PPAD, *Porphyromonas gingivalis* and the Subgingival Microbiome in Periodontitis and Autoantibody-positive Individuals at Risk of Rheumatoid Arthritis

Zijian Cheng

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

School of Dentistry

September, 2018

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

It is a pleasure to thank a number of people who have supported me through my PhD study. This thesis could not have been completed without their support. Firstly, I would like to express my sincere gratitude and appreciation to my supervisors, Dr. Thuy Do, Dr. Josephine Meade, Prof. Paul Emery and Prof. Deirdre Devine for their valuable guidance, support and encouragement throughout my study.

I am grateful to my colleagues and all staff in the Division of Oral Biology who are friendly and supportive, for making an enjoyable environment during this study period. Special thanks are extended to Shabnum Rashid from Oral Microbiology group and David Sharples from Faculty of Biological Science for the training and help with my lab work. Additionally, my appreciation is extended to Prof. Phil Marsh for his valuable help and suggestions.

This research has been carried out by a team and I would like to thank all the clinicians, nurses and technicians from School of Dentistry, Leeds Dental Institute and Leeds Musculoskeletal Biomedical Research Unit, who have contributed to this project. I also wish to extend my special appreciation to the volunteers and patients who participated in this study.

Most importantly, I am indebted to the Chinese Scholarship Council for financial support. I would also like to thank my friends in UK for their friendship that has made my life enjoyable. Finally, I would like to show my gratitude to my parents and family for their continuous encouragement and support with patience and love

during this period. I would not have ventured to a foreign country for PhD study without their support.

Abstract

There is an epidemiological association between periodontitis and rheumatoid arthritis (RA). The subgingival microbiota may play an important role in the link between the two diseases. *Porphyromonas gingivalis*, which produces a peptidyl-arginine deiminase (PPAD) capable of citrullinating proteins, is considered a key organism inducing the production of antibodies against citrullinated proteins systemically and may initiate the pathogenic autoimmune responses associated with RA. The overall aim of this study was to explore the role of PPAD in *P. gingivalis* physiology and to better understand the links between *P. gingivalis*, periodontitis and risk of developing RA.

P. gingivalis W83 and the corresponding Δ*ppad* mutant were grown in batch and continuous culture, to assess pH regulation, bacterial growth, gene expression and arginine gingipain (Rgp) and dipeptidyl-peptidase (DPP) activities. In a collaborative clinical study, the shotgun metagenomic approach was used to observe subgingival microbial profiles in individuals with and without periodontitis, with and without RA, and in those with autoantibodies against citrullinated peptides (CCP) at risk of developing RA.

Based on *in vitro* studies, PPAD may citrullinate Rgp and DPP11, impair their activities and subsequently affect the alkali-promoting activity of *P. gingivalis*. Furthermore, both environmental pH and PPAD deficiency were able to regulate *P. gingivalis* gene expression, promoting adaptation to environmental changes and facilitating bacterial growth. In the clinical study, periodontitis occurred more often in anti-CCP positive at-risk individuals than in healthy controls and the subgingival

microbiomes of those individuals were perturbed, indicating that periodontitis and related microbial dysbiosis precede the onset of RA. *P. gingivalis* and its PPAD in established periodontitis conditions may play an important role in the initiation of RA. Moreover, PAD or PAD-like enzymes present in bacterial species other than *P. gingivalis*, e.g. *Prevotella* spp. exhibited some citrullination activity *in vitro* in a similar manner to PPAD.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Overview of periodontal disease

Periodontal disease, alongside dental caries, is one of the two most common and significant oral diseases contributing to the global burden of chronic disease (Pihlstrom et al., 2005, Bratthall et al., 2006). Periodontal disease is defined as the microbially-induced inflammatory conditions that causes damage to the gingivae (gums), periodontal ligament and alveolar bone, all of which form the supporting tissues of the teeth. The complex multi-factorial aetiology of periodontal disease is related to an imbalance between the resident subgingival microbial communities and the host responses to them. The bacterial biofilm (also called dental plaque) which forms on the surfaces of teeth, causes a chronic microbial stimulus that induces a local inflammatory response. In addition to pathogenic microorganisms in the biofilm, genetic and environmental factors such as smoking, contribute to the development of these diseases.

The term periodontal disease describes a spectrum of inflammatory conditions. Gingivitis, the mildest form of periodontal disease, is an inflammatory response to the accumulation of dental plaque at the gingival margin. It is reversible and can be eradicated by maintaining good oral hygiene. Gingivitis acts as a precursor for the initiation of periodontitis which is a more advanced inflammatory form of periodontal disease, although not all gingivitis progresses to periodontitis (Schatzle et al., 2009).

Unlike gingivitis, periodontitis causes irreversible tissue damage and gingival epithelial migration (Figure 1.1). Clinical manifestations of periodontitis include the

deepening of periodontal pockets and loss of attachment, progressively leading to loosening of teeth and, ultimately, to tooth loss.

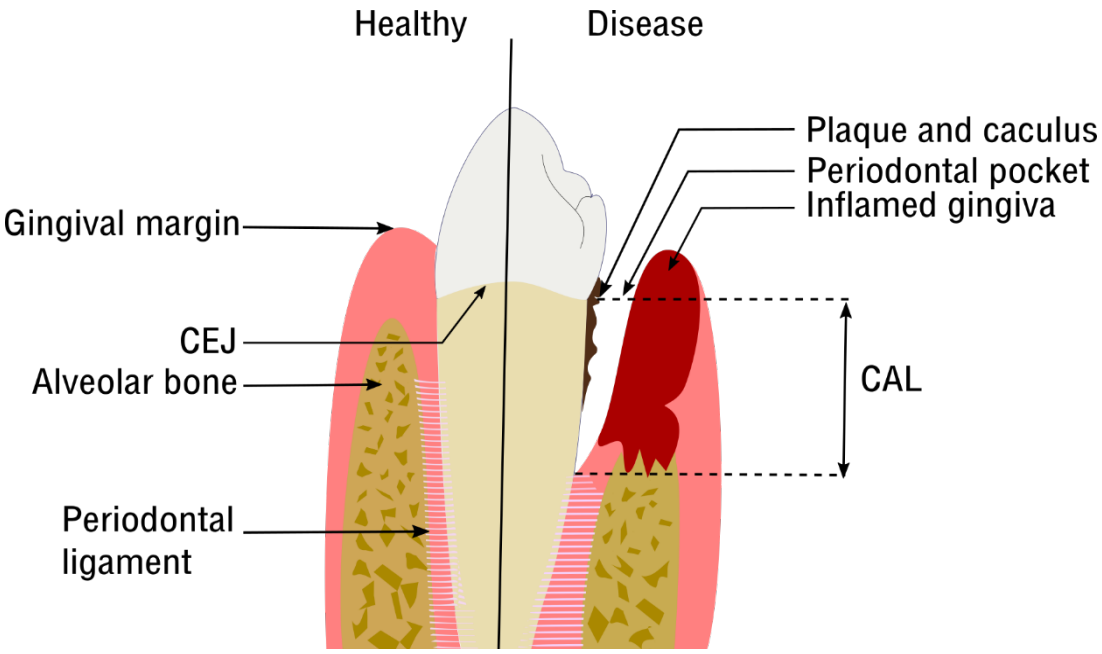


Figure 1.1 Diagram comparing a periodontally healthy site (left panel) with a periodontitis site (right panel) (adapted from Cheng et al. 2017) (Cheng et al., 2017).
The activities of subgingival plaque and the host defences lead to inflammation and tissue damage. Clinical attachment loss (CAL) is the distance from the cemento-enamel junction (CEJ) to the base of the periodontal pocket.

The most common form of periodontitis is chronic periodontitis, which is assessed as mild, moderate or severe depending on the extent of bleeding on probing (BOP), periodontal pocket formation, radiographic bone loss and clinical attachment loss (CAL). The prevalence of severe chronic periodontitis varies according to world regions, from 10% to 15% in adult populations based on World Health Organization (WHO) epidemiological data (Petersen and Ogawa, 2012). A systematic review revealed that between 1990 and 2010 the global prevalence of severe periodontitis increased gradually with age and reached the peak at approximately 40 years of age (Kassebaum et al., 2014). Aggressive periodontitis is a less common but severe form of the disease, characterised by rapid periodontal tissue destruction at a relatively

young age (under 25 years) in systematically healthy individuals who have a high genetic susceptibility (Armitage and Cullinan, 2010).

It has long been accepted that systemic disease has an influence on the severity of periodontal disease, but recent studies also indicate that periodontitis can affect the pathogenesis of major systemic diseases (Cullinan and Seymour, 2013). Associations have emerged between periodontitis and a growing list of systemic diseases or conditions including cardiovascular disease, diabetes mellitus and rheumatoid arthritis (RA) (Lundberg et al., 2010, Genco and Van Dyke, 2010, Lalla and Papapanou, 2011).

1.2 Microbiology of periodontal disease

1.2.1 Dental plaque and microbial communities

More than 700 bacterial species have been identified from the human mouth (Paster et al., 2006), but only 50-60% of these organisms can currently be cultured, possibly because they have evolved to live within a biofilm community rather than in monoculture (Wade, 2002). Oral microbial biofilms are three-dimensional structured bacterial communities attached to mucosal and dental surfaces and are embedded in an exo-polysaccharide matrix (Wood et al., 2000, Do et al., 2013). Living within a biofilm provides bacteria with significant advantages, i.e. protection from host defences and antimicrobial agents, broader habitat range, more efficient metabolism, and enhanced virulence (Marsh, 2005, Marsh et al., 2011).

Dental plaque is a structurally- and functionally-organized biofilm that develops on the surface of the tooth and tooth root. Dental plaque forms in an ordered way by

means of early, intermediate, and late colonizing species and has a diverse microbial composition that, in health, remains relatively stable over time (Marsh, 2006). The early species that colonize teeth are predominantly *Streptococci*, particularly *Streptococcus mitis* and *Streptococcus oralis*. Early colonizers can serve as additional binding sites for intermediate and late colonizers after establishing themselves on the tooth surface (Kolenbrander et al., 2010). The supra-gingival plaque community grows above the gingival-tooth margin (Kolenbrander et al., 2006) and subgingival plaque is derived from supra-gingival plaque that spreads down into the gingival sulcus (Kolenbrander et al., 2006, Aas et al., 2005). During the development of periodontitis, there is a transition from the predominantly Gram-positive facultative populations associated with health to plaque that is dominated by obligately anaerobic, proteolytic Gram-negative rods and spirochetes (Ellen and Galimanas, 2005, Marsh, 1994). Tissue damage and disease progression occur as a result of the combined activities of organisms within subgingival dental plaque and host responses to them (Dixon et al., 2004, Kirkwood et al., 2007). The generation of the deep periodontal pockets contributes to creating an anaerobic environment and the inflammatory processes in periodontitis also provide environmental and ecological stimuli (e.g. increased pH, lower redox potential, increased gingival crevicular fluid flow, increased availability of peptide nutrients and haemin sources) that drive bacterial successions within subgingival plaque and the emergence of populations associated with disease (Marsh, 2003). As described above, the “ecological plaque hypothesis” was proposed to explain the development of the periodontal disease (Marsh, 1994).

1.2.2 Identification of periodontal pathogens

Periodontitis is a complex polymicrobial condition and many organisms have been implicated in its aetiology. Colonization of the pathogens triggers a response by the host's innate immune system. During the development of the disease, host defense pathways that were originally meant to protect against the bacterial challenge are derailed into an uncontrolled catabolic process that leads to damage of the supporting tissues, tooth mobility and ultimately tooth loss.

In early studies, the microbial search for periodontal pathogens relied heavily on culture-based methods. Socransky *et al.* characterized periodontal microbial communities based originally on culture methods and subsequently extended by large scale DNA: DNA checkboard hybridization (Socransky et al., 1998). A group of red-complex bacteria (*Porphyromonas gingivalis, Treponema denticola* and *Tannerella forsythia*) were identified as associated with the severe form of periodontal disease. The red complex is, to some extent, dependent on earlier colonization of the pocket by a complex of somewhat less pathogenic organisms called the orange complex which includes *Fusobacterium nucleatum*, *Prevotella nigrescens* and *Prevotella intermedia* (Socransky and Haffajee, 2005).

The advent of non-culture-based strategies, such as polymerase chain reaction (PCR), Sanger sequencing, the more recent developments in next generation sequencing (NGS), as well as metagenomics, has changed the scenario. Kumar *et al*. found that, in addition to species in the red or orange complexes, six unculturable novel phylotypes and eight recognized species were strongly associated with disease (Kumar et al., 2003). Further studies of unculturable organisms have indicated that

members of the *TM7* phylum are associated with the early stages of disease (Brinig et al., 2003) and methanogenic bacteria with increasingly severe disease (Lepp et al., 2004). Several recent studies based on pyrosequencing of 16S ribosomal RNA (rRNA) gene amplicons provided a much broader picture of the overall diversity of the subgingival microbiota and revealed new species strongly associated with periodontitis (Abusleme et al., 2013a, Griffen et al., 2012, Park et al., 2015). Spirochetes have long been recognized as key players in periodontal disease and many important species from this phylum cannot be cultured (Ellen and Galimanas, 2005). In addition, it has been proposed that herpesviruses play a significant role in periodontal disease (Slots, 2005). A systematic review has suggested that there were 17 newly identified species/phylotypes associated with periodontitis and four of these microorganisms are not-yet-cultivable (Pérez-Chaparro et al., 2014). However, some periodontal pathogens including the red-complex bacteria can also be found in healthy individuals which indicated that their presence alone is not responsible for disease (Kilian et al., 2006, Socransky and Haffajee, 2005). It is very difficult to allocate pathogenic roles to individual periodontal organisms within the complex communities that are associated with disease. The changes in microbial community associated with the transition from healthy to disease status has attracted intense research interest and the stability of the dental-plaque community may act as a good predictor of periodontal health (Kumar et al., 2006). Metagenomics (also referred to as environmental and community genomics) is the genomic analysis of microorganisms by direct extraction and cloning of DNA from an assemblage of microorganisms. These techniques have facilitated the study of the physiology and ecology of environmental microorganisms. Currently, the research endeavour based on culture-independent methods is expanding beyond asking “Who is there?” to include the more difficult question “What are they doing?”. Szafrański *et al.* utilized

metatranscriptomics to identify the functional shift from health to periodontitis as well as the response of individual species to dysbiosis (Szafrański et al., 2015).Yost *et al.* compared metatranscriptomic profiles of subgingival plaque from active and inactive sites in patients with chronic periodontitis and found metabolic changes in the microbial community associated with the initial stages of dysbiosis (Yost et al., 2015). These studies are the starting point to explore microbial communities behaviours and will give insight into how environmental signals modify the behaviour of the community (Solbiati and Frias-Lopez, 2018).

1.2.3 Keystone pathogens

In contrast to predominant species that affect inflammation by their abundant presence, recently, a "keystone-pathogen hypothesis" has gained traction in which keystone pathogens, such as *P. gingivalis*, disproportionally influence the whole microbial community and lead to periodontitis. Studies in a murine model, suggest that even a low number of *P. gingivalis* can disrupt the complement system, impairing host defences, leading to overgrowth of oral commensal bacteria and compositional changes in the microbiota. These changes can result in complement-dependent inflammation and consequently, trigger the development of periodontitis (Hajishengallis et al., 2011). The theory of keystone-pathogen was further developed giving rise to a polymicrobial synergy and dysbiosis (PSD) model of periodontitis aetiology, which suggests that periodontitis is initiated by a synergistic and dysbiotic microbial community (Hajishengallis and Lamont, 2012). After the initiation of pathogenicity by colonization with keystone pathogens such as *P. gingivalis*, communication between a keystone pathogen and other accessory pathogens, such as *Streptococcus gordonii*, enhances community virulence and facilitates the development of pathogenicity. However, the full range of interactions between *P.*

gingivalis and other periodontal microbial community members is yet to be revealed and these notions have been derived mainly from experimental animal models of disease. Environmental factors, such as smoking and diet may also manipulate the homeostatic balance (Divaris et al., 2013, Stabholz et al., 2010).

1.2.4 Aggressive periodontitis and *Aggregatibacter actinomycetemcomitans*

In addition to chronic periodontitis, there is evidence to suggest that *A. actinomycetemcomitans* plays a prominent role in the initiation and development of aggressive periodontitis and may function as a keystone pathogen in localized aggressive periodontitis (Fine et al., 2013, Shaddox et al., 2012). *A. actinomycetemcomitans* is a gram-negative rod which produces a leukotoxin that has lethal effects on human leukocytes including monocytes, polymorphonuclear leukocytes and T cells, and thereby facilitates *A. actinomycetemcomitans* evasion of the host defence system (Herbert et al., 2016). This organism displays significant genetic diversity. A JP2 genotype of *A. actinomycetemcomitans,* which is defined by a 530-bp deletion in the promoter region of the leukotoxin operon, is highly leukotoxic (Brogan et al., 1994). The JP2 genotype of *A. actinomycetemcomitans* has a particularly strong association with disease in people of North and West African descent (Kilian et al., 2006). It has been identified that the individuals infected by JP2 genotype strains of *A. actinomycetemcomitans* have a significantly higher risk of developing aggressive periodontitis than individuals infected by strains of the non-JP2 genotype (Hoglund Aberg et al., 2014).

