17 (0.73)	0 (1.00)	1.000	**0.005**	**0.013**
Capnocytophaga sputigena	0.03 (2.20)	0.15 (0.87)	0 (0.49)	0.263	0.849	**0.033**
Eubacterium nodatum	0 (0.26)	0.02 (0.65)	0 (0.01)	**0.004**	1.000	**0.002**
Campylobacter gracilis	0.04 (0.62)	0.04 (0.48)	0 (0.04)	1.000	**0.017**	**0.013**
Prevotella oris	0.05 (1.19)	0.08 (1.00)	0 (0)	1.000	**0.002**	**0.000**
Campylobacter concisus	0.11 (0.52)	0.06 (8.45)	0 (0.15)	1.000	0.037	0.184
AM420165_s (Aggregatibacter)	0.19 (0.83)	0.07 (1.74)	0 (0.11)	1.000	**0.011**	0.071
Kingella oralis	0.02 (0.58)	0.05 (1.13)	0 (0)	1.000	**0.028**	**0.007**
Streptococcus tigurinus	0.20 (1.55)	0.10 (0.81)	0 (0.47)	0.234	**0.001**	0.076
Campylobacter showae	0.02 (0.25)	0.08 (0.38)	0 (0.02)	0.171	0.165	**0.001**
ADCM_s (Paludibacter)	0.01 (0.63)	0.04 (0.65)	0 (0.15)	0.589	0.457	**0.031**
Cardiobacterium hominis	0.01 (0.70)	0.02 (0.40)	0 (0)	0.876	0.099	**0.006**
Fusobacterium canifelinum	0.10 (1.31)	0.09 (1.33)	0 (0.03)	1.000	**0.012**	**0.019**
Treponema denticola	0 (0.13)	0.02 (0.37)	0 (0)	**0.031**	0.435	**0.001**
EF016847_s (Alicyclobacillaceae)	0.10 (1.20)	0 (1.16)	0 (2.38)	**0.000**	0.067	1.000
Acinetobacter oryzae	0 (0)	0 (3.42)	0 (0)	**0.029**	1.000	0.091

Species/phylotype	Relative abundance (%)[a]			*p* value[b]		
	HM (*n* = 16)	OM (*n* = 20)	OT (*n* = 9)	HM-OM	HM-OT	OM-OT
Stomatobaculum longum	0.02 (0.45)	0.07 (1.26)	0 (0.07)	0.363	0.743	**0.038**
Eikenella corrodens	0.02 (0.37)	0.05 (3.70)	0 (0)	0.994	**0.013**	**0.000**
DQ241813_s (Bergeyella)	0.07 (0.98)	0.03 (0.13)	0 (0.34)	0.513	**0.039**	0.457
4P003193_s (Prevotella)	0.03 (2.07)	0.01 (1.27)	0 (0)	1.000	**0.010**	**0.028**
Prevotella intermedia	0.04 (1.11)	0.13 (0.59)	0 (0)	0.761	**0.030**	**0.001**
AM420132_s (Saccharimonas)	0.02 (2.09)	0.03 (0.63)	0 (0.02)	1.000	0.061	**0.022**
Lachnoanaerobaculum orale	0.07 (0.53)	0.04 (1.18)	0 (0.21)	1.000	**0.021**	0.101
DQ012324_s (Capnocytophaga)	0.01 (0.70)	0 (2.71)	0 (0)	1.000	**0.024**	0.150
Actinomyces meyeri	0.01 (0.11)	0.09 (2.86)	0 (0.03)	0.095	0.451	**0.003**
Actinomyces massiliensis	0 (0.67)	0.03 (0.61)	0 (0.03)	1.000	0.092	**0.010**
Oribacterium sinus	0.08 (0.38)	0.11 (0.54)	0 (0.13)	1.000	**0.010**	**0.007**
Treponema socranskii	0 (0.07)	0.03 (0.17)	0 (0.01)	0.059	1.000	**0.018**
AM420062_s (Lachnospiraceae)	0 (0.52)	0 (2.54)	0 (0)	0.587	0.430	**0.028**
Lachnoanaerobaculum saburreum	0 (0.86)	0 (0.93)	0 (0)	1.000	**0.048**	0.059
Prevotella oulorum	0 (0.83)	0.03 (0.78)	0 (0)	0.261	0.078	**0.001**
4P004975_s (Actinomyces)	0.03 (1.18)	0 (1.12)	0 (0)	0.160	**0.011**	0.483
Gemella morbillorum	0.01 (0.15)	0.02 (0.67)	0 (0.03)	1.000	0.124	**0.033**
4P003196_s (Actinomyces)	0.05 (0.56)	0.07 (0.64)	0 (0.17)	1.000	0.052	**0.017**
AM420271_s (Lachnospiraceae)	0 (0.10)	0.01 (0.18)	0 (0.04)	**0.036**	1.000	0.080
4P004512_s (Corynebacterium)	0.01 (1.90)	0 (0.14)	0 (0)	0.077	**0.017**	0.953
AM420230_s (Tannerella)	0 (1.01)	0.01 (0.69)	0 (0)	1.000	0.062	**0.029**
Leptotrichia trevisanii	0 (0.02)	0.02 (0.33)	0 (0.07)	**0.023**	1.000	0.219
Prevotella pallens	0.03 (0.37)	0.04 (0.61)	0 (0.11)	0.995	0.167	**0.015**
Leptotrichia buccalis	0 (0.03)	0.04 (3.62)	0 (0.08)	0.069	1.000	**0.023**
HQ757980_s (Streptococcus)	0.04 (1.15)	0 (0.11)	0 (0.03)	0.062	**0.011**	0.832
AF287757_s (Propionibacterium)	0 (0.45)	0 (0.65)	0 (0)	0.653	0.408	**0.030**
Prevotella nanceiensis	0.06 (0.22)	0.05 (0.28)	0 (0.10)	1.000	**0.044**	0.129
Alloprevotella rava	0 (0.10)	0.01 (0.45)	0 (0)	0.220	0.319	**0.005**
Neisseria bacilliformis	0 (0.50)	0.01 (0.18)	0 (0)	0.573	0.202	**0.008**
AF385572_s (Leptotrichia)	0 (0.14)	0.02 (0.74)	0 (0)	0.051	0.447	**0.001**
Mogibacterium timidum	0 (0.29)	0.03 (0.18)	0 (0)	0.200	0.332	**0.004**
Abiotrophia defective	0 (0.12)	0 (1.12)	0 (0)	0.510	0.643	**0.045**
Veillonella rodentium	0.02 (0.82)	0.01 (0.12)	0 (0.03)	0.845	**0.026**	0.207
Catonella morbi	0 (0.25)	0.03 (0.26)	0 (0)	0.880	**0.046**	**0.002**
Cutibacterium acnes	0 (0)	0 (0)	0 (0.53)	1.000	**0.014**	**0.033**
Atopobium parvulum	0.01 (0.36)	0.01 (0.66)	0 (0)	1.000	**0.031**	**0.006**
ACIQ_s (Oribacterium)	0 (0.19)	0.02 (0.46)	0 (0)	0.859	0.158	**0.011**
Selenomonas sputigena	0 (0.04)	0.02 (0.20)	0 (0)	**0.004**	1.000	**0.001**
4P003133_s (Streptococcus)	0.01 (0.10)	0.01 (1.21)	0 (0)	1.000	0.082	**0.022**
Selenomonas infelix	0 (0.05)	0 (0.41)	0 (0)	0.070	1.000	**0.030**
Fretibacterium fastidiosum	0 (0.06)	0 (0.10)	0 (0)	0.091	1.000	**0.034**
Prevotella baroniae	0 (0.16)	0 (0.22)	0 (0)	0.724	0.387	**0.032**
GU561335_s (Capnocytophaga)	0.01 (0.19)	0 (0.75)	0 (0)	0.222	**0.024**	0.621
Prevotella tannerae	0 (0.27)	0 (0.21)	0 (0)	1.000	0.228	**0.036**

[a]Expressed as median and range, [b]by Kruskal-Wallis test followed by Mann-Whitney U test with Bonferroni adjustment, [c]$p < 0.05$ by Bonferroni adjustment, [d]Species associated with gingivitis or periodontitis are bolded.

3. Features of intratissue bacterial communities in OLP

3.1. Alpha and beta diversities of mucosal surface and intratissue communities within OLP lesions

To compare the oral microbiota between the mucosal surface and the tissue of OLP lesions within the same patient, both samples from 7 patients with OLP were re-analyzed.

From sequencing of 16S rRNA gene, a total of 473,926 filtered sequences with an average length of 406 bp were obtained, which presented ≥99% Good's coverage for each sample. The bacterial loads were also quantified by real-time PCR. The copy number of tissue samples was significantly decreased compared with mucosa samples (Figure 10A). Not only Chao1 index [331.6 (256.5) vs. 107.0 (133.2)] but Shannon [3.6 (1.43) vs. 3.4 (1.19)] and Simpson's [0.05 (0.07) vs. 0.08 (0.05)] index of tissue communities were comparable to those of mucosa (Figure 10B). In Figure 6, Chao1 index between OM and OT groups showed a significant differences. However, when compared paired samples, they did not show differences. In the PCoA, the tissue samples were not completely separated from the mucosal samples and scattered in a big area (Figure 10C, left), presenting a significantly increased UniFrac distance (0.065

± 0.003 vs. 0.095 ± 0.006) (Figure 10C, right).

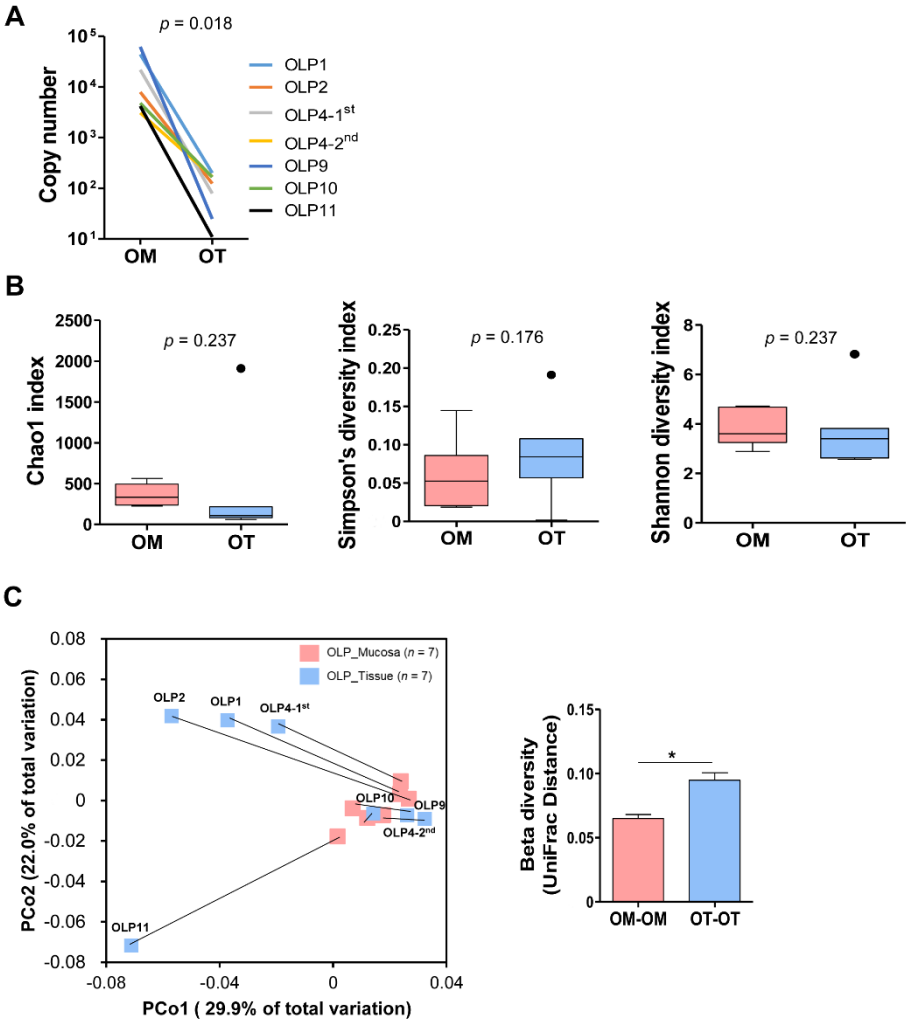


Figure 10. Alpha and beta diversities of oral mucosa and intratissue communities within OLP lesions. Oral mucosa and tissue samples were obtained from seven OLP patients. Bacterial genomic DNA was prepared and subjected to sequencing of the 16S rRNA gene. (A) Total bacterial load in each

sample was estimated by real-time PCR using universal primers targeting the bacterial 16S rRNA gene. The total bacterial load was expressed as the 16S rRNA gene copy number in the total DNA obtained from each sample. (B) The species richness and diversity of two communities were estimated by Chao1, Simpson's and Shannon index. (C) PCoA plot was generated using weighted UniFrac metric. Samples from the same subject are connected with a solid line (left). The inter-subject UniFrac distances of the oral mucosa and intratissue communities were obtained using a weighted metric (right). p value and $*p < 0.05$ by two-tailed Wilcoxon signed-rank test.

3.2. Differences in oral microbiota composition between mucosa and tissue in OLP

From total samples, 46 phylum, 929 genus, 1981 species/phylotypes were identified. Comparison of the relative abundance of each taxon showed many differences in the bacterial composition between the mucosa and tissue samples. At the phylum level, the changes in the relative abundance between mucosa and tissue within each patient were clearly observed (Figure 11A, upper). When compared with OM group, the abundances of Spirochaetes and Synergistetes were significantly decreased in the tissue, while the abundance of Proteobacteria tends to increase in the tissue without statistical significance (Figure 11A, bottom, B). At the genus level, the abundance of *Escherichia*

among 8 genus differently distributed between mucosa and tissue was only clearly increased in the tissue, whereas the abundances of 7 genera were obviously decreased in the tissue (Figure 12). At the species level (≥ 0.1% of relative abundance), 52 species of total bacteria identified in the samples showed a significant differences between 2 groups. Most species identified in samples were significantly decreased in the tissues, while *E. coli* was highly enriched in the tissue community, composing 0.02% to 64.19% of the total bacteria (Figure 13). Interestingly, *E. coli* was not detected in OLP 4-2nd tissue, although it took 21.1% in the OLP4-1st tissue. In addition, the relative abundances of *Streptococcus pneumonia* and *Weissella kandleri* tended to increase in tissue community of OLP 4-2nd, OLP9 and OLP10, respectively. To confirm the sequencing data, the relative ratio of *E. coli* was estimated by real-time PCR. The quantification result by PCR was not perfectly corresponded with the sequencing data. However, a high amount of *E. coli* in OLP 1, 2, and 4-1st samples than that of the others was consistent with the sequencing data (Figure 14).

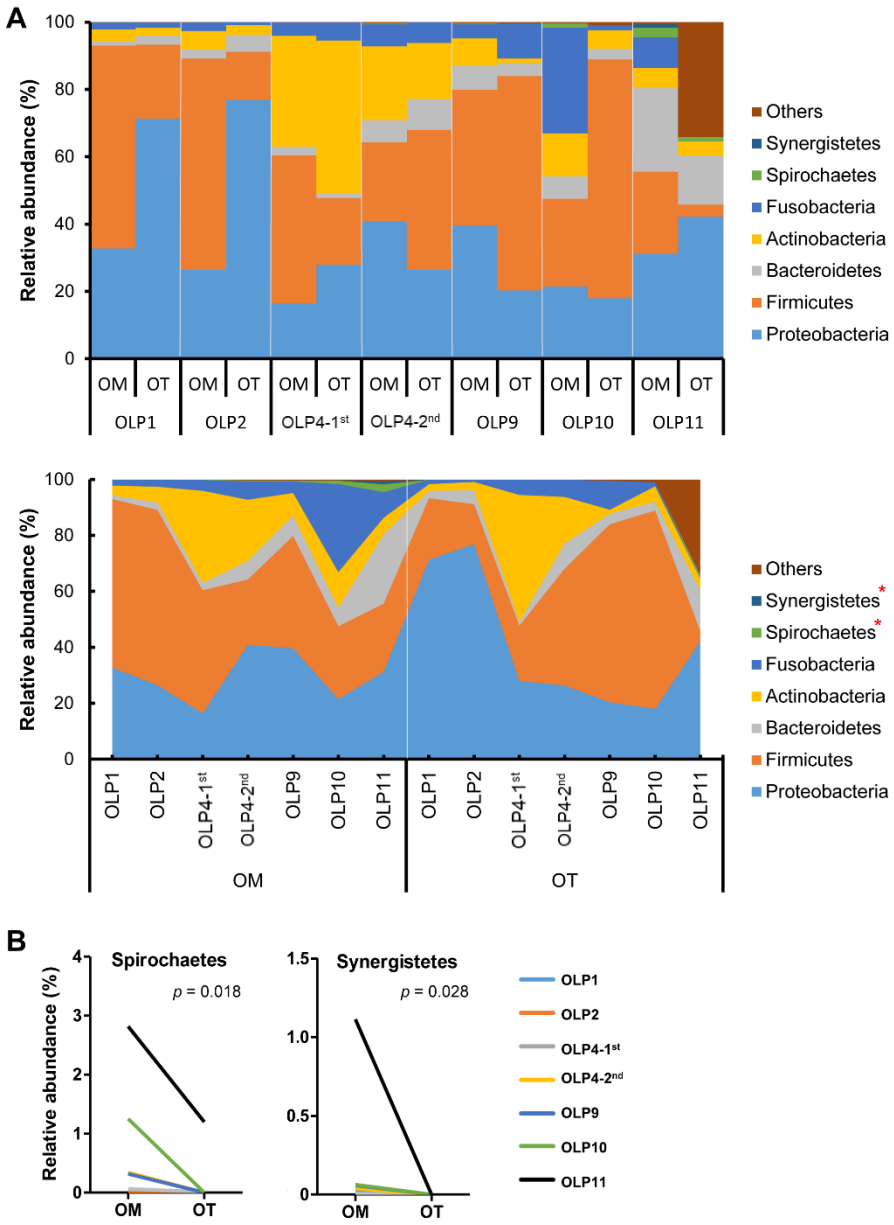


Figure 11. Differences in bacterial composition at the phylum level between the oral mucosa and intratissue communities. The relative abundances of bacterial phyla were compared. (A) The members of top 7 phyla are shown. (B)

Two phyla were differently distributed between the 2 communities. *p* value and *$p < 0.05$ by two-tailed Wilcoxon signed-rank test.

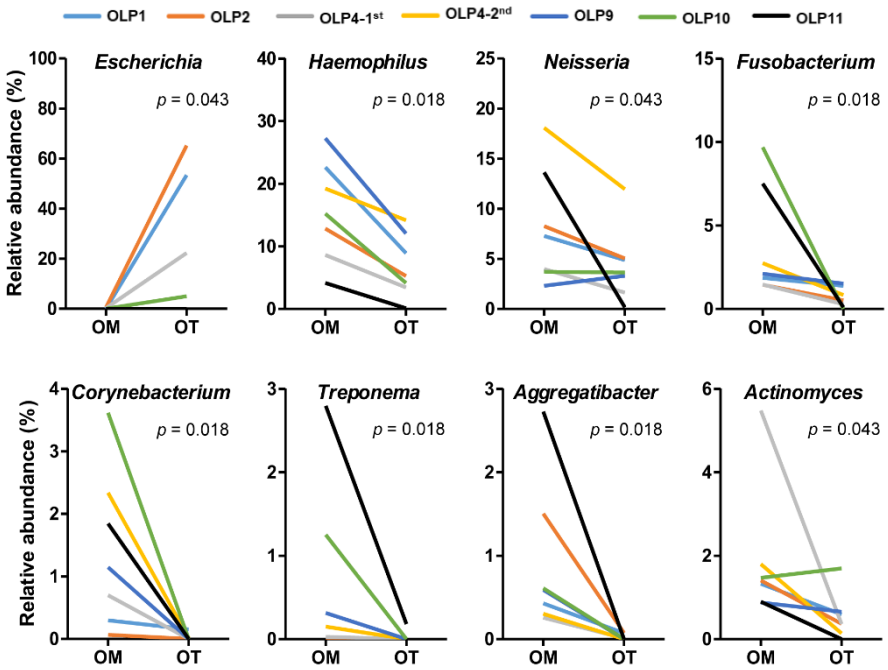
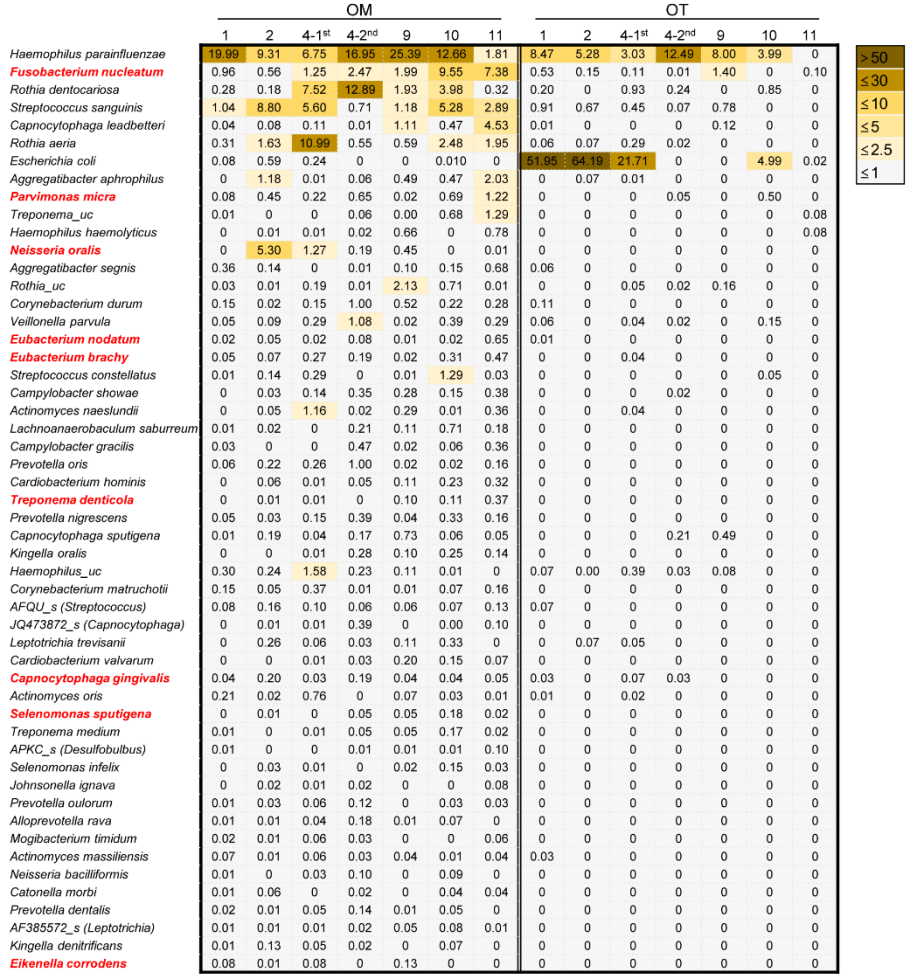


Figure 12. Differences in bacterial composition the genus level between the oral mucosa and intratissue communities. The relative abundances of bacterial genera were compared. The members of top 8 genera are shown. *p* value and *$p < 0.05$ by two-tailed Wilcoxon signed-rank test.