1.2.5 ***Porphyromonas gingivalis***

P. gingivalis, formerly named *Bacteroides gingivalis*, is a non-motile, asaccharolytic, Gram-negative, rod-shaped, obligately anaerobic bacterium which forms black-pigmented colonies on blood agar plates. The abundance of *P. gingivalis* has been shown to increase in sites with periodontitis while it is present at lower levels or is non-detectable in periodontally healthy sites (Schmidt et al., 2014). *P. gingivalis* has been demonstrated to constitute a higher proportion of the total microbiota in deep compared with shallow periodontal pockets (Ali et al., 1996).

P. gingivalis is known to produce a vast arsenal of virulence factors that could penetrate the gingivae and cause tissue destruction either directly, or indirectly by the induction of inflammation (Hajishengallis et al., 2012). Important virulence factors include LPS, capsular polysaccharide (CPS), fimbriae and gingipains.

1.2.5.1 Capsule

Capsule, also known as CPS or K-antigen, is a major virulence determinant of *P. gingivalis* because it is able to facilitate the evasion of host immune system activation, promote the invasion of the bacterium within host cells, and lift the virulence (Singh et al., 2011). Non-encapsulated strains have been shown to be less virulent, and mostly caused non-invasive, localized abscesses whereas encapsulated strains caused invasive, spreading infections in murine models (Laine and van Winkelhoff, 1998). Encapsulated *P. gingivalis* strains were shown to be able to modulate the host response to bacteria by decreasing the synthesis of cytokines interleukin-1 (IL-1), IL-6, and IL-8 by human fibroblasts, which enables *P. gingivalis* to limit any inflammatory response at stages of the infection when this

would be beneficial to its survival (Brunner et al., 2010). Based on the capacity of CPS to stimulate systemic immunoglobulin G (IgG) antibody responses, at least six distinct CPS serotypes have been described (K1-K6) (Laine et al., 1996). Differences in CPS serotypes stimulated differential capacities in chemokine production by murine macrophages (d'Empaire et al., 2006) and dendritic cells exposed to different *P. gingivalis* CPS serotypes elicited distinct T-cell responses (Vernal et al., 2014). A study based on an Indonesian population has shown that the K5 serotype of *P. gingivalis* within clinical isolates was detected with a higher prevalence than other serotypes while this distribution might vary with the study population (Van Winkelhoff et al., 1999).

1.2.5.2 Fimbriae

The fimbriae of *P. gingivalis* are thin, filamentous cell-surface protrusions involved in nearly all interactions between the bacterium and the host, as well as other bacteria (Hamada et al., 1998). The adhesive properties of fimbriae allow *P. gingivalis* to bind and invade host cells, which may subsequently help the bacterium escape the host immune surveillance (Zenobia and Hajishengallis, 2015, Amano, 2010). *P. gingivalis* fimbriae are also vital to the biofilm formation. They are implicated in the cohesive interaction (coaggregation) of *P. gingivalis* with other bacteria other plaque-forming bacteria, such as *Actinomyces viscosus*, *Treponema medium*, *T. denticola*, and *Streptococcus oralis* (Amano, 2007).

There are two types of *P. gingivalis* fimbriae which are encoded by *fimA* gene (major fimbriae) and *mfa1* gene (minor fimbriae) separately. Based on the amino terminal and the DNA sequences, *P. gingivalis* major fimbriae were further classified into six types: types I–V and Ib (Nakagawa et al., 2000, Nakagawa et al.,

2002, Amano et al., 1999). Studies have shown that *P. gingivalis* strains possessing type II were more predominant in periodontitis patients (Amano et al., 1999).

1.2.5.3 Lipopolysaccharide (LPS)

Like all Gram-negative bacterial species, *P. gingivalis* is sheathed by an outer membrane, an asymmetric lipid bilayer of which the outer leaflet is composed of LPS, which comprises an important component recognized by host cell receptors that then triggers intracellular signalling events. In general, bacterial LPS consists of a distal polysaccharide (or O-antigen), a non-repeating "core" oligosaccharide and a hydrophobic domain known as lipid A (or endotoxin) (Figure 1.2) (How et al., 2016). *P. gingivalis* LPS exhibits unique structural features compared with the LPS of other species, especially the lipid A structures (Dixon and Darveau, 2005). The heterogeneous lipid A structures in the LPS of *P. gingivalis* have distinct and opposing effects on toll-like receptors (TLR) playing a critical role in the early innate immune response to invading pathogens (Olsen and Singhrao, 2018). Unlike well-studied LPS of *Escherichia coli,* which is recognised by TLR4 receptor and then lead to innate host defence mediator production, the receptors of *P. gingivalis* LPS have been reported to be either TLR4 or TLR2 (Darveau et al., 2004). In addition, the LPS of *P. gingivalis* is also able to antagonize TLR4 activation (Triantafilou et al., 2007). Furthermore, the heterogeneity of lipid A has also been related to the micro-environmental concentration of haemin (Al-Qutub et al., 2006) and to an extent to the environmental temperature, which in turn are influenced by inflammation (Curtis et al., 2011). Although conflicting results have been reported regarding whether TLR4 or TLR2, or both can be activated by *P. gingivalis* LPS, (possibly due to the use of different forms of LPS and different experiment models) (Nativel et al., 2017), the heterogeneity of lipid A from LPS may facilitate *P.*

gingivalis adaption and survival in different host environments through immunomodulation. *P. gingivalis* LPS also exhibits inhibition on osteoblastic differentiation and mineralisation in periodontal ligament stem cells, which is important in periodontal tissue regeneration (Kato et al., 2014).

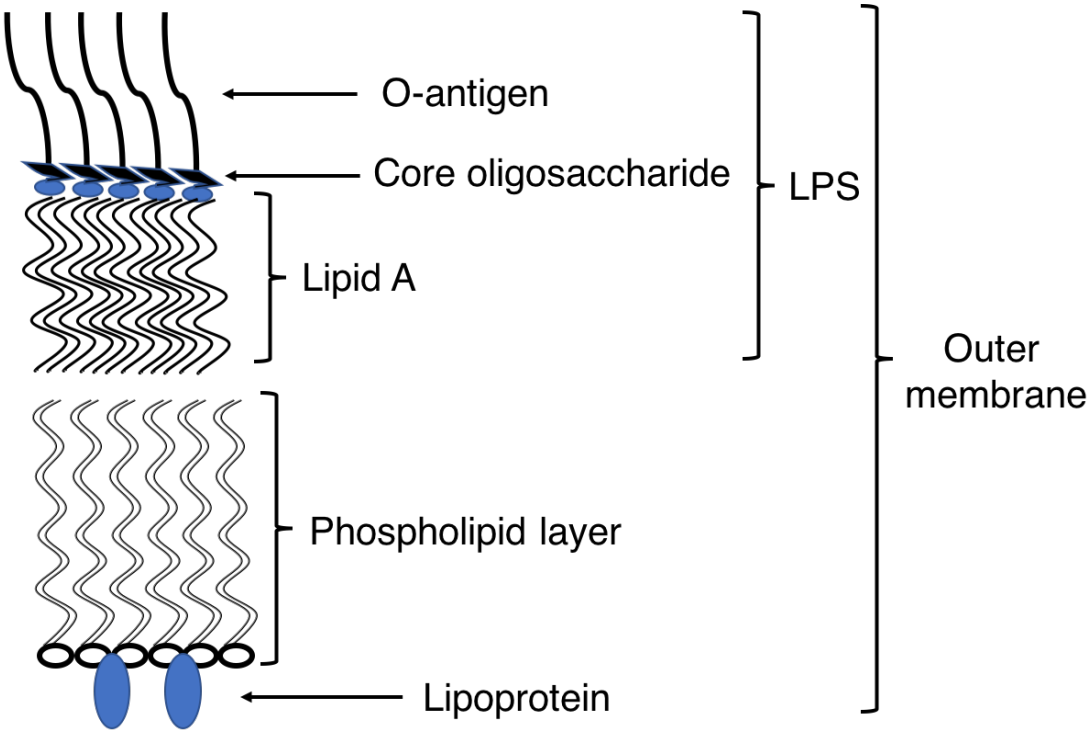


Figure 1.2 Schematic structure of lipopolysaccharide (LPS) of the outer membrane of *P. gingivalis* (adapted from How et al. 2016) (How et al., 2016)

1.2.5.4 Gingipains

Gingipains are a group of cysteine proteinases, also described as "trypsin-like" enzymes, which are major virulence factors of *P. gingivalis.* They account for 85% of the total proteolytic activity of *P. gingivalis* (Potempa et al., 1997). Based on substrate specificity, gingipains are divided into arginine-specific (Rgp) and lysine-specific (Kgp) gingipains, which cleave polypeptides at the C-terminus after an arginine or a lysine residues, respectively (Guo et al., 2010, Curtis et al., 2001). The Rgps are encoded by two homologous genes, *rgpA* and *rgpB*, and Kgp by *kgp*. The

translated products of *rgpA* and k*gp* both contain a catalytic and an adhesion domain while the adhesion domain is missing in the product of *rgpB* (Figure 1.3). Depending on *P. gingivalis* strains, gingipains are either predominantly attached to the bacterial surface or released into the medium in a soluble form. In strain HG66, non-glycosylated RgpB is released into the extracellular milieu in the soluble form; in all other strains, RgpB is glycosylated and remains bound to the cell-surface (Potempa et al., 1995). Several reports have indicated the presence of gingipains in outer membrane vesicles (OMV) which can be internalized into host cells and these OMV-associated gingipains may contribute to tissue destruction in periodontal diseases (Nakao et al., 2014). In general, gingipains play important roles in most phases of the pathogenesis of periodontal disease, from adherence and colonization through to nutrient acquisition and neutralization of host defences.

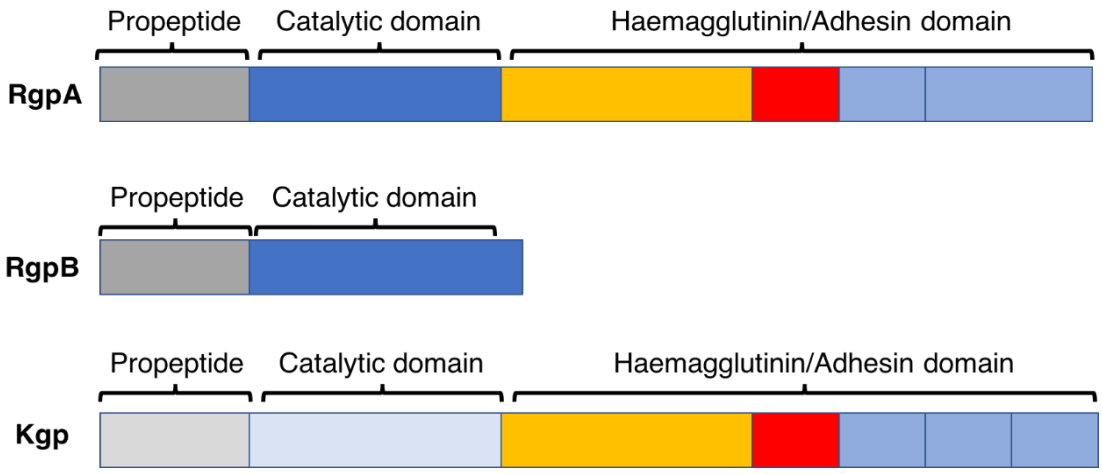


Figure 1.3 Schematic diagram of the gingipains domain structure (adapted from Li and Collyer 2011) (Li and Collyer, 2011).
The domains with high similarities are shown in the same colour.

In the initial phase of infection, gingipains mediate adherence of *P. gingivalis*, either directly or indirectly, to different sites within the oral cavity and facilitate colonization of the bacterial biofilm in the gingival crevice (Guo et al., 2010).

Gingipains themselves, are potent adhesins that can bind several extracellular matrix proteins such as fibrinogen, fibronectin, laminin (Pathirana et al., 2006). Co-aggregation among bacterial cells caused by the adherence of one bacterial species to another can be directly mediated by adhesin domains of RgpA and Kgp (Abe et al., 2004, Kamaguchi et al., 2003). *P. gingivalis* co-aggregation with selected oral bacteria is also mediated by the fimbrial adhesins and the Rgp is indispensable for maturation of fimbriae (Nakayama et al., 1996).

Gingipains are involved in both the destruction of periodontal tissues and interrupting host-defence mechanisms through the degradation of immunoglobulins and complement factors leading eventually to disease progression. Gingipain activity promotes *P. gingivalis* survival through the degradation of antibacterial peptides, such as neutrophil-derived α-defensins, complement factors, such as C3 and C4, T cell receptors, such as CD4 and CD8 (Hajishengallis et al., 2013). Gingipains are also suggested to contribute to the bleeding tendency at the diseased gingiva through degradation of fibrinogen and fibrin (Imamura, 2003). In human plasma, Kgp has the strongest effect on fibrinogen/fibrin, compared with the action of other types of gingipains (Imamura et al., 1995). Many experiments have also indicated that gingipains have seemingly contradicting actions on the innate immune responses, which can possibly be explained by the existence of a concentration gradient of gingipains in the tissue (Pathirana et al., 2010). In addition, *in vitro* experiments showed that Rgp gingipains cleave polypeptide chains at internal arginine residues, generating peptides with terminal arginines that are susceptible to citrullination by *P. gingivalis* peptidylarginine deiminase (to be discussed in detail in section 1.2.5.6) (McGraw et al., 1999, Wegner et al., 2010).

The major habitat of *P. gingivalis* is the subgingival plaque of the human oral cavity where sugar is scarce. *P. gingivalis* derives energy from the fermentation of amino acids for energy (Bostanci and Belibasakis, 2012). In *P. gingivalis*, nutritional extracellular proteins are initially degraded to oligopeptides by gingipains, these oligopeptides are then degraded by dipeptidyl peptidase (DPP), tripeptidyl peptidase, and acylpeptidyl oligopeptidase (AOP) generating di- and tri-peptides, the main incorporated forms in *P. gingivalis* (Nemoto and Ohara-Nemoto, 2016). The gingipain triple null (rgpA-, rgpB-, kgp-) mutant KDP136 was reported as unable to grow in defined medium with human albumin as the sole carbon source (Shi et al., 1999).

Like other anaerobes in the subgingival plaque, *P. gingivalis* also requires haem or haemin in its nutrient milieu for growth. Iron is utilized by *P. gingivalis* in the form of haem or haemin and has been shown to play a crucial role in its growth and virulence. By using chemostat cultures, Marsh *et al*. showed that *P. gingivalis* grown under conditions of haemin-excess, were always more virulent than when grown in haemin-limited conditions (Marsh et al., 1994).

Unlike other Gram-negative bacteria, *P. gingivalis* does not produce siderophores which are small chemical structures synthesized intracellularly to transport iron across cell membranes. Instead it utilizes specific outer membrane receptors, particularly gingipains, to acquire iron/haem (Olczak et al., 2005). Haemoglobin is the most abundant reservoir of haem in the periodontal pocket and inflamed gingival crevice (Hanioka et al., 2005). The bacterium is able to agglutinate erythrocytes by the adhesion domains of Kgp and RgpA; and lyse the erythrocytes to release

haemoglobin. Gingipains can bind haemoglobin with high affinity, which is mediated by the haemagglutinin-adhesin-2 or haemoglobin receptor (Nakayama et al., 1998). Oxyhaemoglobin, a form of haemoglobin released from erythrocytes, is then oxidized to methaemoglobin mainly by Rgps and subsequently hydrolysed by proteases (mainly by gingipains) to release haem. Finally, liberated haem from haemoglobin can be captured with high affinity by hemagglutinin-adhesin-2 (Paramaesvaran et al., 2003). In addition, gingipains can be utilized by *P. gingivalis* to degrade haptoglobin, transferrin, and hemopexin to get extracellular iron or released haem for growth *in vitro*.

1.2.5.5 Exopeptidases

P. gingivalis expresses various exopeptidases (DPP4, DPP5, DPP7, DPP11, prolyl tripeptidyl peptidase A (PtpA), and periplasmic AOP), which release di- and tri-peptides from most oligopeptide substrates. These peptides are then hydrolysed in the cytoplasm into single amino acids and used by *P. gingivalis* for carbon and energy metabolism (Takahashi and Sato, 2001, Takahashi et al., 2000). Studies have shown that the triple-knockout mutant for DPP4, DPP7, and PtpA showed dramatically reduced growth on media supplemented with albumin and IgG as the only carbon sources and the growth was reverted by addition of purified exopeptidases, demonstrating the key role provided by the peptidases (Oda et al., 2009).

All four DPPs and AOP activities have been detected within *P. gingivalis* cells, but not in culture medium (Nemoto and Ohara-Nemoto, 2016). DPP5 and DPP11 are localized in the periplasmic space of the cell (Ohara-Nemoto et al., 2014, Ohara-Nemoto et al., 2011). These exopeptidases have various substrate specificities which

benefit *P. gingivalis* in its need to obtain energy and carbon sources from the nutritionally limited subgingival environment (Nemoto and Ohara-Nemoto, 2016).

DPPs generally cleave oligopeptides without N-terminal modification and the penultimate P1-position residue from the N-terminus of the substrate is critical for the recognition by DPPs and the N-terminal P2-position residue additionally affects the activity (Ohara-Nemoto et al., 2011). The substrate specificity or preference of the DPPs that have been identified so far are listed in Table 1.1. The crystal structures and the amino acid residues, critical for hydrolysing activity and substrate specificity of the DPPs, have been investigated and this information provide a starting point for the development of DPP inhibitors. All those exopeptidases contribute to bacterial growth but there is no more information about the regulation of their gene expression. Although there were studies which reported that the substrate specificity of *P. gingivalis* DPP11 is primarily mediated by Arg^{673} , it is not known yet if *P. gingivalis* can regulate the enzyme activity by citrullination of those arginine residues.

Table 1.1 Summary of *P. gingivalis* DPPs.

Substrate specificity	Family	Crystal structure or determinant of substrate specificity	Ref.

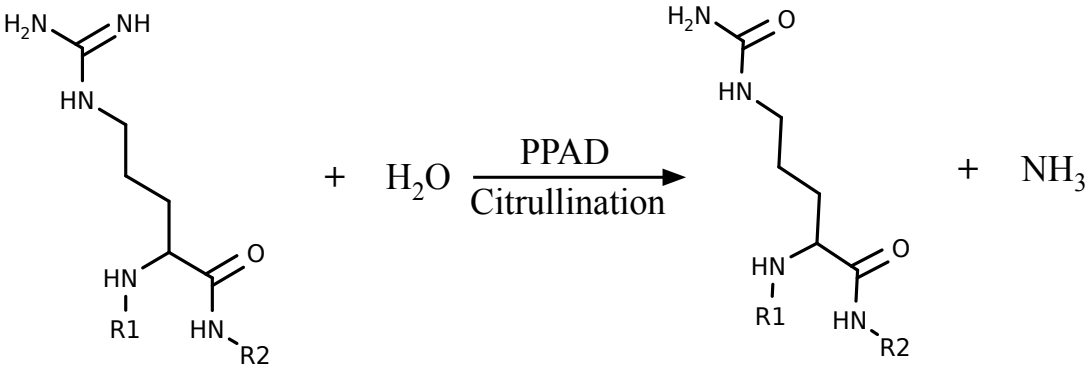
DPP4	Pro at the P1 position; biologically active peptides (substance P, fibrin inhibitory peptide, and ß-casomorphin)	S9	Ser^{593}, $His7^{00}$ and Asp^{668} make up the catalytic triad	(Rea et al., 2017, Banbula et al., 2000)
DPP5	Ala and hydrophobic residues at the P1 position	S9	-	(Ohara-Nemoto et al., 2014)
DPP7	Aliphatic or aromatic residue at the P1 and P2 positions	S46	Gly^{673} is critical but unique determinant of the substrate specificity	(Rouf et al., 2013)
DPP11	Acidic P1 residues (Asp> Glu)	S46	The acidic P1-position preference is primarily mediated by Arg^{673}	(Sakamoto et al., 2015)

1.2.5.6 *P. gingivalis* peptidylarginine deiminase (PPAD)

Recently, considerable interest has been focused on peptidylarginine deiminase (PAD) expressed by *P. gingivalis* (PPAD) which is able to modify proteins by deimination of peptidylarginine residues to produce peptidylcitrulline and ammonia (Figure 1.4). This posttranslational modification (PTM) leads to a reduction of positive charge, reduction in hydrogen-bonding ability and subsequently affects conformation and function of the protein (Vossenaar et al., 2003, Anzilotti et al., 2010). For a long time, PPAD was considered unique among prokaryotes, with *P. gingivalis* being the only bacterium known to produce and secrete such an enzyme. However, it has recently been shown that PPAD homologues was found in other

*Po*rphyromonas species (Gabarrini et al., 2018). The structure of PPAD is composed of four domains (Figure 1.5) (Montgomery et al., 2016) and the enzyme has both a secreted and a cell or OMV-associated forms (McGraw et al., 1999). PPAD is a substrate of type IX secretion system (T9SS) (Sato et al., 2013). During export N-terminal signal peptide (SP) directs the protein to the general secretion system and conserved C-terminal domain (CTD) will be recognized by T9SS. After translocation through the inner membrane CTD directs the protein for further translocation across the outer membrane through T9SS. Finally, CTD is cleaved off by PorU sortase and a secreted protein is modified by adding a A-LPS anchor allowing attachment to the cell surface (Lasica et al., 2017).

Unlike mammalian PADs, which act only upon arginine residues within the polypeptide chain in a calcium-dependent manner, PPAD functions in the absence of calcium and primarily citrullinates C-terminal residues and is able to modify free L-arginine (Bicker and Thompson, 2013, Abdullah et al., 2013). In addition, the proteolytic activity of the arginine gingipains secreted by *P. gingivalis* (section 1.2.5.4) were shown to be necessary for α-enolase citrullination. Rgp is able to cleave polypeptide chains at internal arginine residues, exposing carboxyl-terminal arginine residues, which are the preferential targets of PPAD (Wegner et al., 2010).



Peptidyl-arginine **Peptidyl-citrulline**

Figure 1.4 The process of citrullination by *P. gingivalis* peptidylarginine deiminase (PPAD).
PPAD converts peptidylarginine to peptidylcitrulline in a process called citrullination that also produces a free ammonia.

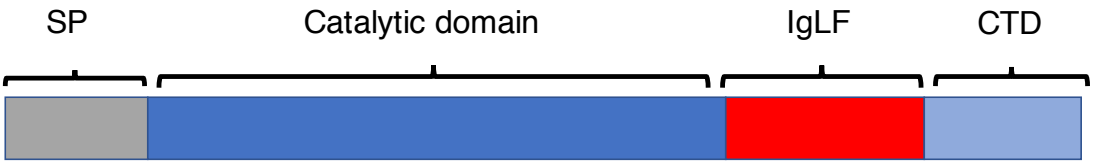


Figure 1.5 Schematic diagram of the PPAD domain structure (adapted from Montgomery *et al*., 2016).
The PPAD comprises four domains, from N- to C-terminal end: the signal peptide (SP), the catalytic domain, the Ig-like fold (IgLF), and the C-terminal domain (CTD).

PPAD is regarded as a virulence factor because citrullination by PPAD abrogates epidermal growth factor (EGF) which is important in periodontal repair and regeneration (Pyrc et al., 2013), interferes with complement activity (Bielecka et al., 2014) and contributes to infection of gingival fibroblasts and induction of prostaglandin E2 synthesis (Gawron et al., 2014). Additionally, a side effect of citrullination is ammonia production, which has a negative effect on neutrophil function and is protective for the bacteria during the acidic cleansing cycles of the mouth (McGraw et al., 1999, Abdullah et al., 2013).

P. gingivalis has attracted much interest of late as its PPAD enzyme has been reported to be able to citrullinate both bacterial and host proteins, thus providing a molecular mechanism for generating antigens that may drive the autoimmune response in RA (to be discussed in detail in section 1.5.1) (Montgomery et al., 2016, Wegner et al., 2010). A PPAD-deficient mutant of *P. gingivalis* W83 was created by replacement of the entire *ppad*-encoding DNA sequence with an antibiotic cassette and was used to assess the role of the enzyme in human and bacterial protein citrullination (Wegner et al., 2010, Bielecka et al., 2014, Stobernack et al., 2016). This mutant has also been utilized to investigate the contribution of PPAD to human gingival fibroblast infection, activation of prostaglandin E2, as well as development of collagen-induced arthritis in a mouse model (Gawron et al., 2014, Maresz et al., 2013). However, it is not known yet if the PPAD deficiency has any influence on the growth or gene expression profile of *P. gingivalis*.

1.3 Roles of neutrophils in periodontal diseases

Neutrophils act as a first protective barrier in periodontal diseases and are important regulators of both innate and adaptive immunity. Neutrophils account for 90% of the leucocytes in gingival crevicular fluid (GCF), and their concentration increases 15-fold in periodontally diseased sites (Pisano et al., 2005). Impaired neutrophil chemotaxis has been reported in periodontitis, and various strategies are employed by periodontal pathogens to disrupt neutrophil chemotaxis and/or function (Hajishengallis et al., 2015). Neutrophils generate neutrophil extracellular traps (NETs), which are web-like structures of DNA, histones, the contents of intracellular granules and antimicrobial peptides. Increased NET (Brinkmann et al., 2004) formation, or delayed NET clearance, may contribute to inflammatory responses as NETs provide an extracellular reservoir of inflammatory components,

such as LL-37, bacterial components, ds-DNA and hypercitrullinated proteins (White et al., 2015). In addition to their importance in periodontal diseases, neutrophils and periodontal bacteria have been implicated in mechanisms that increase the generation of autoantibodies that are important in the development of RA (to be discussed in detail in section 1.5.2).

1.4 Rheumatoid arthritis (RA)

RA is a systemic autoimmune disease characterized by chronic joint inflammation leading to destruction of bone and cartilage causing a reduction of functional capacity. RA affects 0.5%–1% of the overall population (Silman and Pearson, 2002) and the peak age of incidence is during the fifth decade of life (Tedeschi et al., 2013, Goemaere et al., 1990). RA disproportionately affects females compared with males, with a higher prevalence in women (Alpízar-Rodríguez et al., 2017, Goemaere et al., 1990). Moreover, the disease activity and progression of RA tend to be more severe in females compared with males (Sokka et al., 2009).

The aetiology of RA is multifactorial, complex and not fully understood. Known risk factors include certain genetic profiles (e.g. the presence of human leukocyte antigen [HLA]-DR), environmental factors (e.g. smoking) and the presence of autoantibodies (e.g. rheumatoid factor [RF] and anti-citrullinated protein antibodies [ACPA]).

1.4.1 HLA-DR

The most strongly associated genetic risk factor for RA is the presence of the HLA-DRB1 allele, which encodes common amino acid sequences (the shared epitope [SE]) in the third hypervariable region of the DRB1 molecule (Gregersen et al.,

1987). SE is found in numerous alleles and in this set of alleles the SE is a sequence of amino acids in the peptide binding groove of this type of major histocompatibility complex (MHC) Class II molecule (Raychaudhuri et al., 2012). Approximately up to 40% of risk for RA has been attributed to this genetic risk factor (Deane et al., 2017). A recent meta-analysis has confirmed the association of the SE with susceptibility in ACPA-positive RA patients while no robust associations were found in ACPA-negative RA patients (van der Woude et al., 2010). However, controversy exists regarding the possible protective effects of certain HLA–DRB1 alleles (van der Woude et al., 2010, Mattey et al., 2001).

1.4.2 Smoking

Smoking is the best characterized environmental risk factor for RA (Vessey et al., 1987). The increased risk was reported to occur after smoking for a long duration (equal or more than 20 years) and persisted for 10-19 years after cessation (Stolt et al., 2003, Svendsen et al., 2017). The association between smoking and RA was greatly enhanced in the presence of the SE and was dependent on the amount of smoking (Kallberg et al., 2011, Padyukov et al., 2004). In a recent Swedish population-based case-control study, smoking increased the risk of both ACPA-positive and ACPA-negative RA with a more pronounced influence on the risk of the former (Hedstrom et al., 2018).

1.4.3 Autoantibodies associated with RA

The lack of immunological tolerance in RA represents the first step toward the development of autoimmunity. Genetically susceptible individuals, under the influence of environmental factors, develop autoimmune phenomena that result in the presence of autoantibodies.

Protein citrullination is essential for many physiological processes (Gyorgy et al., 2006). However, citrullination may alter the three-dimensional architecture of the proteins and their solubility in aqueous solutions, and may lead to the generation of neo-epitopes, thus breaching immunological tolerance to citrullinated proteins. There are at least five isotypes of human PADs capable of citrullinating mammalian proteins (PAD1, 2, 3, 4 and 6), among which PAD2 and PAD4 are associated with the production of citrullinated proteins in RA (Foulquier et al., 2007). Neutrophils express several isoforms of PADs, and calcium-associated hyper-activation of neutrophil PADs can promote intra- and extracellular citrullination (Konig and Andrade, 2016). ACPA are detectable in approximately 70% of RA patients and are highly specific to this disease (Schellekens et al., 2000, Payet et al., 2014). In clinical practice, ACPA-positivity is defined by measuring antibodies against synthetic cyclic citrullinated peptide (CCP). Anti-CCP antibodies have been reported to be more specific markers for RA than RF, although both types of autoantibodies have been detected in the sera of asymptomatic individuals more than 10 years prior to disease onset (Nielen et al., 2004, Rantapää-Dahlqvist et al., 2003). Testing for both these types of antibodies has been included as a serologic criterion in the recently published 2010 RA classification criteria (Aletaha et al., 2010).

The information regarding ACPA-negative RA however is limited and other potent biomarkers need to be characterized for this manifestation of RA. Recently, a new protocol detecting autoantibodies against carbamylated proteins (anti-CarP) has been described (Shi et al., 2013) but has not yet been implemented for commercial use. Carbamylation is an enzyme-independent PTM in which cyanate binds to the primary amine of lysine, forming carbamyl groups, generating peptidyl-homocitrulline against which autoantibodies are subsequently induced (Trouw and

Mahler, 2012). Neutrophil myeloperoxidase (MPO) can enhance protein carbamylation by promoting the generation of cyanate from thiocyanate (Figure 1.6) (Wang et al., 2007). Similar to citrullination, carbamylation may result in changes to the functioning of proteins, e.g. carbamylation of IgG can inhibit the classical pathway of complement activation (Koro et al., 2014). It has been reported that anti-CarP autoantibodies were present in approximately 45% of RA patients, and importantly, detected in up to 30% of ACPA-negative RA patients (Shi et al., 2011). In a longitudinal study, presence of anti-CarP was shown to be able to predict the development of RA independently of anti-CCP antibodies (Shi et al., 2012).

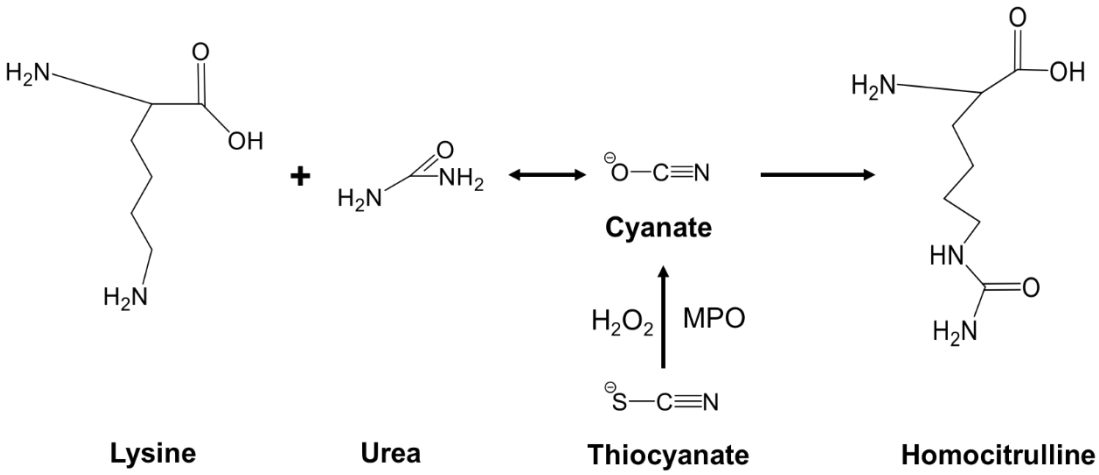


Figure 1.6 Illustration of carbamylation (adapted from Shi et al. 2011) (Shi et al., 2011).
Carbamylation is an enzyme-independent posttranslational modification (PTM) in which cyanate binds to the primary amine of lysine, forming carbamyl groups, generating peptidyl-homocitrulline. Urea is a source of cyanate in host and is in equilibrium with ammonium cyanate. During inflammation, neutrophil myeloperoxidase (MPO) can enhance protein carbamylation by promoting the generation of cyanate from thiocyanate

1.4.4 Individuals at risk of developing RA

RA-related autoantibodies such ACPA and markers of systemic or local subclinical inflammation (e.g. magnetic resonance imaging parameters) can be present months

or years before diagnosis of the disease (Nam et al., 2016, Deane et al., 2017). The development of RA is a multistep process. The European League Against Rheumatism (EULAR) study group differentiated the following phases: (1) presence of genetic and environmental risk factors for RA, (2) systemic autoimmunity associated with RA, (3) symptoms such as joint pain but without clinical arthritis (arthralgia), (4) unclassified arthritis and finally (5) RA (Gerlag et al., 2012).
It is thought that early treatment with disease-modifying anti-rheumatic drugs (DMARDs) and anti-inflammatory steroids can prevent progression of the disease and may even change or prevent the development of erosive joint damage (Heidari, 2011). The phase of arthralgia preceding clinical arthritis is the first opportunity to clinically recognize patients who are at risk for progression to RA and these high-risk individuals may be identified for preventive interventions (Hunt and Emery, 2014, Mankia and Emery, 2016). In contrast to the other phases that have been studied extensively, this phase is less well explored. Previous studies have shown that the risk for developing RA is even higher when the arthralgia is combined with ACPA positivity (Bos et al., 2010).