	OM							OT						
	1	2	4-1st	4-2nd	9	10	11	1	2	4-1st	4-2nd	9	10	11
Haemophilus parainfluenzae	19.99	9.31	6.75	16.95	25.39	12.66	1.81	8.47	5.28	3.03	12.49	8.00	3.99	0
Fusobacterium nucleatum	0.96	0.56	1.25	2.47	1.99	9.55	7.38	0.53	0.15	0.11	0.01	1.40	0	0.10
Rothia dentocariosa	0.28	0.18	7.52	12.89	1.93	3.98	0.32	0.20	0	0.93	0.24	0	0.85	0
Streptococcus sanguinis	1.04	8.80	5.60	0.71	1.18	5.28	2.89	0.91	0.67	0.45	0.07	0.78	0	0
Capnocytophaga leadbetteri	0.04	0.08	0.11	0.01	1.11	0.47	4.53	0.01	0	0	0	0.12	0	0
Rothia aeria	0.31	1.63	10.99	0.55	0.59	2.48	1.95	0.06	0.07	0.29	0.02	0	0	0
Escherichia coli	0.08	0.59	0.24	0	0	0.010	0	51.95	64.19	21.71	0	0	4.99	0.02
Aggregatibacter aphrophilus	0	1.18	0.01	0.06	0.49	0.47	2.03	0	0.07	0.01	0	0	0	0
Parvimonas micra	0.08	0.45	0.22	0.65	0.02	0.69	1.22	0	0	0	0.05	0	0.50	0
Treponema_uc	0.01	0	0	0.06	0.00	0.68	1.29	0	0	0	0	0	0	0.08
Haemophilus haemolyticus	0	0.01	0.01	0.02	0.66	0	0.78	0	0	0	0	0	0	0.08
Neisseria oralis	0	5.30	1.27	0.19	0.45	0	0.01	0	0	0	0	0	0	0
Aggregatibacter segnis	0.36	0.14	0	0.01	0.10	0.15	0.68	0.06	0	0	0	0	0	0
Rothia_uc	0.03	0.01	0.19	0.01	2.13	0.71	0.01	0	0	0.05	0.02	0.16	0	0
Corynebacterium durum	0.15	0.02	0.15	1.00	0.52	0.22	0.28	0.11	0	0	0	0	0	0
Veillonella parvula	0.05	0.09	0.29	1.08	0.02	0.39	0.29	0.06	0	0.04	0.02	0	0.15	0
Eubacterium nodatum	0.02	0.05	0.02	0.08	0.01	0.02	0.65	0.01	0	0	0	0	0	0
Eubacterium brachy	0.05	0.07	0.27	0.19	0.02	0.31	0.47	0	0	0.04	0	0	0	0
Streptococcus constellatus	0.01	0.14	0.29	0	0.01	1.29	0.03	0	0	0	0	0	0.05	0
Campylobacter showae	0	0.03	0.14	0.35	0.28	0.15	0.38	0	0	0	0.02	0	0	0
Actinomyces naeslundii	0	0.05	1.16	0.02	0.29	0.01	0.36	0	0	0.04	0	0	0	0
Lachnoanaerobaculum saburreum	0.01	0.02	0	0.21	0.11	0.71	0.18	0	0	0	0	0	0	0
Campylobacter gracilis	0.03	0	0	0.47	0.02	0.06	0.36	0	0	0	0	0	0	0
Prevotella oris	0.06	0.22	0.26	1.00	0.02	0.02	0.16	0	0	0	0	0	0	0
Cardiobacterium hominis	0	0.06	0.01	0.05	0.11	0.23	0.32	0	0	0	0	0	0	0
Treponema denticola	0	0.01	0.01	0	0.10	0.11	0.37	0	0	0	0	0	0	0
Prevotella nigrescens	0.05	0.03	0.15	0.39	0.04	0.33	0.16	0	0	0	0	0	0	0
Capnocytophaga sputigena	0.01	0.19	0.04	0.17	0.73	0.06	0.05	0	0	0	0.21	0.49	0	0
Kingella oralis	0	0	0.01	0.28	0.10	0.25	0.14	0	0	0	0	0	0	0
Haemophilus_uc	0.30	0.24	1.58	0.23	0.11	0.01	0	0.07	0.00	0.39	0.03	0.08	0	0
Corynebacterium matruchotii	0.15	0.05	0.37	0.01	0.01	0.07	0.16	0	0	0	0	0	0	0
AFQU_s (Streptococcus)	0.08	0.16	0.10	0.06	0.06	0.07	0.13	0.07	0	0	0	0	0	0
JQ473872_s (Capnocytophaga)	0	0.01	0.01	0.39	0	0.00	0.10	0	0	0	0	0	0	0
Leptotrichia trevisanii	0	0.26	0.06	0.03	0.11	0.33	0	0	0.07	0.05	0	0	0	0
Cardiobacterium valvarum	0	0	0.01	0.03	0.20	0.15	0.07	0	0	0	0	0	0	0
Capnocytophaga gingivalis	0.04	0.20	0.03	0.19	0.04	0.04	0.05	0.03	0	0.07	0.03	0	0	0
Actinomyces oris	0.21	0.02	0.76	0	0.07	0.03	0.01	0.01	0	0.02	0	0	0	0
Selenomonas sputigena	0	0.01	0	0.05	0.05	0.18	0.02	0	0	0	0	0	0	0
Treponema medium	0.01	0	0.01	0.05	0.05	0.17	0.02	0	0	0	0	0	0	0
APKC_s (Desulfobulbus)	0.01	0	0	0.01	0.01	0.01	0.10	0	0	0	0	0	0	0
Selenomonas infelix	0	0.03	0.01	0	0.02	0.15	0.03	0	0	0	0	0	0	0
Johnsonella ignava	0	0.02	0.01	0.02	0	0	0.08	0	0	0	0	0	0	0
Prevotella oulorum	0.01	0.03	0.06	0.12	0	0.03	0.03	0	0	0	0	0	0	0
Alloprevotella rava	0.01	0.01	0.04	0.18	0.01	0.07	0	0	0	0	0	0	0	0
Mogibacterium timidum	0.02	0.01	0.06	0.03	0	0	0.06	0	0	0	0	0	0	0
Actinomyces massiliensis	0.07	0.01	0.06	0.03	0.04	0.01	0.04	0.03	0	0	0	0	0	0
Neisseria bacilliformis	0.01	0	0.03	0.10	0	0.09	0	0	0	0	0	0	0	0
Catonella morbi	0.01	0.06	0	0.02	0	0.04	0.04	0	0	0	0	0	0	0
Prevotella dentalis	0.02	0.01	0.05	0.14	0.01	0.05	0	0	0	0	0	0	0	0
AF385572_s (Leptotrichia)	0.01	0.01	0.01	0.02	0.05	0.08	0.01	0	0	0	0	0	0	0
Kingella denitrificans	0.01	0.13	0.05	0.02	0	0.07	0	0	0	0	0	0	0	0
Eikenella corrodens	0.08	0.01	0.08	0	0.13	0	0	0	0	0	0	0	0	0

Figure 13. Differences in bacterial composition at the species level between the oral mucosa and intratissue communities. The relative abundances of bacterial species were compared (≥ 0.1% of relative abundance). The relative abundances of 52 species/phylotypes differently distributed ($p < 0.05$ by two-tailed Wilcoxon signed-rank test) between 2 groups are shown. Species associated with gingivitis or periodontitis are red-colored.

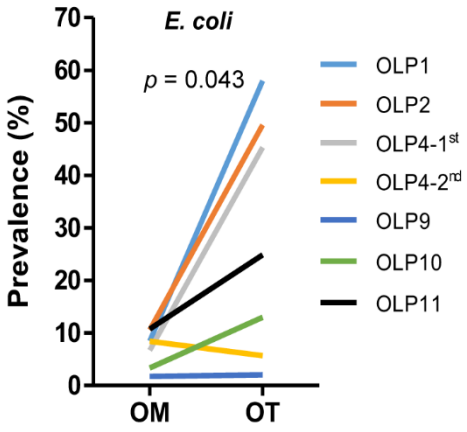


Figure 14. The prevalence of *E. coli* estimated by real-time PCR. To quantify the relative ratio of *E. coli* in total bacteria, real-time PCR was performed using universal and *E. coli*-specific primers. The gene copy number calculated using standard curves, and the prevalence was estimated. *p* value by two-tailed Wilcoxon signed-rank test.

4. Isolation and confirmation of *E. coli* within the tissues from OLP lesions

In the previous study, *P. gingivalis* established as a periodontal pathogen was highly enriched in the tissue of patient with periodontitis. A highly increased relative abundance of *E. coli* in the tissue (0.02% to 64.2% of total community) suggest a possibility that *E. coli* may have a role in the pathogenesis of OLP. It is known that *E. coli* shows a diversity in their genotype and phenotype within same species, with varying serotype and pathotype depending on the strain (64). To investigate the role of *E.coli* in the pathogenesis of OLP, *E. coli* was isolated from the tissue of OLP lesion. From the tissues of 2 patients, four clinical isolates of *E.coli* (OLPC3, OLPC73, OLPC76, and OLPC84) were obtained and confirmed by sequencing of 16S rRNA gene. Three strains of clinical isolates (OLPC73, 76, and 84) were isolated from OLP7 patient and the other (OLPC3) was isolated from OLP5 patient (Table 1).

5. Genomic characterization of *E. coli* isolates by whole genome sequencing

5.1. General genomic features of *E. coli* isolates

To analyze the genomic features of clinical isolates, whole genomes of four clinical isolates were sequenced. Four clinical isolates have an average 4.6-Mb genome with 50.7% G+C content. Plasmids were not found in all isolates. The general genomic features of four strains are described in Table 3.

Table 3. General genomic features of four *E.coli* isolates

Strain	OLPC3	OLPC73	OLPC76	OLPC84
Genome size (bp)	4,688,958	4,687,682	4,685,982	4,685,872
No. of plasmids	0	0	0	0
No. of CDSs[a]	4,361	4,364	4,360	4,367
No. of rRNAs	5	6	7	5
No. of tRNAs	66	65	64	63
No. of contigs	126	131	134	123
N_{50} (bp)	105,630	105,630	95,009	94,980
GC ratio (%)	50.7	50.7	50.7	50.7

[a]CDSs: protein-coding regions

5.2. Comparison of gene contents among *E. coli* isolates

To compare the genome contents among 4 *E. coli* isolates, pan-genome analysis was performed. The core genome contained 4,243 genes from total genes. Each of *E. coli* strains OLPC3, OLPC73, and OLPC84 had two unique genes, respectively. OLPC3 strain had a putative prophage Qin-packaging protein NU1 like protein and a hypothetical protein. OLPC73 strain had a DNA-directed RNA polymerase and a hypothetical protein, whereas OLPC84 had two hypothetical proteins (Figure 15A). To characterize the distribution of functional genes in *E. coli,* the classification of ortholog clusters was analyzed using the clusters of orthologous groups (COGs) database. With the exception of uncharacterized category, the most abundant COG categories were [G] carbohydrate transport and metabolism, [E] amino acid transport and metabolism, followed by [K] transcription. Whereas COG categories [V] defense mechanism, [N] cell motility, and [Q] secondary metabolites biosynthesis, transport and catabolism were lower in 4 genomes (Figure 15B).

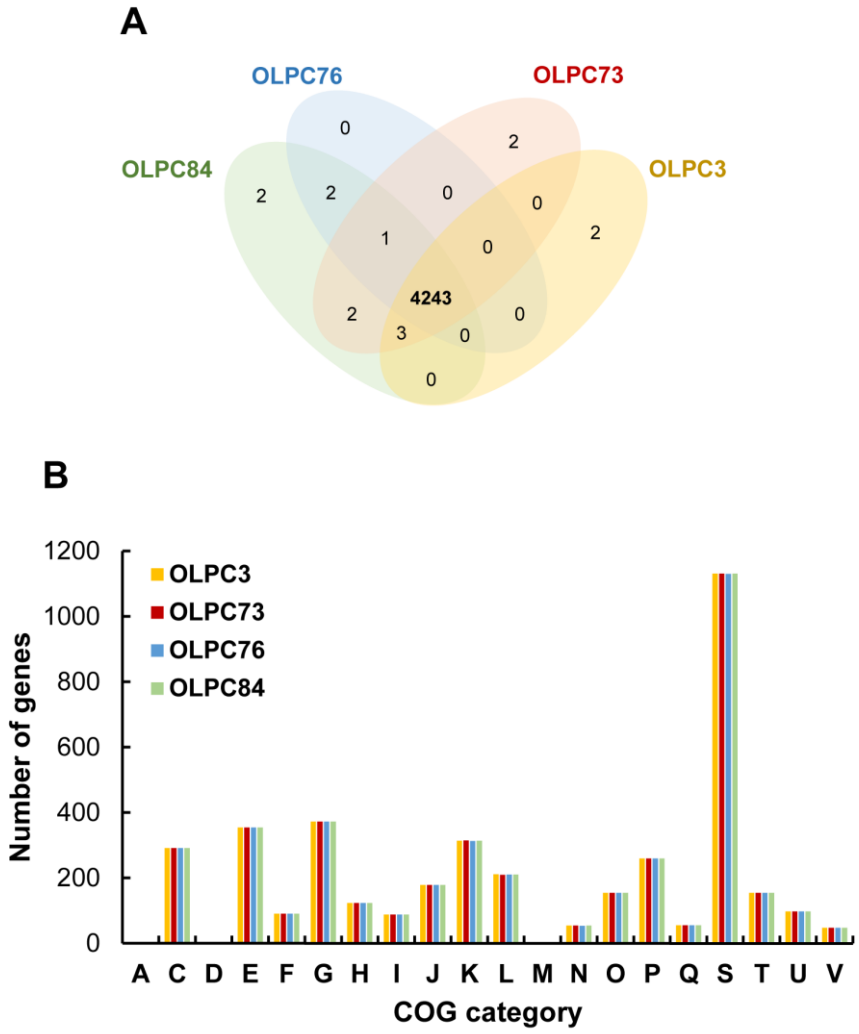


Figure 15. Comparative genomic analysis of 4 *E. coli* strains. (A) Venn diagram shows the number of strain-specific gene in each genome. The number of core genome is represented in bold. (B) COG distribution in the pan genome of 4 *E.coli* isolates. A: RNA processing and modification; C: energy production and conversion; D: cell cycle control, cell division, chromosome partitioning; E: amino acid transport and metabolism; F: nucleotide transport and

metabolism; G: carbohydrate transport and metabolism; H: coenzyme transport and metabolism; I: lipid transport and metabolism; J: translation, ribosomal structure and biogenesis; K: transcription; L: replication, recombination and repair; M: cell wall/membrane/envelope biogenesis; N: cell motility; O: posttranscriptional modification, protein turnover, chaperones; P: inorganic ion transport and metabolism; Q: secondary metabolites biosynthesis, transport and catabolism; S: function unknown; T: signal transduction mechanisms; U: intracellular trafficking, secretion, and vesicular transport; V: defense mechanisms.

5.3. Phylogenetic relationships between strains of *E. coli* group

To infer the evolutionary relationships of 4 clinical isolates among *E. coli* group, multiple sequence alignments of the whole genome were conducted. The sequence of total 15 strains in *E. coli* group were used and listed in Table 4. A phylogenetic tree of 15 *Escherichia* species showed that 4 clinical isolates were highly similar to K12 strains, including subspecies MG1655 and ER3454, followed by labolatory strain BL21 (DE3). For this reason, four clinical isolates of *E. coli* were named as K12-5.1 (OLPC3), K12-7.1 (OLPC73), K12-7.2 (OLPC76), and K12-7.3 (OLPC84), respectively (Figure 16).

Table 4. General features of *E. coli* group strains

Species	Strain	Source of isolation	Genome size (bp)	No. of contigs	G+C content (%)	Plasmid	Pathogenicity
Escherichia coli	OLPC3	Tissue of OLP lesion (human)	4,688,958	126	50.7	0	
Escherichia coli	OLPC73	Tissue of OLP lesion (human)	4,687,682	131	50.7	0	
Escherichia coli	OLPC76	Tissue of OLP lesion (human)	4,685,982	134	50.7	0	
Escherichia coli	OLPC84	Tissue of OLP lesion (human)	4,685,872	123	50.7	0	
Escherichia coli	K-12 substr. MG1655	Fecal (human)	4,641,652	1	50.8	0	Non-pathogenic
Escherichia coli	ATCC BAA-460 (Sakai)	Fecal (human)	5,594,477	3	50.5	2	Enterohemorrhagic
Escherichia coli	BL21(DE3)	Laboratory	4,558,947	1	50.8	0	Non-pathogenic
Shigella flexneri	2457T	Fecal (human)	4,599,354	1	50.9	0	Enterohemorrhagic
Escherichia albertii	KF1	Fecal (human)	4,701,875	1	49.7	0	Potentially enteropathogenic
Escherichia coli	Nissle 1917	Fecal (human)	5,441,200	1	50.6	0	Non-pathogenic
Escherichia coli	K-12 substr. ER3454	Unknown	4,619,729	1	50.8	0	Non-pathogenic
Escherichia coli	CFT073	Blood (human)	5,231,428	1	50.5	0	Uropathogenic
Escherichia coli	JJ2434	Unknown (human)	5,317,099	3	50.8	2	Pathogenic
Escherichia coli	13E0767	Cattle	4,942,246	1	50.6	0	Shiga toxin-producing
Escherichia coli	DSM 103246 (=E28)	Chicken carcass	5,445,834	1	50.5	0	Avian pahogenic

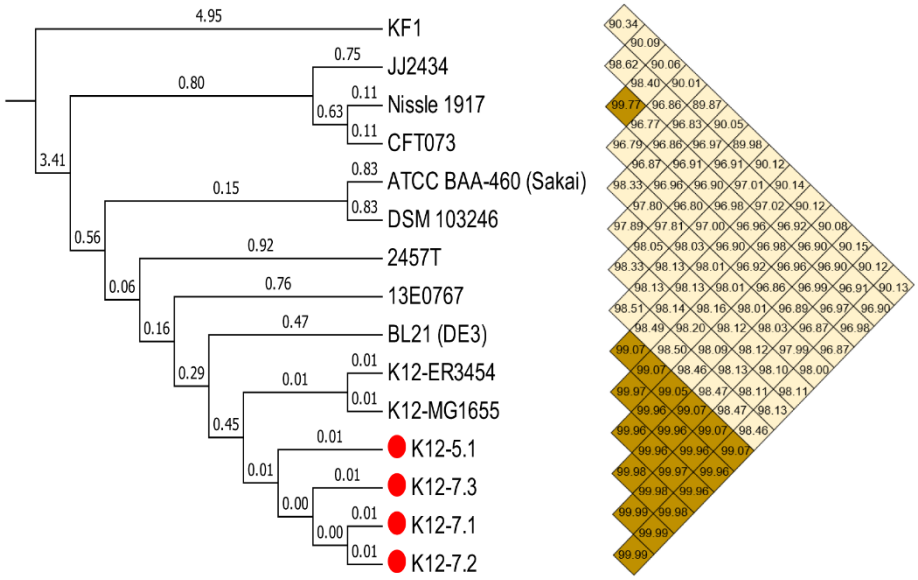


Figure 16. Phylogenic analysis of 15 strains in *E. coli* group. UPGMA dendrogram based on the OrthoANI values of 15 strains in *E. coli* group. The numbers on the branches show the branch length and the numbers in the heat-map indicate the OrthoANI values between two genomes.

5.4. Comparative genomic analysis of *E. coli* isolates

To characterize the genome contents of *E. coli* isolates, comparative genomic analysis was conducted using pathogenic Sakai and CFT073, commensal K12-MG1655, and lab-adapted BL21 (DE3) strains.

When compared with pathogenic strains (Sakai and CFT073), *E. coli* isolates (K12-5.1 and K12-7.1) had a higher number of shared-genes with

pathogenic strains than those of K12-MG1655 strain (Figure 17).

The presence or absence of genes encoding elements involved in flagella, fimbriae, secretion system, antibiotic resistance, and lipopolysaccharide biosynthesis was compared. The genes encoding flagella components were totally different between BL21 (DE3) and the others. All of 17 genes analyzed in this study were lacked in BL21 (DE3) strain, and 4 isolates contained all of genes related flagella expression. When the secretion systems were examined, the results showed that 4 isolates contains genes related with type II secretion system (*gspD* and *gspH*). In 8 *E. coli* strains, 12 genes involved in antibiotic resistance were detected. Sakai, CFT073, and 4 isolates contained 9, 11, and 7 genes encoding for antibiotics resistance, respectively. Finally, the genes involved in lipopolysaccharide (LPS) biosynthesis were compared. K12 strains including 4 isolates showed similar profiles in expression of LPS-related genes, presenting a different profiles from pathogens Sakai and CFT073, and commensal BL21 strains (Figure 18).

It has been reported that uropathogenic *E. coli* (UPEC) forms biofilm-like intracellular bacterial communities (IBCs), which protect their bodies against immune cells, environmental stress, and antibiotics during urinary tract infection (65). The expression of genes encoding elements involved in the formation of IBCs (65-76) may contribute to persistence of bacteria within the tissue. CFT073 strain contained various genes for IBC formation in their genome, but K12 strains and BL21 had a small number of genes than those in

CFT073. Among 32 genes examined in this study, Sakai strain also had 22 genes (Figure 19).

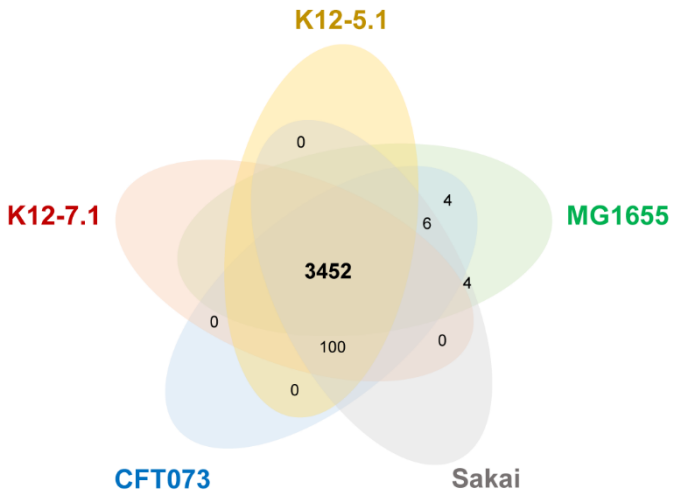


Figure 17. Comparative genomic analysis of 5 *E. coli* strains. Venn diagram shows the number of shared-genes among strains. The number of core genome is represented in bold.