1.5 The relationship between periodontitis and RA

RA and periodontitis display some pathogenic similarities, such as the host immune response leading to soft tissue inflammation with subsequent hard tissue destruction, and certain shared risk factors, including smoking, the HLA-DRB1 allele and obesity (Chaffee and Weston, 2010, Marotte et al., 2006, Cheng et al., 2017). Periodontitis and RA are known to be significantly associated at the epidemiological level (Mikuls et al., 2016, Fuggle et al., 2016, Araujo et al., 2015), although the alternative conclusion has been drawn in some studies, possibly due to inconsistent diagnosis of periodontitis (Mikuls et al., 2016, Eriksson et al., 2016). A recent

systematic review has confirmed an increased risk of periodontitis in RA patients compared with systemically healthy controls (Fuggle et al., 2016).

The link between the two diseases was further highlighted in a recent study where patients with periodontitis and arthralgia who later developed RA had higher levels of disease activity and were more likely to receive methotrexate at RA diagnosis compared to patients without periodontitis (Hashimoto et al., 2015). Furthermore, a recent meta-analysis showed non-surgical periodontal treatment in patients with periodontitis and RA could lead to improvements in markers of disease activity in RA (Kaur et al., 2014). Understanding the common mechanisms that underlie periodontitis and RA could present new possibilities for the treatment and prevention of RA.

1.5.1 *P. gingivalis*, RA and autoantibody production

P. gingivalis, a keystone pathogen for periodontitis, has attracted a lot of attention from researchers exploring the link between these two diseases as it can express its unique PPAD, which is capable of citrullinating both bacterial and host protein/peptides (Wegner et al., 2010). PPAD activity has been detected in GCF from periodontitis patients and at lower levels in healthy controls (Laugisch et al., 2016). Human fibrinogen and α-enolase, two of the proteins targeted by ACPAs in RA (Wegner et al., 2010), are also substrates of PPAD and antibodies against auto-citrullinated *P. gingivalis* proteins cross react with citrullinated human α- enolase autoantibodies (Bright et al., 2015).

Recombinant PPAD has been reported to be capable of auto-citrullinating some of its 18 arginine residues (Quirke et al., 2014), although there is evidence that anti-

PPAD antibodies are not directed against the citrullinated form of PPAD and that N-terminal processing protects PPAD expressed by *P. gingivalis* from auto-citrullination (Konig et al., 2015).

The hypothesis that *P. gingivalis* is crucial in the aetiology of RA is supported by studies in animal models. Inoculation with *P. gingivalis* expressing PPAD accelerated progression and enhanced severity of collagen-induced arthritis in mice and was associated with higher levels of citrullinated proteins at diseased sites (Maresz et al., 2013).

To better understand the role for *P. gingivalis* in the development of RA, a limited number of studies have been performed on individuals at risk of RA. An increased level of anti-*P. gingivalis* antibodies were found in individuals at genetic risk of developing RA (Mikuls et al., 2012). Furthermore, higher anti-RgpB IgG levels, indicating the presence of this *P. gingivalis* virulence factor, were reported in the blood of both individuals at risk of and those with established RA, compared with healthy controls; while ACPA levels increased with time, anti-RgpB antibody levels did not and they decreased following diagnosis (Johansson et al., 2016). In contrast, no association between anti-RgpB and risk of RA was found in another study of a Southern European cohort (Fisher et al., 2015).

Although different approaches have been used to investigate the association between presence of *P. gingivalis* and RA, the results are equivocal (Cheng et al., 2017). In a recent study, periodontitis, but not the subgingival presence of *P. gingivalis*, was more prevalent in patients who later progressed to classifiable RA (Hashimoto et al., 2015). Similarly, De Smit *et al* concluded that, while there was evidence that periodontitis may precede symptomatic RA, there was insufficient evidence to

confirm a role specifically for *P. gingivalis* in disease progression (de Smit et al., 2015). Thus, while the link between periodontitis and RA is established, the specific roles of *P. gingivalis* or PPAD are less clear.

1.5.2 Other oral pathogens and multiple mechanisms underlying the link

Apart from *P. gingivalis*, another periodontal pathogen *A. actinomycetemcomitans*, has emerged as a candidate bacterial trigger of autoimmunity in RA. The leukotoxin-A (LtxA) produced by *A. actinomycetemcomitans* has been implicated in inducing leukotoxic hypercitrullination (protein citrullination spanning the broad range of molecular weights), and exposure to *A. actinomycetemcomitans* was associated with ACPA and RF (Konig et al., 2016). A recent study has reported that the periodontal pathogen *P. intermedia*, not but not *P. gingivalis*, was associated with antibody responses to a novel citrullinated peptide related to RA (Schwenzer et al., 2017).

Neutrophils are important in both RA and periodontitis pathogenesis. The production of RA-related autoantibody can be mediated by neutrophils via multiple routes. Immune dysregulation and tissue damage associated with periodontitis also in part attributed to the action and the function of neutrophils. Interference with the neutrophils is employed by many periodontal bacteria as an important pathogenic strategy, and some of these species may in turn promote neutrophil-mediated autoantibody production, e.g. via the pore-forming leukotoxin of *A. actinomycetemcomitans* (Konig and Andrade, 2016); *Filifactor alocis* promotion of neutrophil degranulation (Armstrong et al., 2016); and *P. gingivalis*, *A.*

actinomycetemcomitans and *F. nucleatum* triggering the release of neutrophil extracellular traps (White et al., 2016).

Periodontitis is a complex disease, mediated by a consortia of co-operating bacteria and the host responses to them. It is, therefore, logical to widen consideration of the influence of the microbiota beyond that of a single, albeit important, bacterium. With the help of NGS techniques, the changes of oral microbial composition have become a target for analysis from health to disease. Using 16S rRNA sequence analysis of the entire subgingival microbiome, Scher *et al.* found that the microbiome of RA patients was similar to healthy subjects with similar periodontal status, but, specific *Prevotella* and *Leptotrichia* operational taxonomic units (OTUs) were only found in new-onset RA patients, and *Anaeroglobus geminatus* was correlated with the presence of ACPA and RF, and with periodontitis (Scher et al., 2012). Another large-scale study using metagenomic shotgun sequencing identified compositional and functional alterations in RA-associated oral microbiomes, which were partly resolved by DMARD treatment; thus, this big data approach suggests that the oral microbiome composition could be important in the prognosis and diagnosis of RA (Zhang et al., 2015). However, the oral microbiome's contribution to the individuals at risk of development of RA has not been investigated yet.

1.6 Aims of this study

With the overall aim of better understanding the links between *P. gingivalis*, periodontal disease and risk of developing RA, as well as understanding the role of PPAD in *P. gingivalis* physiology, there were two main arms of the research project: *in vitro* studies and analysis of clinical samples.

1.6.1 The aim of the *in vitro* study

To investigate the role of PPAD in *P. gingivalis* physiology, beginning with testing the hypothesis that PPAD contributes to maintaining an environmental pH that favours *P. gingivalis* growth.

1.6.2 The aims of study of clinical samples

The main aim was to characterise the subgingival microbiome related to the development of RA or underlying the link between RA and periodontitis, which may lead in the long term to identifying potential markers for early diagnosis and treatment of RA. To do this, I used a shotgun metagenomic approach to compare the microbial composition and functional capability of subgingival biofilms in patients with and without periodontitis, with and without RA, and in individuals identified as anti-CCP positive and at risk of developing RA.

Chapter 2

Materials and Methods

2.1 Bacterial strains, storage and batch culture conditions

All the bacterial strains used in this study were listed in Table 2.1. The *P. gingivalis* W83 Δ*ppad* mutant (Wegner et al., 2010) was kindly gifted by Prof. Jan Potempa (University of Louisville). The *ΔrgpA+rgpB* mutant of *P. gingivalis* W50 (E8) (Aduse-Opoku et al., 2000) was kindly gifted by Prof. Michael A. Curtis (King's College London).

For long term storage, the colonial growth of a pure culture was innoculated in a Microbank cryovial (Pro-Lab Diagnostics, Canada) containing cryopreservative solution and beads. The bacteria were bound to the porous beads after inverting the vials 4-5 times. The excess cryopreservative was aspirated and the cryovials were kept at -80 °C for extended storage.

For batch culture, all the strains were grown on Columbia blood agar base (CBA) plate (Oxoid, UK) and in brain heart infusion (BHI) broth (Oxoid, UK) at 37°C in an A45 anaerobic work station (Don Whitley Scientific, UK) under 80% N_2, 10% H_2 and 10% CO_2 . To prepare CBA plates, Columbia agar base was autoclaved at 121°C for 15 minutes and 5% (v/v) defibrinated horse blood (Oxoid, UK) was added. All the media were pre-reduced by placing in the anaerobic workstation at least overnight prior to use. A bacterial strain was revived by streaking a bead from the cryovial on CBA plates and grown for 3 days. Bacterial strains were inoculated from CBA plates into BHI broth enriched with 5 mg/L haemin (Sigma-Aldrich, America) and 1 mg/L menadione (Sigma-Aldrich, America). The batch culture in BHI broth at different growth stages was used for further measurments.

Table 2.1 List of strains used in the study.

Strains	Source
Porphyromonas gingivalis W83	Gifted by Prof. Jan Potempa, type strain, parent strain of the Δ*ppad* mutant,
Porphyromonas gingivalis W83 Δ*ppad* mutant	Gifted by Prof. Jan Potempa (Wegner et al., 2010)
Porphyromonas gingivalis W50	Gifted by Prof. Michael A. Curtis, type strain parent strain of E8 mutant,
Porphyromonas gingivalis ΔrgpA+rgpB mutant (E8)	Gifted by Prof. Michael A. Curtis (Aduse-Opoku et al., 2000)
Prevotella corporis A818	Stored clinical isolates (Conrads et al., 1997)
Prevotella intermedia ATCC 25611	Stored type strain
Prevotella intermedia OMZ 326	Stored clinical isolate from subgingival plaque
Prevotella nigrescens OMZ 227	Stored clinical isolate from subgingival plaque
Prevotella melaninogenica NCTC 12963	Stored type strain

2.2 Continuous culture of *P. gingivalis* W83 and its Δ*ppad* mutant using a chemostat

2.2.1 Assembly of the chemostat

To determine the potential role of PPAD in pH regulation and its effect on the bacteria growth, *P. gingivalis* W83 and its Δ*ppad* mutant were grown in a chemostat (Applikon, the Netherlands) separately, under the same conditions. The chemostat vessel was attached to an Applikon ADI Biocontroller (Applikon, the Netherlands) which was able to monitor and record the temperature, pH and redox potential by probes. A motor controller was also mounted on the vessel to power a rotor submersed within the culure allowing agitation (Figure 2.1).

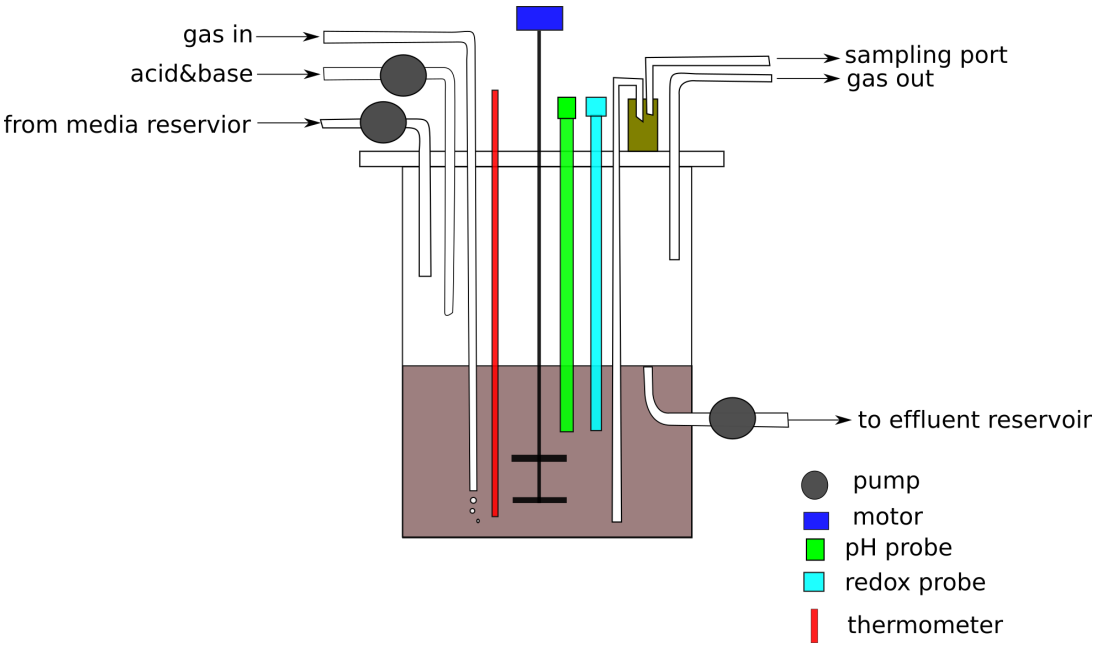


Figure 2.1 Diagram of chemostat.
A 2-liter-capacity chemostat was operated at a working volume of 700 mL. Fresh medium was continuously added and culture liquid was continuously removed at the same rate by pumps to get the dilution rate of 0.05/h. The pH of the culture was controlled by the automatic addition of 1 M NaOH and 0.5 M HCl, and the temperature was controlled at 37 ± 0.1°C. The culture vessel was sparged with a gas mixture of N_2 (95%, v/v) and CO_2 (5%, v/v) to maintain anaerobic conditions. A motor controller was mounted on the chemostat vessel allowing agitation. The temperature, pH and redox potential were monitored and recorded by corresponding probes.

2.2.2 Operation of the chemostat

The medium used was BHI broth supplemented with 5 mg/L of haemin (Sigma-Aldrich, America) to achieve excess haemin levels. The inoculum was prepared by growing the bacteria to late log phase in BHI broth, and 100 mL of this batch culture was used to inoculate the chemostat containing 300 mL of the same medium. A 10-L reservoir of medium feed was conrolled via a Watson Marlow 101U pump (Watson Marlow, UK) with a very slow rate (15 mL/h) initially; this rate was maintained overnight to reach the required working volume of 700 mL; once this volume was attained, the medium flow rate was increased to 35 mL/h. An outflow reservoir was also attached to the vessel via a Watson Marlow 120S/DM2 pump (Watson Marlow, UK) and its flow rate was adjusted to give the desired dilution rate of 0.05/h (dilution rate is calculated by dividing the flow rate by the culture volume).

The culture vessel was sparged with a filter-sterilized gas mixture of 95% (v/v) N_2 and 5% (v/v) CO_2 to maintain anaerobic conditions. The temperature was controlled at 37 ± 0.1°C and the culture was agitated at 40 rpm. The pH of the culture was maintained initially at 7.25 ± 0.05 by the automatic addition of 1 M NaOH and 0.5 M HCl and the chemostat was allowed to achieve a steady state, defined as at least three optical density at a wavelenghth of 600 nm (OD_{600}) measurements varying by around ±10% or less during at least two or three consecutive residence times (1/dilution rate). After three days at the first steady-state, the pH control was removed and the culture was left to reach the second steady-state. Redox potential, pH and temperature of the culture were measured by means of the BioXper software package (Applikon, the Netherlands).

The purity of the culture was checked daily by Gram staining and viable culture on CBA plates to confirm that there was no contamination from any other microoganism.

2.2.3 Sampling and analyses of the chemostat culture

Samples of the chemostat culture were taken daily from each steady state and the transition stage between the two steady states for analysis, by using a syringe at the end of the air filter connected to the sampling port. Viable counts were determined after 1 mL fresh culture was transported to an anaerobic cabinet quickly and 10-fold serial dilutions were prepared in BHI broth to a dilution of 10^{-10}. Aliquots (100 μL) of 10^{-7}, 10^{-8}, 10^{-9} and 10^{-10} diluted cultures were spread on separate CBA plates in triplicate. Colony forming units (CFU) on the plates were counted after incubation for 3-5 days. Only plates with 30-300 colonies were counted.

The enzyme activities of the chemostat cultures were assessed as described in sections 2.3.1 and **Error! Reference source not found.**. Samples were also stored to allow subsequent analysis of RNA. For this, 12 mL bacterial culture was mixed with 8 mL RNA bacterial protect reagent (Qiagen, Germany). Aliquots of 2 mL sample were centrifuged at 10000 g for 2 min to remove supernatant and then stored at -80 °C.

2.3 Measurement of bacterial enzyme activities

2.3.1 *P. gingivalis* peptidylarginine deiminase (PPAD)

PPAD enzymatic activity was measured using a colorimetric assay that detects formation of the ureido group of citrulline (Knipp and Vasak, 2000). PPAD activity buffer was prepared by dissolving 2.07 g 2-(Ncyclohexylamino)-ethanesulfonic acid

(CHES) (Sigma-Aldrich, America) and 0.308 g dithiothreitol (DTT) (Sigma-Aldrich, America) in 200 mL of deionized water, providing 50 mM CHES containing 10 mM DTT; the pH was adjusted to 9.5 using 1 M NaOH and HCl. N-a-benzoyl-L-arginine ethyl ester (BAEE) was used as the substrate which was dissolved in the buffer to get 6 mM substrate working solution.