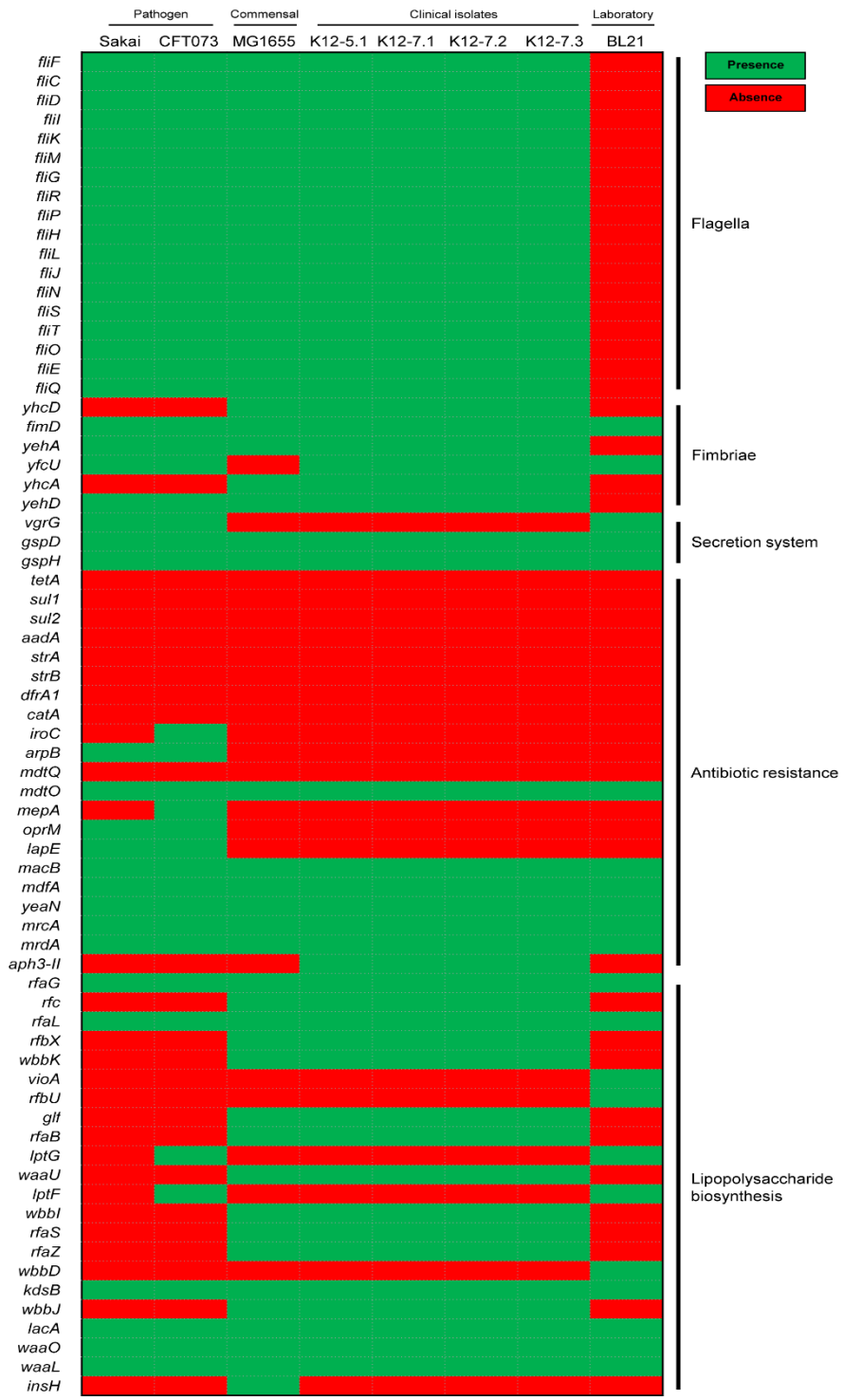

Pathogen
Commensal
Clinical isolates
Laboratory
Sakai
CFT073
MG1655
K12-5.1
K12-7.1
K12-7.2
K12-7.3
BL21
Presence
Absence
fliF
fliC
fliD
fliI
fliK
fliM
fliG
fliR
fliP
fliH
fliL
fliJ
fliN
fliS
fliT
fliO
fliE
fliQ
yhcD
fimD
yehA
yfcU
yhcA
yehD
vgrG
gspD
gspH
tetA
sul1
sul2
aadA
strA
strB
dfrA1
catA
iroC
arpB
mdtQ
mdtO
mepA
oprM
lapE
macB
mdfA
yeaN
mrcA
mrdA
aph3-II
rfaG
rfc
rfaL
rfbX
wbbK
vioA
rfbU
glf
rfaB
lptG
waaU
lptF
wbbI
rfaS
rfaZ
wbbD
kdsB
wbbJ
lacA
waaO
waaL
insH
Flagella
Fimbriae
Secretion system
Antibiotic resistance
Lipopolysaccharide biosynthesis

Figure 18. Comparison of genes involved in flagella, fimbriae, secretion system, antibiotic resistance, and LPS biosynthesis in *E. coli*. The presence or absence of bacterial genes encoding for flagella, fimbriae, secretion system, antibiotic resistance, and LPS biosynthesis in *E. coli* were compared. The color indicates the presence or absence of gene; green indicates the presence of gene, red indicates the absence of gene.

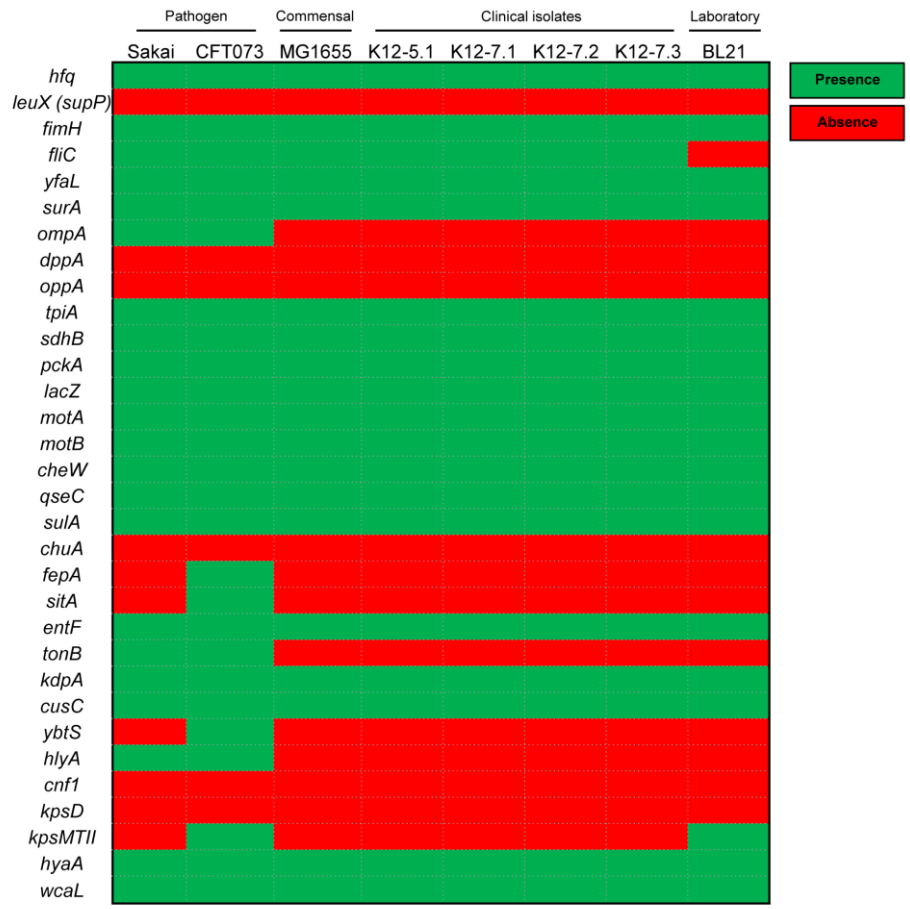


Figure 19. Genes encoding elements involved in formation of intracellular bacterial communities (IBC). Various genes known as involving in the formation of intracellular bacterial communities were compared. CFT073 strain has been known as uropathogenic *E. coli* (UPEC) and can form the IBC during urinary tract infection. The color indicates the presence or absence of gene; green indicates the presence of gene, red indicates the absence of gene.

6. Development of animal model for OLP

To establish the animal model for OLP, female ICR mice were orally inoculated with 10^9 viable *E. coli* isolate K12-7.2. After inoculation of *E. coli*, the immune cell infiltration and *E. coli* invasion within the tongue tissue of mice were examined by H&E stain and *in situ* hybridization using a DIG-labeled *E. coli*-specific probe. After preparation for DIG-labeled *E.coli*-specific probe, the sensitivity and specificity of probe were confirmed by dot blot assay. A probe against *E. coli* 16S rRNA was less sensitive than the control DIG-labeled probe provided in the kit, and further experiment for *in situ* hybridization was performed using 50 ng/μl of *E. coli*-specific probe (Figure 20, upper). In addition, this probe was found to specifically hybridize to *E. coli* among the bacterial gDNA of nine oral bacterial species, however, an *E. coli*-specific probe showed a differences in the hybridization efficacy depending on *E. coli* strain (Figure 20, bottom).

Using this probe, the bacterial invasion within the tongue tissue of mice was examined by *in situ* hybridization. The *E. coli* invasion within the tissue was not different between *E. coli*-inoculated and sham mice. In addition, the immune infiltration into the tissue was not observed (Figure 21).

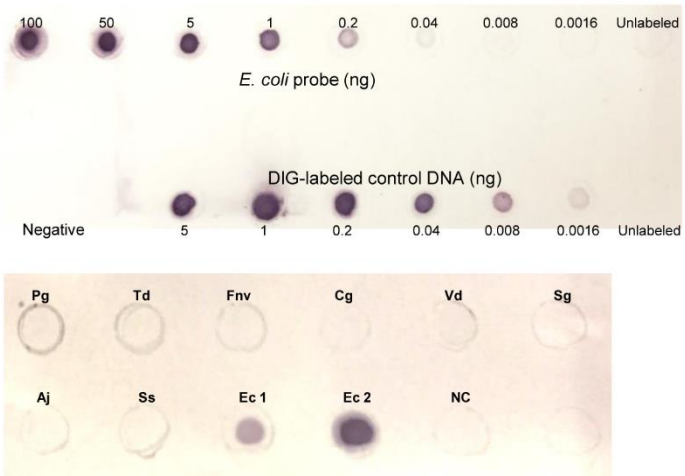


Figure 20. Sensitivity and specificity test of *E. coli*-specific probe. The sensitivity (upper) and specificity (bottom) of digoxigenin-labeled probe for *E. coli* were confirmed by dot blot assay. Bacterial gDNA from nine bacterial species were used for specificity test. Pg: *Porphyromonas gingivalis*; Td: *Treponema denticola*; Fnv: *Fusobacterium nucleatum* Sub. *vincentii*; Cg: *Capnocytophaga gingivalis*; Vd: *Veillonella dispar*; Sg: *Streptococcus gordonii*; Aj: *Acinetobacter johnsonii*; Ss: *Streptococcus salivarius*; Ec1: *E. coli* SG13009; Ec2: *E. coli* K12-5.1; NC: negative control.

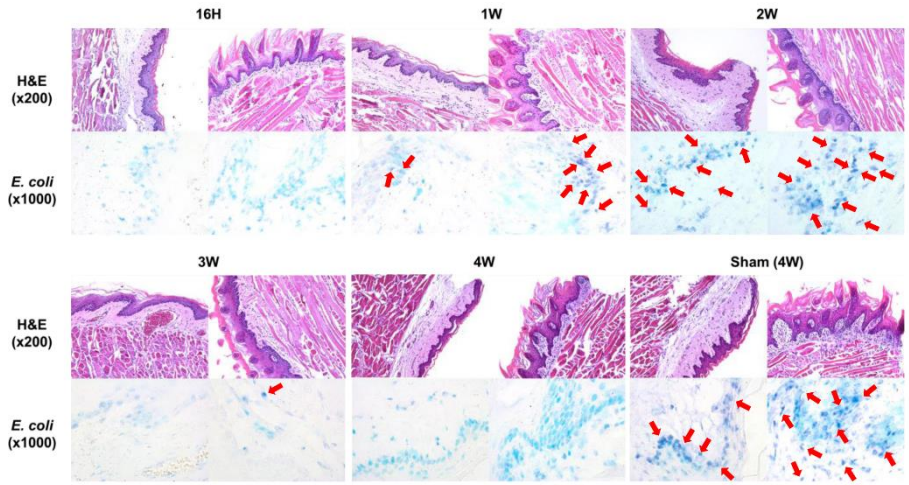


Figure 21. Histological images of animal experiment. Mice were orally inoculated with 10^9 viable *E. coli* K12-7.2. Tongue sections from mice were stained with H&E or subjected to *in situ* hybridization with an *E. coli*-specific probe. As a negative control, hybridization was also performed with the probe mixed with a 10-fold excess of unlabeled probe. Positive signals were shown in violet (red arrows).

IV. Discussion

The present study compared that the bacterial communities between the mucosal surface and tissues of OLP lesions in order to understand the role of oral bacteria in the pathogenesis of OLP. Analysis of bacterial communities by sequencing of 16S rRNA gene revealed that the alterations in the oral microbiota of OLP patients, as well as within the tissues of OLP lesions. Additionally, decreased total bacterial loads and increased inter-subject variability were observed in OT community as compared to the OM community. Moreover, the relative abundance of *E. coli* within the tissues was uniquely higher than on the mucosal surface.

Previously, only 4 research groups have studied the oral microbiota profiles of saliva and mucosa in OLP patients by high-throughput sequencing (1, 28-30). Some results of the present study concur with previous findings despite differences in the populations, sampled site, and the sequenced region of 16S rRNA gene. For instance, a tendency for increased bacterial diversity, a decrease in *Streptococcus*, and an increase in *Leptotrichia* were observed in OM compared to HM (40). Therefore, these are considered to be common features of the mucosal surface compared with healthy controls.

In terms of bacterial composition, the most abundant phyla of the control mucosa were Firmicutes, Proteobacteria, Actinobacteria, Bacteroidetes, and Fusobacteria, which composed over approximately 99% of the total bacteria.

This bacterial composition of the control mucosa corresponded with the results of Human Microbiome Project (HMP) (77). He et al. concluded that 4 genera, including *Actinomyces*, *Veillonella*, *Lautropia,* and *Leptotrichia,* are the core microbiota in the buccal mucosa highly associated with OLP, and *Actinobacillus* is the healthy mucosa-specific genera (30). In the present study, the relative abundances of *Actinomyces*, *Veillonella*, and *Lautropia* in OM were not different compared to HM, while there were significant differences in the abundance of *Leptotrichia* between HM and OM. This discrepancy may be attributed to differences in the sampling method, ethnicity, and sequenced regions of the 16S rRNA gene. In addition, present results at the species level were quite different from previously reported findings. Choi et al. reported that several species associated with gingivitis or periodontitis including *F. nucleatum*, *Neisseria oralis*, *C. gingivalis*, and *Eikenella corrodens* showed a different distribution in the OLP mucosa compared to control (28). Although this study examined the analysis together with the data of Choi et al. (28), the relative abundances of these 4 species were not different between HM and OM. However, the abundances of *T. denticola* and *Selenomonas sputigena* were increased in OM, and might be due to inter-subject variability. It is known that the periodontal status of OLP patients is worse compared to the control group (78, 79). In addition, the prevalence of the periodontitis-associated bacteria, including *A. actinomycetemcomitans*, *P. gingivalis*, *P. intermedia*, *T. forsythia*, and *T. denticola* was higher in the subgingival plaque of OLP patients with

gingivitis or periodontitis compared to those in non-OLP patients (27). Although the periodontal status of OLP patients was not evaluated in the present study, an increase in the relative abundance of *T. denticola* and *S. sputigena* in the OM community compared with HM, may reflect poor oral hygiene in OLP patients. However, all the gingivitis/periodontitis-associated bacteria (12 species, Table 2) were significantly decreased in the OT community compared with OM. These findings can be explained by two hypotheses. The first is that the gingivitis/periodontitis-associated bacteria may not be associated with the pathogenesis of OLP. The second is that they play a role only in the initiation of OLP development. To better understand the role of periodontal bacteria in the pathogenesis of OLP, further research is needed.

The most interesting finding of this study was the enrichment of *E. coli* within the tissues of OLP lesions, despite their decreased relative abundance in the mucosal community of OLP patients compared with that of healthy controls. *E. coli* is a gram-negative, rod-shaped facultative anaerobic bacteria, which is a common microflora of the gastrointestinal tract, and is rarely found in the oral cavity. It is well-known that *E. coli* show considerable genetic and phenotypic diversity depending on the strains. In addition, *E. coli* has various virulence factors including fimbrial adhesins, secretion systems, and toxins (80). Phylogenetic analysis in this study revealed that the four isolated strains of *E. coli* belonged to avirulent K-12 strains. Among the bacterial components, LPS of *E. coli* consists of lipid A, core oligosaccharide, and O-antigen (80).

Generally, all strains of *E. coli* K-12 are characterized by a rough phenotype due to a mutation within the *rfb* locus. Among the LPS components, the enzymes responsible for O-antigen biosynthesis are encoded in the *rfb* locus. However, *E. coli* K-12 is unable to synthesize the O-antigen due to the disruption of rhamnosyl transferase gene (*wbbL*) by the insertion sequence (IS) 5 element (81). It has been reported that the complementation of this mutation in *E. coli* K-12 MG1655 results in the expression of O-antigen. In addition, the complemented MG1655 could invade the gut of *Caenorhabditis elegans* and kill them (82). In the comparative genomic analysis, only MG1655 but not the OLP clinical strains contained the *insH* gene encoding the IS5 element. This result suggests that the 4 *E. coli* strains may have potential invasive and survival abilities, unlike the K-12 MG1655 strain.

Among the pathotypes of *E. coli*, UPEC CFT073 as well as UTI89, which are causative pathogens for urinary tract infection (UTI) in females, can create specific biofilm-like intracellular bacterial communities (IBCs) within the superficial facet cells of the bladder, and are able to replicate rapidly (83, 84). IBC provide a safety niche for the replication and survival of the bacteria, resulting in their persistence within the bladder. Since the persistence of bacteria within the tissue can cause chronic inflammation, the formation of IBC within OLP lesions may provide a possible etiology for OLP. The comparative genomic analysis showed that the 4 clinical isolates also contained several genes related with IBC formation in their genome, indicating their IBC forming

potential. The ability of *E. coli* to form IBC was not examined in this study, and further studies are required to clarify the pathogenesis of OLP.

In the animal models developed for OLP, infection with *E. coli* did not result in immune cell infiltration. In China, investigators attempted to develop an animal model for OLP by injecting the tissue homogenate of OLP lesion obtained from OLP patients into the hamster's oral submucosa. However, no significant differences were observed between the control and injected groups (85). To further develop the animal model, various factors including the genetic background of mice, infection route, and cell number of *E. coli* must be considered.

The present study has several limitations. First, the functional role of *E. coli* in the pathogenesis of OLP was not examined. Based on the proposed OLP pathogenesis model (Figure 22), further research is needed to clarify the relationship between *E. coli* and the development of OLP. Second, the presence of *E. coli* within the tissues of OLP patients was not confirmed. The amount of *E. coli* within the lamina propria of tissues is needed to be compared between control subjects and patients. Third, two different sequencing platforms had to be used to analyze the microbial community since Roche 454 was discontinued after the analysis of OLP1, 2, and 4 samples. It has been reported that the sequencing accuracy, coverage rate, and performance metrics can differ depending on the sequencing platforms (86). Allali et al. reported that despite similar biological conclusions, different platforms (Illumina MiSeq, Roche 454

GS FLX Titanium, and Ion Torrent PGM) showed differences in the beta diversity and abundance of specific taxa from the same samples (87). Despite this limitation, the bacterial composition of the mucosal surface and tissue was distinguished by the two sequencing platforms.

In case of the OLP4 patient, *E. coli* was not detected in the second tissue sample by sequencing, while the REU score was increased at the second visit. During the hospital visit, the OLP4 patient showed a change in the epithelial morphology with a thickened white plaque. Thus, a biopsy was performed to exclude cancerous transformation, and *E. coli* was not found in this tissue sample. In addition, the tissue samples from OLP 6, 7, and 9-11 have shown a less than 5% relative abundance of *E. coli*. These results can be explained by three hypotheses: First is that multiple species apart from *E. coli*, may be associated with the pathogenesis of OLP. Second is that OLP development may result in the enrichment of *E. coli* within the tissue. Third is that the core microbiota may undergo dynamic changes during the development of OLP.

In conclusion, this is the first study to identify a difference in the bacterial communities between the mucosal surface and within the tissues of OLP patients. Understanding the role of *E. coli* in the development of OLP may provide a new insight into the etiopathogenesis of OLP.

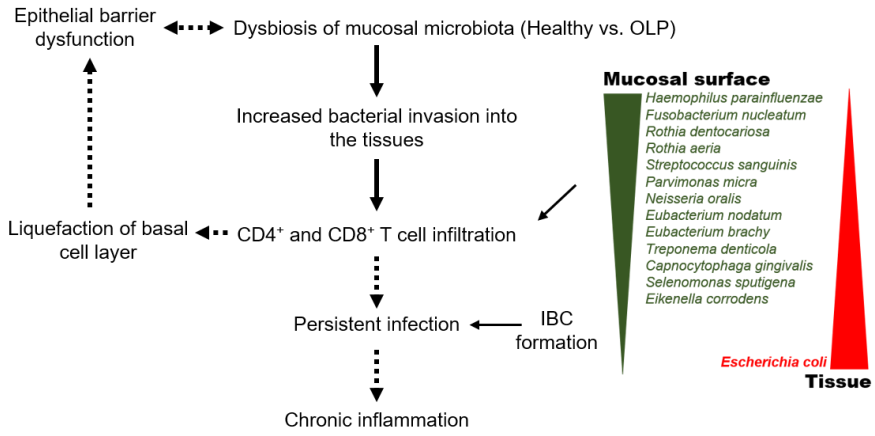


Figure 22. Proposed pathogenesis model for OLP. The diagram is adapted from previous published figure (28). Thick and dashed arrows indicate proved and unproved relations, respectively. Thin arrows indicate proved and expected relations based on results in current study.

V. Reference

1. Wang K, Lu WX, Tu QC, Ge YC, He JZ, Zhou Y, et al. **Preliminary analysis of salivary microbiome and their potential roles in oral lichen planus**. *Sci Rep*. 2016;6:22943.

2. Chiang CP, Chang JYF, Wang YP, Wu YH, Lu SY, Sun A. **Oral lichen planus - Differential diagnoses, serum autoantibodies, hematinic deficiencies, and management**. *J Formos Med Assoc*. 2018;117(9):756-65.

3. Alrashdan MS, Cirillo N, McCullough M. **Oral lichen planus: a literature review and update**. *Arch Dermatol Res*. 2016;308(8):539-51.

4. Burkhart NW, Burker EJ, Burkes EJ, Wolfe L. **Assessing the characteristics of patients with oral lichen planus**. *J Am Dent Assoc*. 1996;127(5):648-62.

5. Carrozzo M, de Capei MU, Dametto E, Fasano ME, Arduino P, Broccoletti R, et al. **Tumor necrosis factor-alpha and interferon-gamma polymorphisms contribute to susceptibility to oral lichen planus**. *J Invest Dermatol*. 2004;122(1):87-94.

6. Sun A, Wu YC, Wang JT, Liu BY, Chiang CP. **Association of HLA-te22 antigen with anti-nuclear antibodies in Chinese patients with erosive oral lichen planus**. *Proc Natl Sci Counc Repub China B*. 2000;24(2):63-9.

7. Porter K, Klouda P, Scully C, Bidwell J, Porter S. **Class I and II HLA antigens in British patients with oral lichen planus**. *Oral Surg Oral Med Oral Pathol*. 1993;75(2):176-80.

8. Watanabe T, Ohishi M, Tanaka K, Sato H. **Analysis of HLA antigens in Japanese with oral lichen planus**. *J Oral Pathol*. 1986;15(10):529-33.