Early stationary phase cultures of *P. gingivalis* W83 and W50 were diluted 1:10 in BHI broth and were incubated anaerobically overnight. Aliquots of the culture were collected at two-hourly intervals. Samples were also collected from the chemostat as described in section 2.2.3. Cell-free supernantants were obtained from 1 mL cultures by centrifugation at 10000 g for 2 min; the cellular pellets were re-suspended in 1 mL PPAD activity buffer after washing twice using the same buffer. 50 µL cells suspension in buffer or 50 µL cell-free supernatant was mixed with 50 µL substrate working solution in a 0.5 mL micro-centrifuge tube, and incubated for 30 min at 37 °C in a heating block (VWR® Analog Dry Block Heaters; VWR International, America). A 200 µL portion of freshly prepared colour developing reagent, containing 1 volume of Solution A and 3 volumes Solution B, was used to quench the reaction. Solution A contained 80 mM diacetyl monoxime (DAMO) (Merck, Austria) and 2 mM thiosemicarbazide (TSC) (Sigma-Aldrich, America). It was prepared as follows: 1.62 g DAMO and 36 mg TSC were dissolved in 200 mL of deionized H_2O and stored in the dark at 4°C. Solution B containing 3 M H_3PO_4, 6M H_2SO_4, and 2 mM $NH_4Fe(SO_4)_2$ (Alfa, UK) was prepared as follows: 200 mL of 85% (w/v) H_3PO_4 was slowly added to 450 mL of deionized H_2O with gentle stirring. To this solution, 330 mL of 98% (w/v) H_2SO_4 was slowly added. 750 mg of $NH_4Fe(SO_4)_2 \cdot 12H_2O$ was dissolved in this mixture. Upon cooling to room

temperature, the solution was carefully adjusted with deionized H_2O to a final volume of 1 L.

The mixtures containing cells or cell-free supernatants were incubated at 95 °C for 15 min in the heating block, and cooled for 3 min on the bench top. They were centrifuged at 10000 g for 2 min and the absorbance at 540 nm of the supernatant was then measured (using a Varioskan plate reader (Thermo Fisher Scientific, America). A standard curve was created using 0, 10, 50, 100, 200 and 400 μM L-citrulline (Alfa, UK) in PPAD activity buffer. A control reaction containing bacteria but no substrate was included. Another control of substrate alone was also included. All activity measurements were performed in triplicate.

The enzymatic activity, V, was calculated using the following equation, where A_{cit} is the measured absorbance, A_0 is the absorbance of controls (substrate along was used as the blank for the activity in cells, bacterial supernatant was used as the blank for activity in supernatant), B is the slope of the standard curve and T is the time of enzymatic reaction.

$$V = \frac{(A_{cit} - A_0)}{B \times T}$$

2.3.2 Citrullination activity in different species

The potential citrullination activity was assessed in the cell pellets of a range of *Prevotella* species and *P. gingivalis*. Different substrates were used to investigate the substrate specificity of those enzymes. PPAD and human PADs have different calcium requirements but the calcium dependency is not clear for activity in

Prevotella species. Therefore, the dependence on calcium for the citrullination activity was also analysed by using reaction buffer with or without calcium.

P. gingivalis W83, W83 Δ*ppad* mutant, W50, E8, *P. corporis* A818, *P. intermedia* ATCC 25611, *P. nigrescens* ATCC25261, *P. nigrescens* OMZ 227 and *P. melaninogenica* NCTC 12963 were grown in BHI until the late log phase. The OD_{600} of the culture was adjusted to 1.0 using BHI broth. The cell pellet was collected from 1 mL culture by centrifugation at 10000 g for 2 min, washed twice using PPAD activity buffer, pH 7.5, and re-suspended in 1 mL same buffer. The colorimetric assay for enzyme activity was used as described above (section 2.3.1), but a variety of substrates were employed: 3mM synthetic arginine substrate BAEE (Sigma-Aldrich, America), 3mM synthetic peptides Arg-Gly-Glu, Met-Arg-Phe, Gly-Arg (Bachem, Switzerland) or 50 mg/mL bovine serum albumin (BSA). When using BAEE and BSA as substrate, the buffer with or without 10 mM $CaCl_2$ was included in the experiment to assess the effect of calcium on the enzyme activity.

The activity of Rgp was determined in the whole culture and cell-free supernatant of chemostat cultures using the substrates carrying chromogenic leaving group p-nitroanillide (pNA; Sigma-Aldrich, America). Gingipain assay buffer was prepared on the day of use, by adding 0.2 mL of 1 M L-Cysteine•HCl stock solution to 10 mL 200 mM Tris•HCl-20 mM $CaCl_2$. L-Cysteine•HCl solution was neutralized before being added to the buffer by mixing 9 volumes of 1 M cysteine solution with 1 volume of 8 M NaOH.

Samples from the chemostat were also collected as described in section 2.2.3. The supernatant was obtained by centrifugation 0.5 mL culture at 10000 g for 2 min. 50

µL sample was added to 100 µL buffer in each well of a 96-well plate, and was mixed well by pipetting up and down a few times. The volume in each well was adjusted to 190 µL by adding deionized water. The plate was then incubated for 10 min at 37 °C allowing warming of the samples and reduction of the cysteine residues.

10 µL 10 mM Bz-L-arginine-pNA (L-BAPNA; Sigma-Aldrich, America) in dimethyl sulfoxide (DMSO) was added to the reaction at a final concentration of 0.5 mM and mixed thoroughly. Activity was measured in 0.1 M Tris-HCl-10 mM L-cysteine-10 mM $CaCl_2$, pH 8.0, containing 0.5 mM L-BAPNA at 30°C using a Varioskan plate reader supplied with an incubation function. The reaction was monitored by the increase in the OD_{405} nm over 30 min. A standard curve was prepared using pNA concentrations of 0, 30, 60, 125, 250 and 500 µM. A control reaction containing bacteria but no substrate was included. Another control of substrate alone was also included. All activity measurements were performed in triplicate.

2.3.3 Dipeptidyl-peptidase (DPP) activity

P. gingivalis W83 and the W83 Δ*ppad* mutant were grown in BHI until they reached the late log phase. The OD_{600} of the sample was adjusted to 1.0 using BHI broth. Cell pellets were collected from each 1 mL sample by centrifugation at 10000 g for 2 min. They were washed twice in ice-cold phosphate buffered saline (PBS), pH 7.4, and re-suspended in 1 mL same PBS.

Peptidase activity was measured as previously reported (Nishimata et al., 2014). Gly-Pro-, Lys-Ala-, Met-Leu-, and Leu-Asp-α-(4-methylcoumaryl-7-amide) (MCA)

were used as substrates (PeptaNova, Germany). 50 µL cell suspension was added to 199 µL reaction buffer composed of 50 mM sodium phosphate, 5 mM EDTA, pH 7.0. Then 1 µL of substrate working solution (5 mM) was added, providing a final substrate concentration of 20 µM. A reaction mixture containing bacteria but no substrate (50 µL cell suspension + 200 µL buffer) was included. Another control for substrate alone (1 µL substrate + 249 µL buffer) was also performed. After incubation for 30 minutes at 37°C, fluorescence intensity was measured with excitation at 380 nm and emission at 460 nm. A standard curve was created using 0, 1.25, 2.5, 5, 10 and 20 µM 4-Methyl-Coumaryl-7-Amide (AMC; PeptaNova, Germany) in reaction buffer. For each bacterial strain, three biological repeats were performed. DPP activities toward Gly-Pro-, Lys-Ala-, Met-Leu-, and Leu-Asp-MCA were defined as the rate of hydrolysis (µM/min).

2.4 Analysis of gene expression

2.4.1 *P. gingivalis* DPP 5 and DPP 11

The expression of the genes for DPP5 and DPP11by *P. gingivalis* W83 and its Δ*ppad* mutant was assessed using quantitative real-time polymerase chain reaction PCR (qRT-PCR).

2.4.1.1 RNA extraction

P. gingivalis W83 and its Δ*ppad* mutant were grown in BHI until they reached the late log phase. RNA was extracted from 1 mL culture of each strain using TRIzol Max Bacterial RNA Isolation Kit (Thermo Fisher Scientific, America) according to

the manufacturers' instructions and the quantity of RNA was estimated by using a Nanodrop 2000 Spectrophotometer (Thermo Fisher Scientific, America).

2.4.1.2 DNase treatment

549.13 ± 236.51 ng/μL RNA was obtained from batch culture sample. Two μg RNA were mixed with 3 μL DNase I (New England Biolabs, America) and 2 μL Alu I (New England Biolabs, America) in 2 μL 10•DNase I reaction buffer (New England Biolabs, America). Additonal diethyl pyrocarbonate (DEPC) treated water was added to make the final reaction volume 20 μL and the mixture was incubated at 37°C for 2 h to digest and eliminate contaminating genomic DNA. The reaction was inactivated by incubation at 80 °C for 20 min.

2.4.1.3 Precipitating RNA

Twenty μL of the RNA sample after DNase treatment was mixed with 2 μL sodium acetate (Alfa, UK) and 55 μL cold 95% ethanol. The mixture was incubated at -20 °C for 20 min. Then the supernatant was removed after centrifuging at 800g for 30 min. The pellet was washed using cold 70% (v/v) ethanol and was allowed to air dry after centrifuging at 800 g for 10 min. Finally, the RNA was dissolved in 10 μL nuclease-free water.

2.4.1.4 Assessment of genomic DNA contamination

PCR was performed to determine if there was any detectable genomic DNA in the RNA samples. The primers of 16S ribosomal RNA (rRNA) were used : PgF (5'-TGGTTTCATGCAGCTTCTTT-3') and PgR (5'-TCGGCACCTTCGTAATTCTT-3'). Each reaction mixture contained 1 μL of template mixed with 10 μL of Dream

Taq Green PCR MasterMix (Thermo Fisher Scientific, America), 2 μL of each primer (final concentration of each was 0.1 μM) and 5 μL of PCR water. The PCR was performed as follows: 95°C 10 min; 95°C 30 sec; 53°C 30 sec; 72°C, 35 sec; 35 cycles, and a final extension step of 10 min at 72°C. Reactions were also performed using water and DNA of *P. gingivalis* as the template, providing negative and positive controls, respectively. A 10 μL aliquot of the PCR product was run on a 3% (w/v) agarose-SB gel (protocol for preparing SB buffer: see Appendix A) at 200V for 5 min.

2.4.1.5 cDNA synthesis and qRT-PCR

One μg purified total RNA was mixed with Randon Hexamer primer (final concentration is 60 μM) for cDNA synthesis by using Transcriptor First Strand cDNA Synthesis Kit (Roche, Switzerland). The concentration of cDNA was quantified by using a Nanodrop 2000 Spectrophotometer. Custom Taqman primers and probes were ordered for *P. gingivalis* DPP 5, DPP 11 and 16S rRNA (Thermo Fisher Scientific, America) (Table 2.2). The cDNA was diluted in nuclease-free water over a 100-fold range (0.01~1). A 1 μL sample of dilutied cDNA sample was mixed with 10 μL TaqMan PCR Master Mix and 1 μL primers with TaqMan probe (Thermo Fisher Scientific, America). The mixtures were prepared in a 96-well reaction plate with additional nuclease-free water to provide a final volume of 20 μL for each cDNA concentration in each well. The reaction was run on a LightCycler 480 II (Roche, Switzerland) as follows: incubation at 50 °C for 2 min, ezyme activation at 95 °C for 10 min and then 40 cycles of 95 °C for 15 sec and 60 °C for 1 min. The average threshold cycle (Ct) value was calculated for the target gene (DPP 5, DPP 11) and reference gene (16S rRNA).

Table 2.2 TaqMan primers and probes.

Target	Sequence of probe (reporter)	Sequence of primer
P. g 16S rRNA	CCTGCTTCGCTCCCC	F: GTATCAAACAGGATTAGATACCCTGGTAG
		R: TGACGGTATATCGCAAACTCCTAGT
P. g DPP 5	TCAGCGTAATCTCCC	F: TGTAAGGAGGCAGAGACCAATCT
		R: CTGTCCCTGCGTGATCTGA
P. g DPP 11	TCCACGCAGTAAACG	F: GTCAAAACAAACAGGCGATGCA
		R: GTATCAAACAGGATTAGATACCCTGGTAG

P. g = *P. gingivalis*; F = forward; R = reverse.

2.4.1.6 Relative quantitation

To use the relative quantitation, the assumption that amplification efficiencies of the target and reference were approximately equal was checked (Livak and Schmittgen, 2001). The ΔCt (Ct_{target} - $Ct_{reference}$) was determined and a plot of the log cDNA dilution versus ΔCt was made. If the absolute value of the slope was close to zero, the efficiencies of the target and reference genes were similar. For both DPP 5 and

DPP 11, Ct value was normalized to that of 16S rRNA gene and the ΔCt were compared between *P. gingivalis* W83 and Δ*ppad* mutant.

2.4.2 RNA sequencing of *P. gingivalis* W83 and its Δ*ppad* mutant growing in chemostat

2.4.2.1 RNA extraction from stored samples

Frozen continous-culture samples of *P. gingivalis* W83 and its Δ*ppad* mutant were thawed on ice. Three biological repeats of two steady states, early and late transition stage were evaluated for each strain. RNA was extracted from each sample using the TRIzol Max Bacterial RNA Isolation Kit and the quantity of RNA was estimated by using Quant-iT™ RiboGreen™ RNA Reagent (Thermo Fisher Scientific, America). Genomic DNA was digested and checked as described above (section 2.4.1).

2.4.2.2 Library preparation

Ribosomal RNA was removed from 2 µg of total RNA using the Ribo-Zero rRNA Removal Kit (Gram-Negative Bacteria; Illumina, America) according to the manufacturer's instructions. High Sensitivity RNA Screen TapeThe concentration of ribosome-depleted RNA was quantified by using Quant-iT™ RiboGreen™ RNA Reagent (Thermo Fisher Scientific, America). Based on the RIN (RNA integrity number) value of the RNA samples (4.03 ± 0.72) , fragmentation, synthesis of first- and second- strand cDNA, the RNA-seq barcode ligation, and other steps of library construction were performed using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina (New England Biolabs, America) according to the manufacturer's protocol. The concentration of DNA libraries were quantified for each sample using

the Quant-iT™ PicoGreen ® dsDNA Reagent and Kits (Thermo Fisher Scientific, America). The size distribution of 4-fold diluted samples was evaluated on the Agilent 2200 TapeStation (Agilent, America) controlled by Agilent 2200 TapeStation Software, using the Agilent High sensitivity D1000 ScreenTape & Reagents. The samples were finally sequenced by using the NextSeq 500 sequencing system (Illumina, America), with single-end 75 bp reads.

2.4.2.3 Reads preprocessing and data analysis

Quality assessment was peformed on the raw data by using the FastQC (version 0.11.5). The adaptor at the 3' end was removed using Cutadapt (version 1.14) (Martin, 2011) and the reads were trimmed using Sickle (version 1.33) with cutoff of 28 and 15 for phred score and length. Reads after quality trimming were mapped to *P. gingivalis* W83 genome sequence by using Burrow-Wheeler Aligner (Li and Durbin, 2010). Reads in different features were counted by using htseq-count (Anders et al., 2015). Differential gene expression analysis was performed on DESeq2 R package (Love et al., 2014). Differentially expressed genes were annotated to the gene ontology (GO) terms by using R package topGO. A Fisher exact test implemented in the package was used to define the GO terms enriched in these up- and down regulated genes. The results of enrichment analysis were summarized and visualized by REVIGO (Supek et al., 2011). The scripts are attached in the Appendix B.

2.5 Metagenomic study of subgingival plaques in relation to RA

2.5.1 Ethical approval

Ethics approval from the local National Research Ethics Service (NRES) committee was sought and gained by the Leeds Musculoskeletal Biomedical Research centre (LMBRC), led by Prof. Paul Emery, application number 06/Q1205/169. R&D approval from the Leeds Teaching Hospital Trust (LTHT) was also gained. All participants provided informed written consent prior to study enrolment.

2.5.2 Study participants

Four groups of participants were recruited in this study in Leeds including: 1) 32 non-RA volunteers (healthy controls), 2) 48 anti-citrullinated protein antibodies (CCP) positive individuals with no clinical synovitis, 3) 26 new-onset RA patients (NORA) who are anti-CCP positive and within 3 months of commencing DMARD therapy, 4) 10 chronic RA patients who had at least 6 months DMARD therapy. The four groups of participants were age, gender and smoking status matched during the recruitment process. All participants underwent a periodontal examination. Diagnosis of gingivitis or periodontitis was given by the dentists based on their consensus decision.

2.5.3 Collection and processing of subgingival dental plaque samples

Sample collection sites were identified by research dental practitioners during the periodontal examination. Periodontally diseased sites were defined as sites with

pockets of 4 mm depth or more with bleeding on probing. Healthy sites were defined as sites of 3 mm depth or less with no bleeding on probing. Sites selected were those that were accessible and where saliva contamination could be controlled. Supragingival plaque was removed with cotton wool pledgets prior to sample collection. Two sterile paper points (Maillefer Pro Taper Paper Points F3; The Dental Directory, UK) were used to collect subgingival plaque from each site avoiding bleeding, and immediately placed in a labelled cryovial containing 0.5 mL of RNA protect Bacteria Reagent. Cryovials containing the paper points were vigorously vortexed and then centrifuged at 8000 g for 1 min. Supernatants were removed by pipetting and cryotubes containing pellets were stored at -80 °C.