9. McCartan BE, Lamey PJ. **Expression of CD1 and HLA-DR by Langerhans cells (LC) in oral lichenoid drug eruptions (LDE) and idiopathic oral lichen planus (LP)**. *J Oral Pathol Med*. 1997;26(4):176-80.

10. Ognjenovic M, Karelovic D, Cindro VV, Tadin I. **Oral lichen planus and HLA A**. *Coll Antropol*. 1998;22 Suppl:89-92.

11. Femiano F, Scully C. **Functions of the cytokines in relation oral lichen planus-hepatitis C**. *Med Oral Patol Oral Cir Bucal*. 2005;10 Suppl 1:E40-4.

12. Kis A, Feher E, Gall T, Tar I, Boda R, Toth ED, et al. **Epstein-Barr virus prevalence in oral squamous cell cancer and in potentially malignant oral disorders in an eastern Hungarian population**. *Eur J Oral Sci*. 2009;117(5):536-40.

13. Adtani P, Malathi N. **Epstein-Barr virus and its association with rheumatoid arthritis and oral lichen planus**. *J Oral Maxillofac Pathol*. 2015;19(3):282-5.

14. Vieira Rda R, Ferreira LL, Biasoli ER, Bernabe DG, Nunes CM, Miyahara GI. **Detection of Epstein-Barr virus in different sources of materials from patients with oral lichen planus: a case-control study**. *J Clin Pathol*. 2016;69(4):358-63.

15. Ghannoum MA, Jurevic RJ, Mukherjee PK, Cui F, Sikaroodi M, Naqvi A, et al. **Characterization of the oral fungal microbiome (mycobiome) in healthy individuals**. *PLoS Pathog*. 2010;6(1):e1000713.

16. Kleinegger CL, Lockhart SR, Vargas K, Soll DR. **Frequency, intensity, species, and strains of oral *Candida* vary as a function of host age**. *J Clin Microbiol*. 1996;34(9):2246-54.

17. Darwazeh AMG, Al-Refai S, Al-Mojaiwel S. **Isolation of *Candida* species from the oral cavity and fingertips of complete denture wearers**. *J Prosthet Dent*. 2001;86(4):420-3.

18. Nittayananta W, Jealae S, Winn T. **Oral *Candida* in HIV-infected heterosexuals and intravenous drug users in Thailand**. *J Oral Pathol Med*. 2001;30(6):347-54.

19. Simon M, Jr., Hornstein OP. **Prevalence rate of *Candida* in the oral cavity of patients with oral lichen planus**. *Arch Dermatol Res*. 1980;267(3):317-8.

20. Artico G, Freitas RS, Santos AM, Benard G, Romiti R, Migliari DA. **Prevalence of *Candida* spp., xerostomia, and hyposalivation in oral lichen planus - A controlled study**. *Oral Dis*. 2014;20(3):e36-e41.

21. Rautemaa R, Rusanen P, Richardson M, Meurman JH. **Optimal sampling site for mucosal candidosis in oral cancer patients is the labial sulcus**. *J Med Microbiol*. 2006;55(10):1447-51.

22. Li C, Ha T, Ferguson DA, Jr., Chi DS, Zhao R, Patel NR, et al. **A newly developed PCR assay of *H. pylori* in gastric biopsy, saliva, and feces. Evidence of high prevalence of *H. pylori* in saliva supports oral transmission**. *Dig Dis Sci*. 1996;41(11):2142-9.

23. Shimoyama T, Horie N, Kato T, Kaneko T, Komiyama K. ***Helicobacter pylori* in oral ulcerations**. *J Oral Sci*. 2000;42(4):225-9.

24. Riggio MP, Lennon A, Wray D. **Detection of *Helicobacter pylori* DNA in recurrent aphthous stomatitis tissue by PCR**. *J Oral Pathol Med*. 2000;29(10):507-13.

25. Attia EA, Abdel Fattah NS, Abdella HM. **Upper gastrointestinal findings and detection of *Helicobacter pylori* in patients with oral lichen planus**. *Clin Exp Dermatol*. 2010;35(4):355-60.

26. Bornstein MM, Hakimi B, Persson GR. **Microbiological findings in subjects with asymptomatic oral lichen planus: a cross-sectional comparative study**. *J Periodontol*. 2008;79(12):2347-55.

27. Ertugrul AS, Arslan U, Dursun R, Hakki SS. **Periodontopathogen profile of healthy and oral lichen planus patients with gingivitis or periodontitis**. *Int J Oral Sci*. 2013;5(2):92-7.

28. Choi YS, Kim Y, Yoon HJ, Baek KJ, Alam J, Park HK, et al. **The presence of bacteria within tissue provides insights into the pathogenesis of oral lichen planus**. *Sci Rep*. 2016;6:29186.

29. Kragelund C, Keller MK. **The oral microbiome in oral lichen planus during a 1-year randomized clinical trial**. *Oral Dis*. 2018;25(1):327-38.

30. He Y, Gong D, Shi C, Shao F, Shi J, Fei J. **Dysbiosis of oral buccal mucosa microbiota in patients with oral lichen planus**. *Oral Dis*. 2017;23(5):674-82.

31. Farhi D, Dupin N. **Pathophysiology, etiologic factors, and clinical management of oral lichen planus, part I: facts and controversies**. *Clin Dermatol*. 2010;28(1):100-8.

32. Le Cleach L, Chosidow O. **Clinical practice. Lichen planus**. *N Engl J Med*. 2012;366(8):723-32.

33. Sugerman PB, Savage NW, Walsh LJ, Zhao ZZ, Zhou XJ, Khan A, et al. **The pathogenesis of oral lichen planus**. *Crit Rev Oral Biol Med*. 2002;13(4):350-65.

34. Zhao ZZ, Sugerman PB, Zhou XJ, Walsh LJ, Savage NW. **Mast cell degranulation and the role of T cell RANTES in oral lichen planus**. *Oral Dis*. 2001;7(4):246-51.

35. Sharma R, Sircar K, Singh S, Rastogi V. **Role of mast cells in pathogenesis of oral lichen planus**. *J Oral Maxillofac Pathol*. 2011;15(3):267-71.

36. Lukac J, Brozovic S, Vucicevic-Boras V, Mravak-Stipetic M, Malenica B, Kusic Z. **Serum autoantibodies to desmogleins 1 and 3 in patients with oral lichen planus**. *Croat Med J*. 2006;47(1):53-8.

37. Lundstrom IM. **Serum immunoglobulins and autoantibodies in patients with oral lichen planus**. *Int J Oral Surg*. 1985;14(3):259-68.

38. Wu KM, Wang YP, Lin HP, Chen HM, Chia JS, Sun A. **Modulation of serum smooth muscle antibody levels by levamisole treatment in patients with oral lichen planus**. *J Formos Med Assoc*. 2013;112(6):352-7.

39. Al-Hashimi I, Schifter M, Lockhart PB, Wray D, Brennan M, Migliorati CA, et al. **Oral lichen planus and oral lichenoid lesions: diagnostic and therapeutic considerations**. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2007;103 Suppl:S25 e1-12.

40. Baek K, Choi Y. **The microbiology of oral lichen planus: Is microbial infection the cause of oral lichen planus?**. *Mol Oral Microbiol*. 2018;33(1):22-8.

41. Mizuki H, Abe R, Kogi S, Mikami T. **Immunohistochemical detection of *Mycoplasma salivarium* in oral lichen planus tissue**. *J Oral Pathol Med*. 2017;46(8):649-56.

42. Wade WG. **The oral microbiome in health and disease**. *Pharmacol Res*. 2013;69(1):137-43.

43. Nath SG, Raveendran R. **Microbial dysbiosis in periodontitis**. *J Indian Soc Periodontol*. 2013;17(4):543-5.

44. Tanner ACR, Kressirer CA, Rothmiller S, Johansson I, Chalmers NI. **The caries microbiome: Implications for reversing dysbiosis**. *Adv Dent Res*. 2018;29(1):78-85.

45. Ramon-Fluixa C, Bagan-Sebastian J, Milian-Masanet M, Scully C. **Periodontal status in patients with oral lichen planus: a study of 90 cases**. *Oral Dis*. 1999;5(4):303-6.

46. Tonetti MS, Jepsen S, Jin L, Otomo-Corgel J. **Impact of the global burden of periodontal diseases on health, nutrition and wellbeing of mankind: A call for global action**. *J Clin Periodontol*. 2017;44(5):456-62.

47. Lamont RJ, Hajishengallis G. **Polymicrobial synergy and dysbiosis in inflammatory disease**. *Trends Mol Med*. 2015;21(3):172-83.

48. Mohammad AR, Brunsvold M, Bauer R. **The strength of association between systemic postmenopausal osteoporosis and periodontal disease**. *Int J Prosthodont*. 1996;9(5):479-83.

49. Moutsopoulos NM, Madianos PN. **Low-grade inflammation in chronic infectious diseases: paradigm of periodontal infections**. *Ann N Y Acad Sci*. 2006;1088:251-64.

50. Socransky SS, Haffajee AD, Cugini MA, Smith C, Kent RL, Jr. **Microbial complexes in subgingival plaque**. *J Clin Periodontol*. 1998;25(2):134-44.

51. Mishima E, Sharma A. ***Tannerella forsythia*** **invasion in oral epithelial cells requires phosphoinositide 3-kinase activation and clathrin-mediated endocytosis**. Microbiology. 2011;157(Pt 8):2382-91.

52. Deshpande RG, Khan MB, Genco CA. **Invasion of aortic and heart endothelial cells by *Porphyromonas gingivalis***. *Infect Immun*. 1998;66(11):5337-43.

53. Gawron K, Bereta G, Nowakowska Z, Lazarz-Bartyzel K, Lazarz M, Szmigielski B, et al. **Peptidylarginine deiminase from *Porphyromonas gingivalis* contributes to infection of gingival fibroblasts and induction of**

prostaglandin E2 -signaling pathway. *Mol Oral Microbiol*. 2014;29(6):321-32.

54. Shin J, Choi Y. **The fate of *Treponema denticola* within human gingival epithelial cells**. *Mol Oral Microbiol*. 2012;27(6):471-82.

55. Lamont RJ, Chan A, Belton CM, Izutsu KT, Vasel D, Weinberg A. ***Porphyromonas gingivalis* invasion of gingival epithelial cells**. *Infect Immun*. 1995;63(10):3878-85.

56. Weinberg A, Belton CM, Park Y, Lamont RJ. **Role of fimbriae in *Porphyromonas gingivalis* invasion of gingival epithelial cells**. *Infect Immun*. 1997;65(1):313-6.

57. Ji S, Choi YS, Choi Y. **Bacterial invasion and persistence: critical events in the pathogenesis of periodontitis?**. *J Periodontal Res*. 2015;50(5):570-85.

58. Baek KJ, Choi YS, Kang CK, Choi Y. **The proteolytic activity of *Porphyromonas gingivalis* is critical in a murine model of periodontitis**. *J Periodontol*. 2017;88(2):218-24.

59. Choi YS, Kim YC, Ji S, Choi Y. **Increased bacterial invasion and differential expression of tight-junction proteins, growth factors, and growth factor receptors in periodontal lesions**. *J Periodontol*. 2014;85(8):e313-22.

60. Baek K, Ji S, Choi Y. **Complex intratissue microbiota forms biofilms in periodontal lesions**. *J Dent Res*. 2018;97(2):192-200.

61. Chao A. **Nonparametric-estimation of the number of classes in a population**. *Scand J Stat*. 1984;11(4):265-70.

62. Simpson EH. **Measurement of diversity**. *Nature*. 1949;163:688.

63. Shannon CE. **A mathematical theory of communication**. *Bell Syst. Tech. J*. 1948;27(4):623-56.

64. Kaper JB, Nataro JP, Mobley HL. **Pathogenic *Escherichia coli***. *Nat Rev Microbiol*. 2004;2(2):123-40.

65. Berry RE, Klumpp DJ, Schaeffer AJ. **Urothelial cultures support intracellular bacterial community formation by uropathogenic *Escherichia coli***. *Infect Immun*. 2009;77(7):2762-72.

66. Justice SS, Hunstad DA, Seed PC, Hultgren SJ. **Filamentation by *Escherichia coli* subverts innate defenses during urinary tract infection**. *P Natl Acad Sci USA*. 2006;103(52):19884-9.

67. Ejrnaes K, Stegger M, Reisner A, Ferry S, Monsen T, Holm SE, et al. **Characteristics of *Escherichia coli* causing persistence or relapse of urinary tract infections phylogenetic groups, virulence factors and biofilm formation**. *Virulence*. 2011;2(6):528-37.

68. Alteri CJ, Smith SN, Mobley HLT. **Fitness of *Escherichia coli* during urinary tract infection requires gluconeogenesis and the TCA cycle**. *PLoS Pathog*. 2009;5(5):e1000448.

69. Kostakioti M, Hadjifrangiskou M, Pinkner JS, Hultgren SJ. **QseC-mediated dephosphorylation of QseB is required for expression of genes associated with virulence in uropathogenic *Escherichia coli***. *Mol Microbiol*. 2009;73(6):1020-31.

70. Rouviere PE, Gross CA. **SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins**. *Genes Dev*. 1996;10(24):3170-82.

71. Robino L, Scavone P, Araujo L, Algorta G, Zunino P, Pirez MC, et al. **Intracellular bacteria in the pathogenesis of *Escherichia coli* urinary tract infection in children**. *Clin Infect Dis*. 2014;59(11):e158-64.

72. Hunstad DA, Justice SS. **Intracellular lifestyles and immune evasion strategies of uropathogenic *Escherichia coli***. *Annu Rev Microbiol*. 2010;64:203-21.

73. Lane MC, Lockatell V, Monterosso G, Lamphier D, Weinert J, Hebel JR, et al. **Role of motility in the colonization of uropathogenic *Escherichia coli* in the urinary tract**. *Infect Immun*. 2005;73(11):7644-56.

74. Kulesus RR, Diaz-Perez K, Slechta ES, Eto DS, Mulvey MA. **Impact of the RNA chaperone Hfq on the fitness and virulence potential of uropathogenic *Escherichia coli***. *Infect Immun*. 2008;76(7):3019-26.

75. Nomura T, Fujita N, Ishihama A. **Expression of the leuX gene in *Escherichia coli*. Regulation at transcription and tRNA processing steps**. *J Mol Biol*. 1987;197(4):659-70.

76. Snyder JA, Haugen BJ, Buckles EL, Lockatell CV, Johnson DE, Donnenberg MS, et al. **Transcriptome of uropathogenic *Escherichia coli* during urinary tract infection**. *Infect Immun*. 2004;72(11):6373-81.

77. Human Microbiome Project C. **Structure, function and diversity of the healthy human microbiome**. *Nature*. 2012;486(7402):207-14.

78. Azizi A, Rezaee M. **Comparison of periodontal status in gingival oral lichen planus patients and healthy subjects**. *Dermatol Res Pract*. 2012;2012:561232.

79. Lopez-Jornet P, Camacho-Alonso F. **Periodontal conditions in patients with oral lichen planus: a pilot study**. *Quintessence Int*. 2012;43(2):147-52.

80. Mainil J. ***Escherichia coli* virulence factors**. *Vet Immunol Immunopathol*. 2013;152(1-2):2-12.81. Jorgenson MA, Young KD. **Interrupting biosynthesis of O antigen or the lipopolysaccharide core produces morphological**

defects in *Escherichia coli* by sequestering undecaprenyl phosphate. *J Bacteriol*. 2016;198(22):3070-9.

82. Browning DF, Wells TJ, Franca FLS, Morris FC, Sevastsyanovich YR, Bryant JA, et al. **Laboratory adapted *Escherichia coli* K-12 becomes a pathogen of *Caenorhabditis elegans* upon restoration of O antigen biosynthesis**. *Mol Microbiol*. 2013;87(5):939-50.

83. Subashchandrabose S, Mobley HLT. **Virulence and fitness determinants of uropathogenic *Escherichia coli***. *Microbiol Spectr*. 2015;3(4).

84. Mysorekar IU, Hultgren SJ. **Mechanisms of uropathogenic *Escherichia coli* persistence and eradication from the urinary tract**. *Proc Natl Acad Sci U S A*. 2006;103(38):14170-5.

85. DongZuo. **Establishment of animal model of oral lichen planus and the detection and clinical significance of salivary enzymes, IL-18 in patients with oral lichen planus** [Thesis]: Hebei Medical University; 2014.

86. Harismendy O, Ng PC, Strausberg RL, Wang X, Stockwell TB, Beeson KY, et al. **Evaluation of next generation sequencing platforms for population targeted sequencing studies**. *Genome Biol*. 2009;10(3):R32.

87. Allali I, Arnold JW, Roach J, Cadenas MB, Butz N, Hassan HM, et al. **A comparison of sequencing platforms and bioinformatics pipelines for**

compositional analysis of the gut microbiome. *BMC Microbiol* 2017;17(1):194.

국문초록

구강편평태선 병인기전에 관여하는 세균 동정

백 금 진

서울대학교 대학원

치의과학과 면역 및 분자미생물학 전공

(지도교수: 최 영 님)

1. 목 적

구강편평태선은 T 림프구가 매개하는 만성 염증성 점막 상피 질환으로, 정확한 원인과 병인기전은 밝혀지지 않았다. 상피에 존재하는 미지의 항원에 의해 $CD8^+$ T 세포가 활성화되어 상피세포와 기저막을 공격하여 파괴되는 것으로 이해되고 있다. 자가 항원, 치과 수복물, 감염 등이 세포 독성 T 세포 반응을 유발하는 미지의 항원

으로 거론되고 있지만, 정확한 원인은 분명하지 않다.

세균 감염과 구강편평태선 발병 사이의 관계에 많은 연구자들이 주목하였고, 최근 염기 서열 분석 기술을 통해 구강편평태선 환자와 건강인의 타액 또는 점막 세균의 차이를 보여주는 연구 결과들이 보고되고 있다. 건강인과 비교하였을 때 구강편평태선 병소 조직 내에서 유의한 세균 증가가 관찰됨에도 불구하고, 구강편평태선 병소 조직 내에 존재하는 세균에 대해서는 보고된 바가 없다. 따라서 본 연구의 목적은 구강편평태선 병소 조직과 점막 표면에 존재하는 세균 공동체를 비교 분석하고, 구강편평태선 병인기전과 관련 있을 것이라 여겨지는 세균을 분리·동정 하는 것이며, 구강편평태선 동물 모델을 확립하는 것이다.

2. 방 법

구강편평태선 환자의 병소 점막 표면과 조직 내에 존재하는 세균의 조성을 비교하기위해, 11명의 구강편평태선 환자로부터 병소 점막에 존재하는 세균 또는 병소 부위 생검을 얻었으며, 두 샘플을 모두 얻은 환자는 총 7명이다. 이 중, 2명 환자의 조직은 세균 분석에 이용하지 않고 조직 내 세균을 분리하는데 이용하였다. 각 샘플로부터 세균의 DNA를 추출하여 세균 분석을 진행하였고, 조직 샘플의

경우 표면에 존재하는 세균을 제거하기위해서 DNA를 추출하기 전에 라이소자임, 항생제, DNA 분해효소를 처리하였다. 추출한 DNA는 세균의 16S rRNA 유전자 V1-V3 또는 V3-V4 가변부위를 중합효소 연쇄 반응으로 증폭시켜, 454 GS FLX Titanium 또는 Illumina MiSeq 시스템을 통해 염기 서열 분석 하였다. 세균 조성 분석을 위해, 이전 연구에서 사용한 11명의 건강인, 13명의 구강편평태선 환자의 점막 세균 데이터를 포함하여 분석하였다. 샘플 내 세균 양은 실시간 중합효소 연쇄 반응을 통하여 정량하였다.

병소 조직으로부터 대장균을 분리하기 위하여, 환자 2명의 조직을 이용하였다. 조직 표면의 세균을 제거한 후에, 조직을 잘게 잘라 액체 배지에서 하루 배양 하고, 배양액을 고체 선택배지에서 배양하였다. 고체 배지 위에 자란 미생물 집락을 대장균-특이적 프라이머를 이용하여 중합효소 연쇄 반응으로 확인한 후, 염기 서열 분석을 통해 세균을 동정하였다. 환자 조직에서 분리·동정한 4개의 대장균 (K12-5.1, 7.1, 7.2, 7.3) 으로부터 DNA를 추출하여 전체 게놈 염기 서열 분석에 이용하였다.

구강편평태선 동물 모델을 확립하기 위해, 살아있는 대장균 K12-7.2를 (10^9/마리 당) ICR 암컷 생쥐 구강 내로 일주일에 한 번씩 주입하였다. 기간 별로 세균의 조직 내 침투와 면역세포의

침윤을 관찰하기 위해, 세균 접종 16시간, 7, 14, 21, 28일 후에 생쥐를 희생하여 혀 조직을 얻고, 대장균-특이적 탐침자를 이용한 가시적 분자 결합화와 H&E 염색을 수행하였다.

3. 결 과

실시간 중합효소 연쇄 반응을 통해 전체 세균 양을 확인하였더니, 구강편평태선 환자의 점막 샘플에서의 세균 양이 건강인의 점막 샘플, 구강편평태선 환자의 조직 샘플에 비해 통계적으로 유의하게 높았다. 세균 조성을 문 (phylum), 속 (genus), 종 (species) 수준에서 비교하였더니, 3개의 문, 13개 속, 90개 종이 통계적으로 유의하게 세 그룹 간에서 차이를 보였다.

각 환자 내에서 점막 표면과 조직 내 존재하는 세균의 변화를 관찰하기 위해서 샘플을 모두 얻은 7명의 세균 데이터만을 다시 비교하였다. 실시간 중합효소 연쇄 반응을 통해 정량한 세균 양은 점막 표면 샘플 군보다 조직에서 유의하게 감소하였고, 세균 종 풍부도와 다양성 지표는 두 그룹간에 차이가 없었다. 그러나 세균 조성 비교에서는, 2개의 문, 8개 속, 52개 종이 통계적으로 유의하게 두 그룹간에 차이를 보였다. 두 그룹 간에서 차이를 보인 52개 종 중에서, 특히 대장균은 조직의 세균 공동체에서 유의한 증가를 보였다.

구강편평태선 환자 병소 조직으로부터 분리한 4개 대장균 임상 균주는 전제 유전체 분석을 통해 대장균 K-12 균주와 유사한

것을 확인하였다. 그러나 임상 균주는 대장균 K-12 MG1655와는 약간 다른 유전자 프로파일을 보였다.

동물 실험에서, 대장균 임상 균주를 구강 내로 주입한 그룹과 대조군 간에 세균의 조직 내 침투와 면역세포 침윤 정도를 비교하였다. 세균의 조직 내 침투는 큰 차이를 발견하지 못하였으며, 실험군에서 면역 세포의 침윤을 관찰하지 못하였다.

4. 결 론

본 연구에서는 최초로 구강편평태선 환자 병소 부위의 점막 표면과 조직 내 사이의 세균 조성 차이를 관찰하였다. 특히 점막 표면과 비교했을 때, 병소 조직 내에서 상대적 존재 비가 크게 증가한 대장균의 역할을 이해한다면, 구강편평태선의 병인기전을 이해하는 데 새로운 시각을 제공 할 수 있을 것이다.

주 요 어: 구강편평태선, 병인기전, 세균 공동체, 대장균, 비교유전체
학 번: 2016-30620