2.5.4 DNA extraction from subgingival plaque

Subgingival plaque samples were thawed on ice from -80°C. For each patient, the plaque samples from diseased sites or healthy sites were pooled together, separately. DNA was extracted from pooled samples using the UltraClean® Microbial DNA Isolation Kit (Qiagen, Germany) as per manufacturer's instructions and quantified by using PicoGreen ® dsDNA Reagent and Kits (Thermo Fisher Scientific, America).

2.5.5 DNA library preparation and sequencing

The DNA was sheared to 200 bp in a small glass vial (microTUBE AFA Fiber Pre-Slit Snap-Cap 6x16mm) by using a S220 Focused-ultrasonicator (Covaris, UK). The size distribution of 4-fold diluted samples was evaluated on the Agilent 2200 TapeStation controlled by Agilent 2200 TapeStation Software A.01.05, using the Agilent High sensitivity D1000 ScreenTape & Reagents.

Depending on the concentration of sheared DNA in the samples, either NEBNext® Ultra™ DNA Library Prep Kit for Illumina® or NEBNext® Ultra™ II DNA Library Prep Kit for Illumina (New England Biolabs, America) was used for library construction including end preparation, adaptor ligation and PCR enrichment. AxyPrep™ Mag PCR Clean-up beads (Corning, America) were used for the clean-up steps during and after the library preparation to remove unincorporated adaptors, primers, adaptor dimers, primer dimers and other contamintants. The size distribution and the quantity of the DNA libraries were checked using the method described above. DNA libraries tagged with different index primers were pooled together and paired-end sequenced on the Illumina Hiseq 3000 machine (Illumina, America).

2.5.6 Metagenomic analysis using an in-house pipeline

An in-house pipeline was used to do the analysis locally where every step and parameter can be strictly controlled. Each sample's raw paired read data were processed with Cutadapt (version 1.14) to remove the adaptor at the 3'end and then quality trimmed using Sickle (version 1.33) in paired-end mode with the cut-off of 28 and 15 for phred score and length, respectively. Trimmed reads of each sample were *de novo* assembled using MEGAHIT (version 1.1.2) (Li et al., 2016). Assembled contigs were aligned against the NCBI non-redundant (NR) protein database (ftp.ncbi.nlm.nih.gov/blast/db/FASTA/nr. gz, November 2017) using DIAMOND (version 0.9.14) in sensitive mode with 95% identity and an e-value of 0.001 (Buchfink et al., 2015). The lowest common ancestor approach implemented in MEGAN6 (L.C.A., version CE.6.10.10) was used to assign aligned reads at different taxonomic levels with default parameters. For functional capability analysis, aligned reads were also assigned to clusters of orthologous genes (COGs)

using EggNOG functional identifier implemented in MEGAN6. The scripts used in the analysis are attached in appendix B.

2.5.7 Statistical analyses

Taxonomic and functional count data were exported from MEGAN6 in the biom format and imported as phyloseq objects (phyloseq, R). Microbial communities were characterized using alpha-diversity indices (i.e., number of observed species and Shannon diversity indices) and beta-diversity (Bray–Curtis dissimilarity) at species level. Alpha-diversity indices were calculated based on raw count data for each sample and compared between groups using the Kruskal-Wallis Test. The Dunn-Bonferroni correction of the *P* value was performed for multiple testing. Raw count data were converted to relative abundances using the number of annotated sequence reads as a denominator on a per-sample basis. β-diversity was determined by Bray-Curtis dissimilarity based on relative abundance after taxa filtering (prevalence equal or more than 3, total counts more than 10). The results were plotted using principal coordinates analysis (PCoA). Permutational multivariate analysis of variance (PERMANOVA) statistical test (*Adonis* function, vegan package, R) was used to test the significantly different β-diversity between groups. Homogeneity of multivariate dispersions was also tested by using *betadisper* function to ensure PERMANOVA assumptions (significant result is not due to differences in group dispersions). The permutation test (One-sided *signassoc* function, indicspecies R-package) was used to test statistical significance of taxa-group association based on relative abundance data. Sidak's correction was applied for multiple testing.

The core microbiome of each group was identified based on the threshold of 50% prevalence, 0.2% relative abundance by using microbiome R package. Co-occurrence network analyses were performed using an R script based on vegan, igraph and Hmisc packages as described previously (Ju et al., 2014). Co-occurrence of the species with the occurrence in at least 20% of samples in each group from periodontally healthy or diseased sites were investigated by Spearman's rank correlation and the *P* values were adjusted with a multiple testing correction using the Benjamini-Hochberg. Network visualization was conducted on the platform of Cytospace (Shannon et al., 2003). Spearman's coefficient (q) < -0.5 or > 0.5 and adjusted $P < 0.05$ were considered to be a strong and significant correlation between species. Topological features for each node were calculated for the networks. For functional capability analysis, raw count data was imported into DESeq2 and normalized by the estimated size factors. The significant difference of the functional capability between groups were tested by the Wald test in DESeq2.

2.5.8 Scan of PAD in subgingival plaque samples using the shotgun sequencing data

Each sample's raw paired read data were trimmed and *de novo* assembled as described in the section 2.5.6. Assembled contigs were aligned against the NCBI NR protein database (ftp.ncbi.nlm.nih.gov/blast/db/FASTA/nr. gz, November 2017) using DIAMOND (version 0.9.14) with 70% identity and an e-value of 0.001 to get a list of proteins accession number for each sample. The definition and annotation information of the identified proteins were retrived by searching agains the NCBI protein database with the proteins accession number. The protein information were recorded if the the names contain PAD on the NCBI database.

Chapter 3

Results

3.1 Growth of *P. gingivalis* W83 and its Δ*ppad* mutant in batch culture

P. gingivalis W83 and its Δ*ppad* mutant were inoculated separately into BHI broth containing 5 mg/L haemin and 1 mg/L menadione, and incubated in an anaerobic cabinet (as described in section 2.1). Samples of culture were withdrawn every two hours and the optical density was measured at 600 nm (Figure 3.1). The rates of growth were similar between the two strains in BHI broth. Culture samples were collected at late log phase for enzyme activity assays.

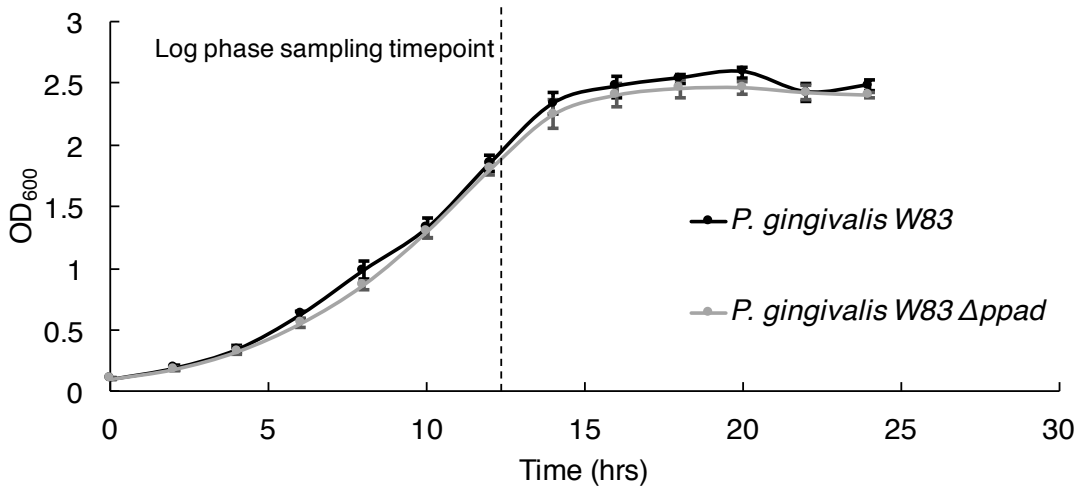


Figure 3.1 Growth curves of *P. gingivalis* W83 and its Δ*ppad* mutant in batch culture.
Each strain was grown anaerobically at 37°C in BHI broth containing 5 mg/L haemin and 1 mg/L menadione. The OD_{600} was recorded every 2 hours. Log phase sampling time point is indicated. Results are expressed as means ± standard deviations (n=3).

3.2 PPAD activity associated with cells and cell-free supernatant of *P. gingivalis* batch culture

To investigate whether PPAD activity in the strains used in this study is cell associated or secreted, or both, and to inform the forthcoming experimental design, PPAD activity was measured in both cells and supernatant which were obtained from the batch culture of *P. gingivalis* W50 or W83 every two hours after inoculation. The OD_{600} of the culture was also monitored for both strains. PPAD activity was detected in both cells and supernatant, but only the activity in the cells was increasing in line with the growth of the bacteria indicating that the majority of PPAD activity was associated with bacterial cells (Figure 3.2). Therefore, in forthcoming experiment the PPAD activity were mainly measured in the cells allowing me to identify any difference in the enzyme activity of different *P. gingivalis* strains.

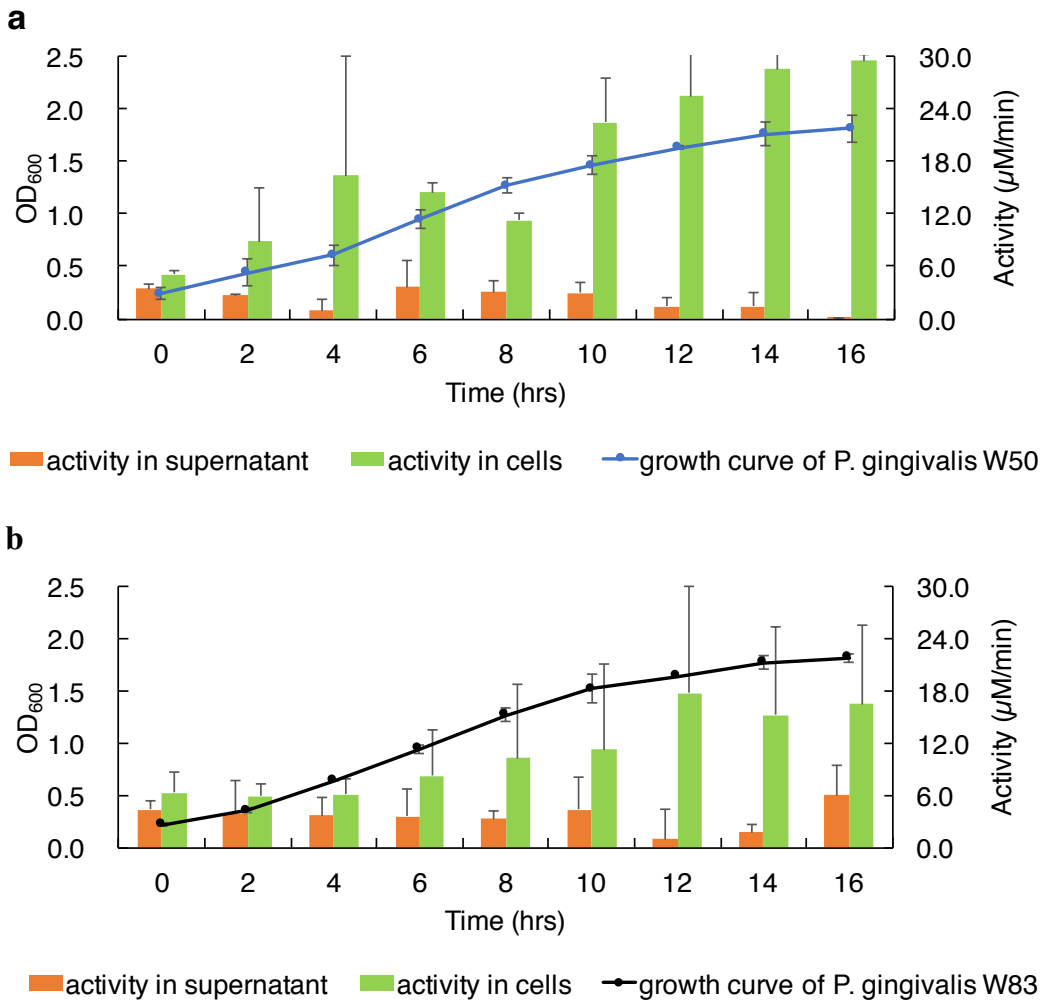


Figure 3.2 PPAD activity in the cells and cell-free supernatant from batch culture.
(a) *P. gingivalis* W50 and (b) W83 were grown in BHI broth anaerobically. Samples were withdrawn at the two-hourly intervals. The OD_{600} of the sample was measured and plotted agains the inoculation time (lines). Culture supernatant and the cellular suspension in the buffer were assayed for PPAD activity at the indicated time points (orange bar: PPAD activity in the supernatant; green bar: activity in the cells). Results are expressed as means ± standard deviations (n=3).

3.3 Stablity of PPAD activity

To check whether it would be necessary to use the fresh bacterial samples for the forthcoming PPAD activity assay, the stability of PPAD activity in the bacterial cells was analysed by comparing the activity between the fresh samples and those

stored at -80°C for two and three weeks. The results showed that there was no significant loss of activity when the bacterial cells were suspended in the PPAD enzyme activity assay buffer (as described in method section 2.3.1) and kept frozen for a maximum of three weeks (P=0.051, Kruskal-Wallis test) (Figure 3.3). Therefore, the cells suspension samples stored at -80°C within three weeks were used for PPAD activity assay in the subsequent experiments.

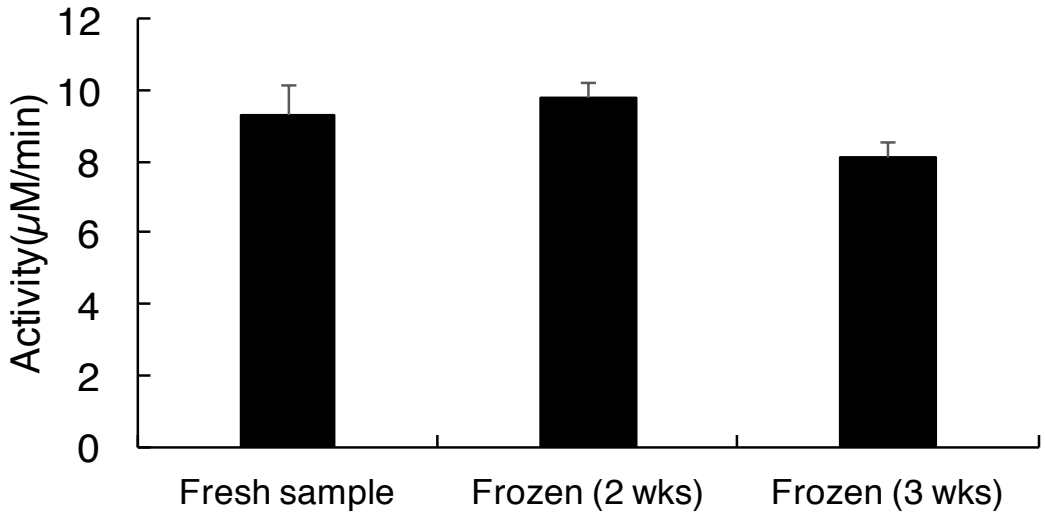


Figure 3.3 Comparison of PPAD activity between fresh and stored samples.
Bacterial cells of *P. gingivalis* W83 were grown in BHI, harvested and re-suspended in the PPAD enzyme activity assay buffer. PPAD activity was measured in the fresh cells suspension samples as well as the samples kept frozen (-80°C) for 2 weeks and 3 weeks. No significant difference was found in the PPAD activity of fresh and stored samples.

3.4 Continuous culture of *P. gingivalis* W83 and its Δ*ppad* mutant in the chemostat system.

To test the hypothesis that the PPAD can help maintain an alkaline local environmental pH which might benefit the growth of *P. gingivalis*, the *P. gingivalis* W83 wild-type strain and its Δ*ppad* mutant were each cultured in the chemostat

system, allowing the pH and other environmental conditions to be monitored and controlled. Three stages of the continuous culture were observed for each strain including two steady states, early and late transition stages situated between the two. The steady states of the continous culture in the chemostat were defined by monitoring the OD_{600} which varied by around ±10% or less during three consecutive residence time (the time it takes to entirely exchange the volume of the reactor).

3.4.1 First steady-state (pH controlled at 7.25 ± 0.05)

During the first steady-state culture the pH was externally controlled at 7.25 ± 0.05 by the automatic addition of acid or base as described in the method section 2.2.2 for both strains (Figure 3.4a). No significant difference was found when comparing the growth (OD_{600} and viable cell numbers) between the two strains ($P > 0.05$, t test) (Table 3.1). The redox potential of the culture was similar during growth of the two strains ($P = 0.20$, Mann-Whitney test) (Table 3.1) (Figure 3.4b).

3.4.2 Second steady-state (without pH control)

To test if *P. gingivalis* can regulate the local pH through PPAD, the external pH control was removed after 3 days of the first steady-state for both *P. gingivalis* W83 and its Δ*ppad* mutant. The pH of the culture decreased after removing the pH control for both strains (Figure 3.4a) but the pH of the mutant culture (6.90 ± 0.06) was higher than that of the wild-type strain (6.69 ± 0.03) ($P < 0.001$, Mann-Whitney test) during the second steady-state. Thus, the present data demonstrated that PPAD had an effect on the pH of the culture although not in line with my hypothesis, which was that PPAD enzyme activity helps *P. gingivalis* W83 maintain an alkaline environment. There must be other strategies *P. gingivalis* can utilize to adjust the environmental pH, on which PPAD may have the negative effect. The redox

potential of the mutant culture decreased in the second steady-state and was significantly lower than that of the culture of the wild-type strain in the second steady state (Table 3.1) (Figure 3.4b) ($P < 0.001$, t test). More interestingly, the OD_{600} of the cultures of both strains increased after removing the pH control ($P < 0.001$, t test) and both strains had an increase in cell numbers in the second steady-state compared with the first steady-state although only in the wild-type the difference was significant comparing the $\log_{10}$ values of the cell numbers between the two steady-states ($P < 0.001$, t test). In the second steady-state, the OD_{600} of the mutant was significantly higher than that of the wild-type strain ($P <0.001$, t test), indicating better growth of the bacteria. This result also supports the evidences that higher environmental pH and lower redox potential were found in the mutant culture compared with the wild-type strain, which may favour *P. gingivalis* growth. However, no significant difference was found comparing the cell numbers between the two strains ($P > 0.05$, t test) (Table 3.1).

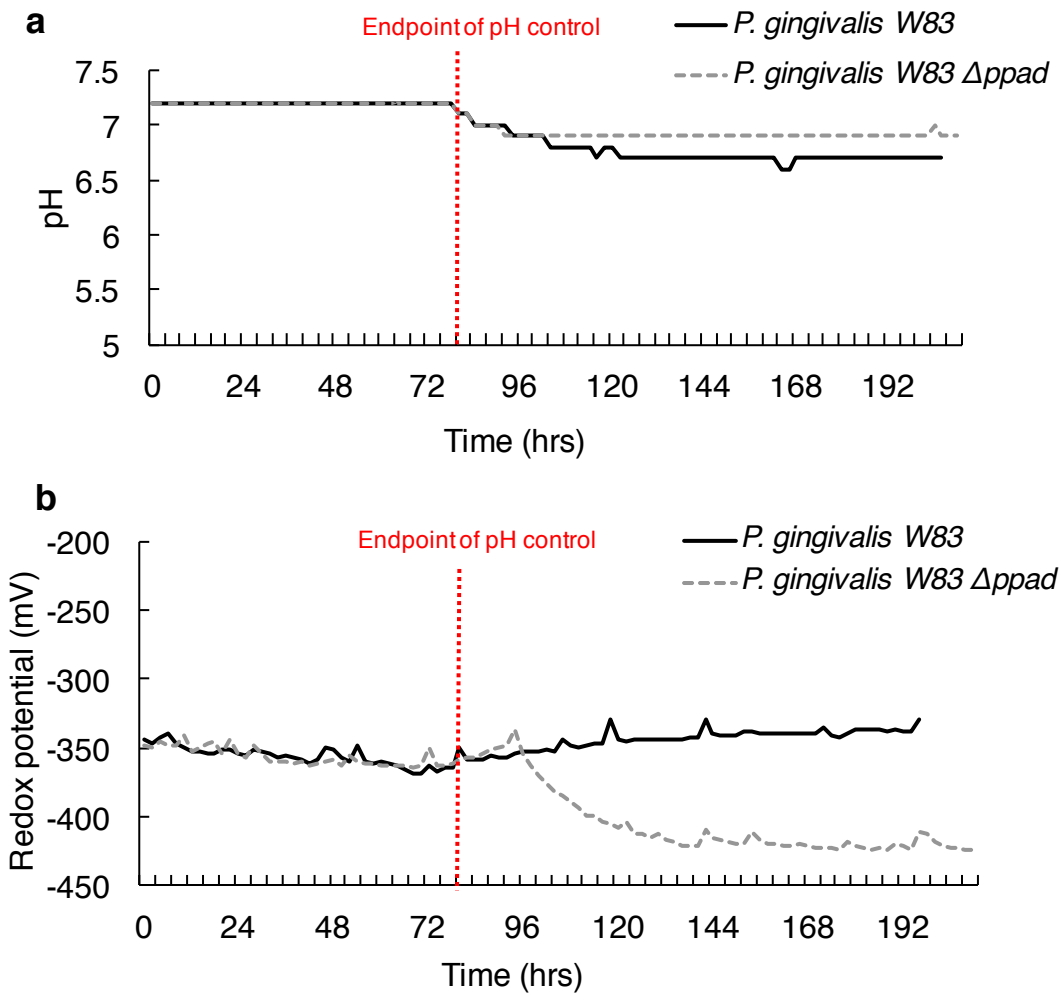


Figure 3.4 Environmental pH and redox potential of the continuous culture of *P. gingivalis* W83 and its Δ*ppad* mutant in the chemostat system. (a) The pH and (b) redox potential of the culture were monitored for *P. gingivalis* W83 (black line) and its Δ*ppad* mutant (grey dashed line). pH was controlled at 7.25 ± 0.05 by automatic addition of acid or base during the first 76 hours (red dash: endpoint of pH control).

Table 3.1 Evaluation of the growth and environmental conditions of the continuous culture of *P. gingivalis* W83 and its Δ*ppad* mutant during the first (pH controlled at 7.25 ± 0.05) and second steady-states (without pH control).

	***P. gingivalis* W83**		**Δ*ppad* mutant**	
Stage	**1st steady-state**	**2nd steady-state**	**1st steady-state**	**2nd steady-state**
Cell number (CFU/mL, log_{10} scale)	10.04 ± 0.22	$^{\#}$ 10.47 ± 0.11	9.89 ± 0.25	10.30 ± 0.17
OD_{600}	1.08 ± 0.03	$^{\#\#***}$ 1.15 ± 0.02	1.15 ± 0.08	$^{\#\#\# ***}$1.82 ± 0.03
pH	7.21 ± 0.01	$^{\#\#\# ***}$6.69 ± 0.03	7.24 ± 0.04	$^{\#\#\# ***}$ 6.90 ± 0.06
Redox potential (mV)	-354.08 ± 6.55	$^{\#\#\# ***}$-338.42 ± 3.05	-355.68± 6.80	$^{\#\#\# ***}$ -419.60 ± 3.70

$^{\#}$: significant difference ($P < 0.05$) between the first and second steady-states (t test), $^{\#\#}$: $P < 0.01$; $^{\#\#\#}$: $P < 0.001$.

***: significant difference between the two strains (t test), $P < 0.001$.

Results are expressed as means ± standard deviations.

3.5 PPAD activity

To further investigate the potential effect of PPAD on the environmental pH regulation, PPAD activity was measured in the cell suspensions of each strain within first and second steady-states. The activity was normalized to viable cell counts. No detectable PPAD activity was observed in the samples of Δ*ppad* mutant. PPAD activity of *P. gingivalis* W83 wild-type was significantly decreased in the second steady-state compared with the first steady-state ($P < 0.001$, t test) (Figure 3.5).

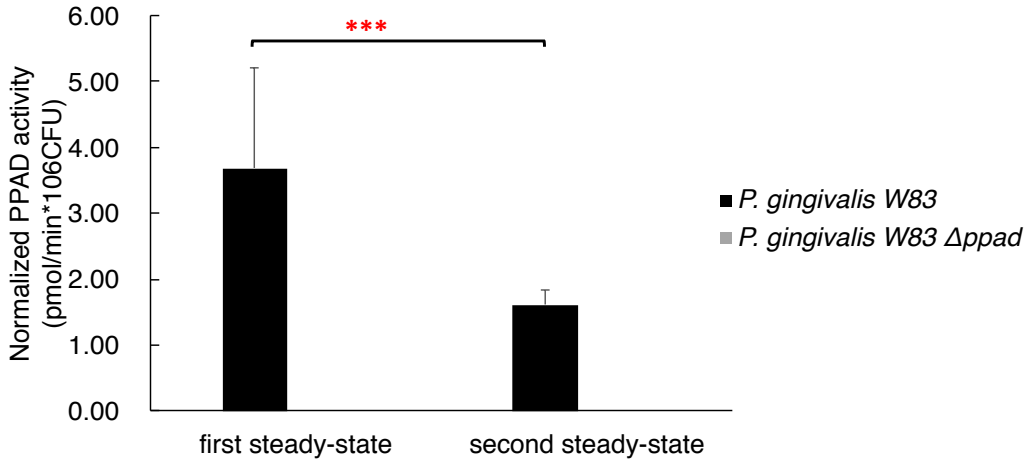


Figure 3.5 PPAD activity in the cells of *P. gingivalis* W83 and its Δ*ppad* mutant sampled from the first (pH controlled at 7.25 ± 0.05) and second steady-states (without pH control).
Stored cells suspension in the enzyme activity buffer were thawed and used to measure the PPAD activity. Results are normalized to viable cell counts and expressed as means ± standard deviations (n=3). No detectable PPAD activity was observed in the samples of Δ*ppad* mutant. PPAD activity of *P. gingivalis* W83 wild-type was significantly decreased in the second steady-state compared with the first steady-state. ***: $P < 0.001$ (t test).

3.6 Rgp activity of *P. gingivalis*

In order to understand the mechanism of how PPAD can influence the environmental pH and the growth of *P. gingivalis,* the activity of *P. gingivalis* Rgp, which is important for its energy metabolism and growth, was investigated using the continuous culture samples of *P. gingivalis* W83 and Δ*ppad* mutant. Whole culture samples from the first (pH controlled at 7.25 ± 0.05) and second steady-states (without pH control) were used to detect the Rgp activity with the chromogenic substrate L-BAPNA as described in method section 2.3.3. In order to compare the Rgp activities between the two strains, the activity in the whole culture sample was further divided by the cell number to normalize all values per 10^7 CFU (Figure 3.6) as the OD_{600} reflected very different viable cell counts in the present data. For both strains the normalized Rgp activity decreased in the second steady-state with lower environmental pH, compared with the first steady-state with a control of environmental pH. Within the second steady-state, normalized Rgp activity in the *P. gingivalis* W83 wild-type was significantly lower than that in the Δ*ppad* mutant. The data present here suggested that amount of Rgp may be reduced (or, less likely but possible, the Rgp produced is less active) due to the changes of environmental pH. Further experiments are required to confirm if the environmental pH influences the related gene expression causing the decrease of the Rgp activity.

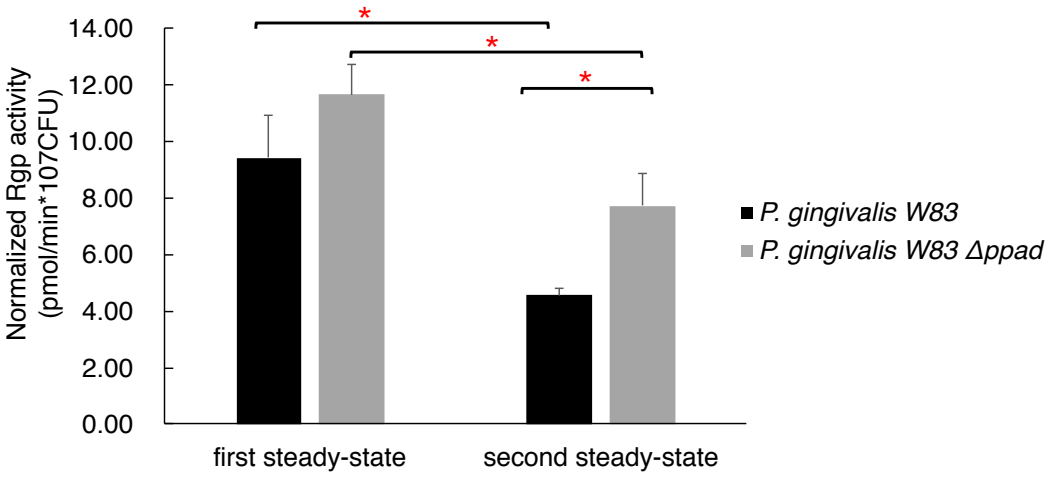


Figure 3.6 Rgp activity in the whole culture of *P. gingivalis* W83 and its Δ*ppad* mutant sampled from the first (pH controlled at 7.25 ± 0.05) and second steady-states (without pH control).
Rgp activity in the whole culture samples from the first steady-state (pH controlled at 7.25) and the second steady-state (without pH control) of the continuous culture was divided by the cell number of each strain to normalize all values per 10^7 CFU. Activity values are shown as the mean ± standard deviations (n=3). *: $P < 0.05$ (t test).

3.7 DPP activities in *P. gingivalis* W83 and its Δ*ppad* mutant from batch culture

DPP is another enzyme that is important for the metabolism of asaccharolytic *P. gingivalis*, the enzyme activity of which might be influenced by PPAD through citrullination. Due to the time constraints during the operation of chemostat culture system, the DPP activity could not be measured in the fresh samples from chemostat culture. To avoid any potential activity loss, instead of using stored chemostat culture samples, DPP activity was investigated in fresh batch cultures samples of *P. gingivalis* W83 and its Δ*ppad* mutant which were collected from late log phase. Cells suspensions in PBS were tested for DPP activities against Gly-Pro-, Lys-Ala-, Met-Leu-, and Leu-Asp-MCA, which are specific or preferential substrates of DPP 4, DPP 5, DPP 7, and DPP 11, respectively (method section 2.3.3). DPP activity to

Leu-Asp-MCA was significantly higher in the Δ*ppad* mutant compared with that in the wild-type ($P < 0.05$, t test) (Figure 3.7), suggesting that PPAD can impair the enzyme activity or gene expression of DPP 11.

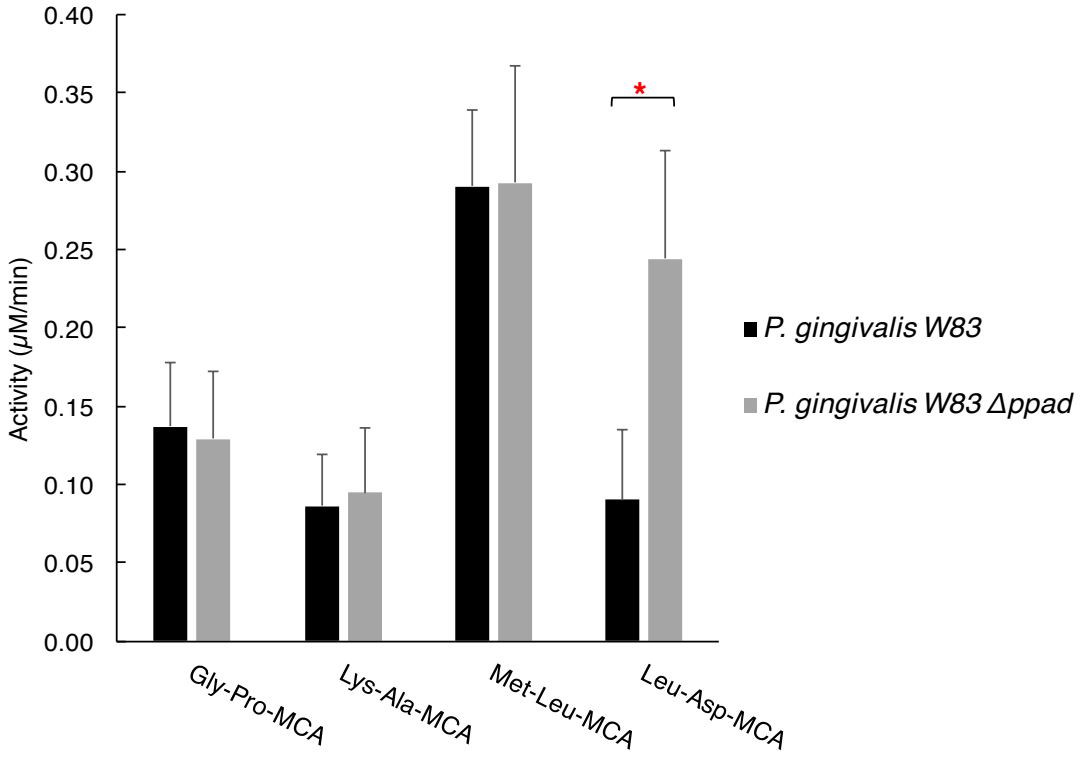


Figure 3.7 DPP activities of *P. gingivalis* W83 and its Δ*ppad* mutant grown in batch culture.
Bacterial cells were collected from late log phase of batch cultures and re-suspended in PBS. DPP activities were measured toward Gly-Pro-, Lys-Ala-, Met-Leu-, and Leu-Asp-MCA which are specific or preferential substrates of DPP 4, DPP 5, DPP 7, and DPP 11, respectively. The activity was quantified through release of MCA groups in a fluorometric assay. Activity values are shown as mean ± standard deviations (n=3). DPP activity against Leu-Asp-MCA, which is the preferential substrate for DPP 11, was significantly higher in the *P. gingivalis* W83 Δ*ppad* mutant compared with that in the wild-type strain (*: $P < 0.05$, t test).

3.8 Gene expression of *P. gingivalis* DPP 5 and DPP 11

To further investigate if PPAD had an effect on the gene expression of DPP, which may cause the difference in the DPP activity identified above, the gene expression

of DPP 5 and 11 were analysed by qRT-PCR as described in the method section 2.4.1. Samples were collected from the batch culture of *P. gingivalis* W83 and its Δ*ppad* mutant during late log phase. The gene expression of the target gene was normalized to the 16S rRNA gene. No significant difference was found in the gene expression of DPP 5 and 11 between the two strains ($P > 0.05$, Mann-Whitney test) (Figure 3.8), indicating that PPAD has no influence on the gene expression of DPP 5 and 11 but it may impair the enzyme activity of DPP 11 by citrullination.

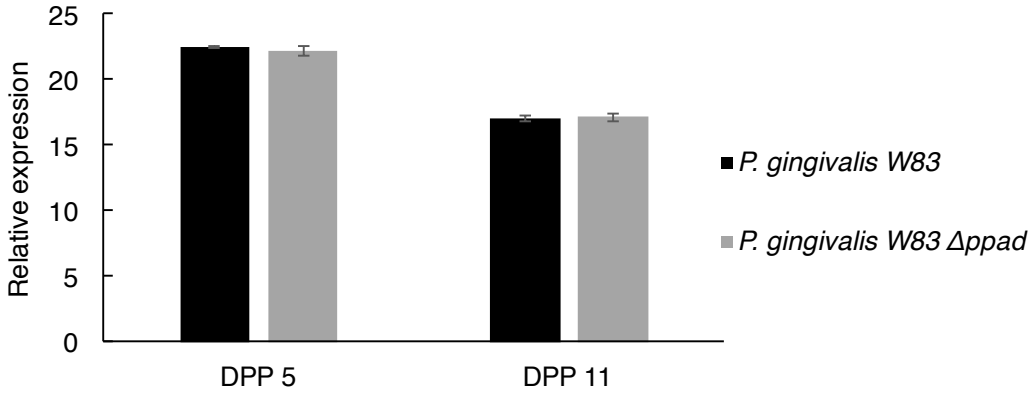


Figure 3.8 Gene expression of DPP 5 and DPP 11 in *P. gingivalis* W83 and its Δ*ppad* mutant.
The gene expression was normalized to 16S rRNA and results are expressed as means ± standard deviations (n=5). No significant difference was found in the gene expression of DPP 5 and 11 between the two strains (Mann-Whitney test).

3.9 RNA sequencing of *P. gingivalis* W83 and its Δ*ppad* mutant growing in the chemostat system.

To better understand the role of PPAD in the physiology of *P. gingivalis* and the interaction with the environmental pH, RNA was extracted and sequenced using the samples of both *P. gingivalis* W83 and its Δ*ppad* mutant from different stages with different pH conditions of the continuous culture.

3.9.1 Principal Component Analysis (PCA)

PCA was performed to identify the first (PC1) and second principal component (PC2), which can be used to explain the most and second most variations in the gene expression of all the samples. The PC1 results showed that samples of *P. gingivalis* W83 and its Δ*ppad* mutant were distinctly separated regardless of different stages of the continuous culture (Figure 3.9). The PC2 demonstrated that gene expression in both strains was related to the different continuous-culture stages (from first steady-state to the second steady-state). Moreover, the samples of wild-type strain were more clustered than those of the mutant according to the PC2, although in a similar pattern. This result indicated that to remove pH control had less influence on gene expression in *P. gingivalis* W83 wild-type strain than its Δ*ppad* mutant.

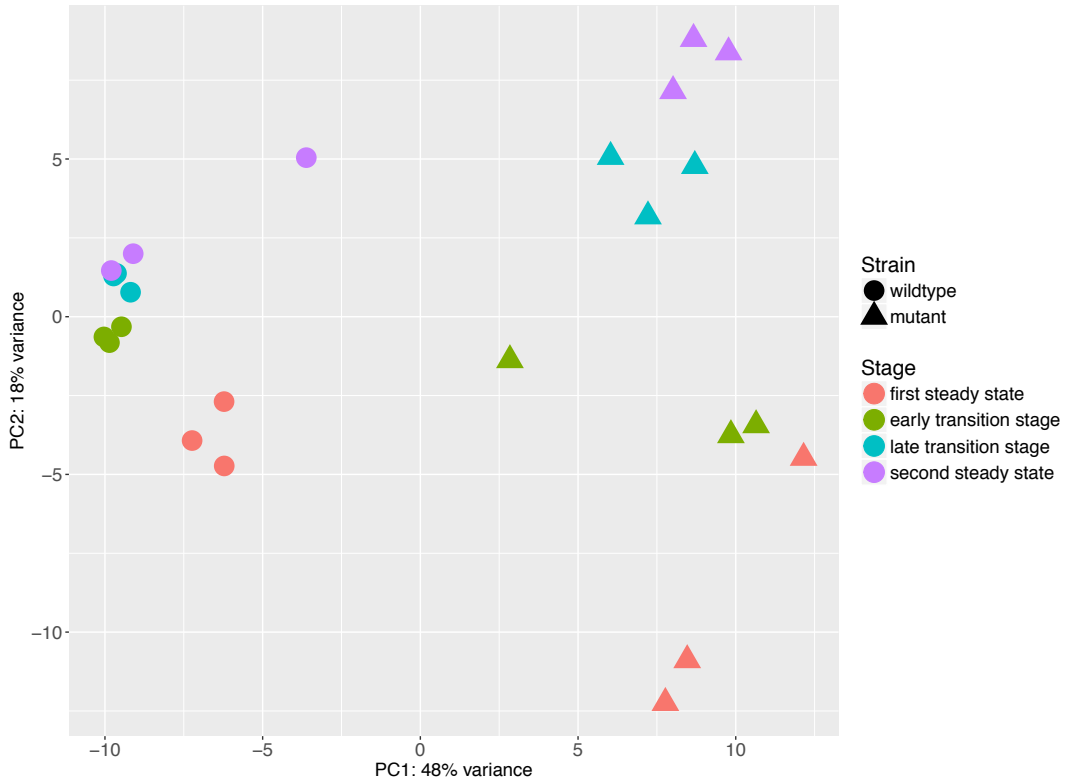


Figure 3.9 Principal Component Analysis (PCA) on the gene expression profiles of *P. gingivalis* W83 and Δ*ppad* mutant samples from different stages of the continuous culture.
The first component (PC1) showed that samples of *P. gingivalis* W83 (dot) and the Δ*ppad* mutant (triangle) were distinctly separated in all culture stages. The second components (PC2) demonstrated that genes expression in both strains were strongly related to the different culture stages (represented by different colours).

3.9.2 Differentially expressed genes between the two strains

In order to identify any differences in gene expression between *P. gingivalis* W83 wild-type and its Δ*ppad* mutant, which may explain the observed difference in the properties between the two strains in continuous culture, the gene expression profiles of the two strains were compared using DESeq2 R package based on the RNA sequencing data. There were 119, 103, 101 and 155 differentially expressed genes identified in the Δ*ppad* mutant culture compared with the wild-type at first steady-state, early transition stage, late transition stage and second steady-state, respectively (log-fold change > 1, adjusted $P < 0.01$). The number of up-and down-

regulated genes in the mutant at different culture stages is shown in Figure 3.10. In the first steady-state, where the pH was controlled at 7.25 ± 0.05, there were more down-regulated genes identified in the Δ*ppad* mutant. When the pH control was removed, more up-regulated genes were found at the transition stage and the second steady-state. Thus, it appears that PPAD deficiency can regulate the gene expression of *P. gingivalis* but the regulation may be affected by the environmental pH of the culture. By overlap analysis, the expression of 25 genes were significantly different between the two strains independent of pH conditions, indicating that the changes of the expression of these genes were mainly caused by the PPAD deficiency (Figure 3.11). Twenty-two of these genes were up-regulated and three were down-regulated in the mutant including the one encoding PAD (Table 3.2).

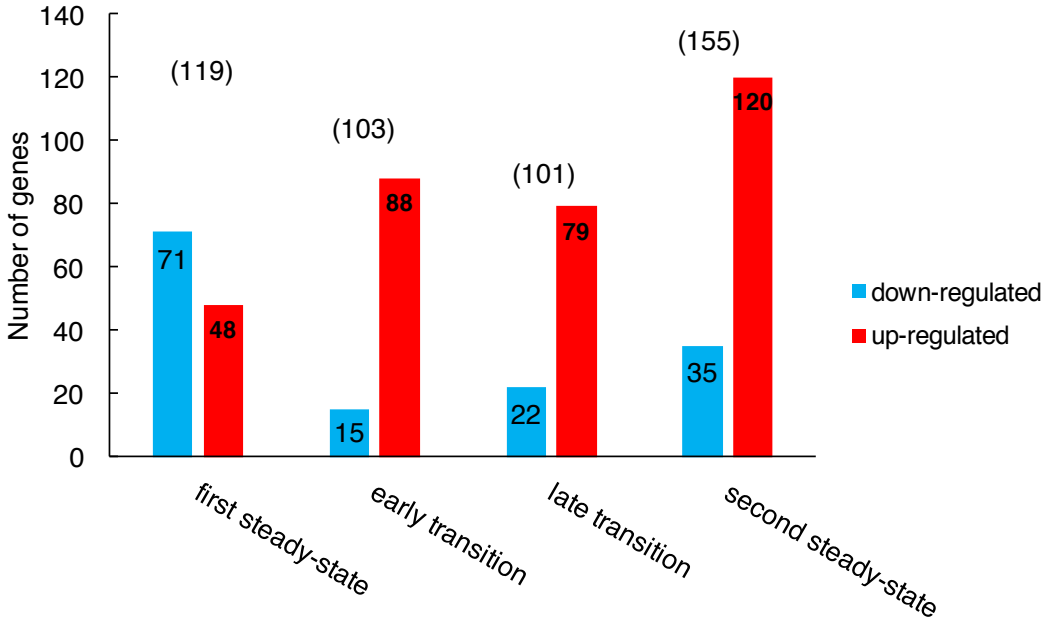


Figure 3.10 Number of genes differentially expressed in *P. gingivalis* W83 Δ*ppad* mutant compared with the wild-type when grown under similar conditions in a chemostat.
Numbers of up-and down-regulated genes are displayed (one-fold or more, adjusted $P < 0.01$, DESeq2). Totals of differentially expressed genes in each stage are shown in brackets.

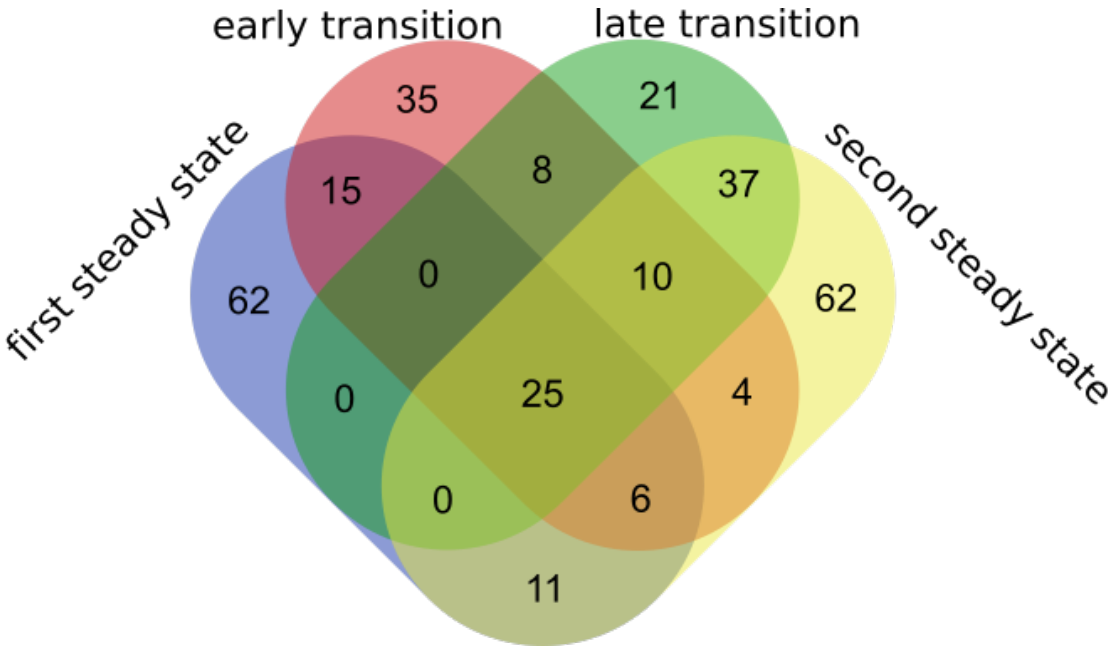


Figure 3.11 Venn diagram of the differentially expressed genes in *P. gingivalis* W83 Δ*ppad* mutant compared with the wild-type when grown under identical conditions in a chemostat.
The number of differentially expressed genes during different culture stages were summarised. Stage denotation: blue: first steady-state (pH externally controlled at 7.25 ± 0.05), red: early transition stage, green: late transition stage, yellow: second steady-state (the steady-state without pH control).

Table 3.2 Differentially expressed genes in *P. gingivalis* W83 Δ*ppad* mutant compared with its wild-type strain throughout the stages of the continuous culture (one-fold or more, adjusted $P < 0.01$, DESeq2).

Locus name	Protein name	log2 Fold Change			
		[a] first steady	[b] early transition	[c] late transition	[d] second steady
PG_1424	peptidyl-arginine deiminase	-7.7118	-7.1763	-7.25969	-6.8408
PG_0432	SAM-dependent methyltransferase	-2.7387	-2.6091	-2.6657	-3.0575
PG_0195	rubrerythrin family protein	-1.4932	-2.0349	-1.6448	-1.6310
PG_1543	acyl-CoA thioesterase	1.3225	2.2099	2.4115	2.4270
htrA	DegQ family serine endoprotease	1.3790	1.8235	2.2980	2.4807

folD	bifunctional methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase	1.3808	1.3661	1.3832	1.3289
PG_0646	ABC transporter ATP-binding protein	1.4939	2.1504	3.3110	3.6830
PG_1178	hypothetical protein	1.5818	3.5586	3.0720	2.5988
PG_0555	MULTISPECIES: histidinol phosphate phosphatase	1.7143	1.6896	2.2802	1.8662
PG_1868	MULTISPECIES: membrane protein	1.9063	2.8092	1.3983	1.8054
PG_0275	thiol reductase thioredoxin	1.9981	2.4441	2.2229	1.9293
PG_0645	adenosylcobinamide amidohydrolase	2.0561	3.5614	4.2718	4.3406
PG_1124	cob(I)yrinic acid a,c-diamide adenosyltransferase	2.1543	2.1754	1.9001	2.2796
PG_0421	DUF2807 domain-containing protein	2.1807	2.4253	2.5650	2.9259
PG_1180	membrane protein	2.2476	3.8918	2.5050	2.8005
PG_0707	TonB-dependent receptor	2.3806	3.3058	2.4387	2.5354
PG_1553	cobaltochelatase subunit CobN	2.5101	3.5516	2.2061	1.7080
PG_0495	T9SS C-terminal target domain-containing protein	2.5792	4.5655	2.8550	1.4768
PG_0686	DUF1858 domain-containing protein	2.8048	3.1056	2.9973	2.8858
hmuR	TonB-dependent receptor	2.8326	3.7629	1.9374	1.6697
PG_0173	DNA-binding protein	3.1840	3.1625	2.1381	1.2681
PG_1179	outer membrane lipoprotein-sorting protein	3.2196	3.8016	2.3684	3.2893

PG_0174	DUF1661 domain-containing protein	3.9931	4.1327	3.6834	3.0286
PG_1858	flavodoxin	4.0137	4.5150	1.8344	2.2097
hmuY	HmuY protein	4.9440	5.0379	2.2992	2.0669

[a] First steady-state: the steady-state under pH control at 7.25 ± 0.05;

[b] Early transition-stage: the first day after removing the pH control;

[c] Late transition-stage: the last day after removing pH control and before reaching the second steady-state;

[d] Second steady-state: the steady-state without pH control.

Within the second steady-state, 120 genes were up-regulated and 35 were down-regulated in the mutant compared with the wild-type strain. These were annotated to the gene ontology (GO) terms related to the biological process, molecular function and cellular component by using R package topGO. A Fisher exact test implemented in the package was used to define the GO terms enriched in these up- and down regulated genes in the mutant. All GO terms with enrichment $P < 0.05$ were summarized and visualized by REVIGO (Figure 3.12). The GO term of cell division was enriched in the up-regulated genes, which is in accordance with the higher OD_{600} values found in the continuous culture of Δ*ppad* mutant in the second steady-state (Table 3.1). This result indicated that PPAD deficiency could upregulate the expression of cell division related genes in the absence of additional pH control and then prompt bacterial growth. Moreover, the GO term of peptidase was also enriched in the up-regulated genes which can benefit *P. gingivalis* for obtaining energy and carbon sources. In addition, the GO term of oxidoreductase activity was enriched in the up-regulated genes which may explain at least in part, the decrease of the redox potential in the culture of mutant within the second steady-state.

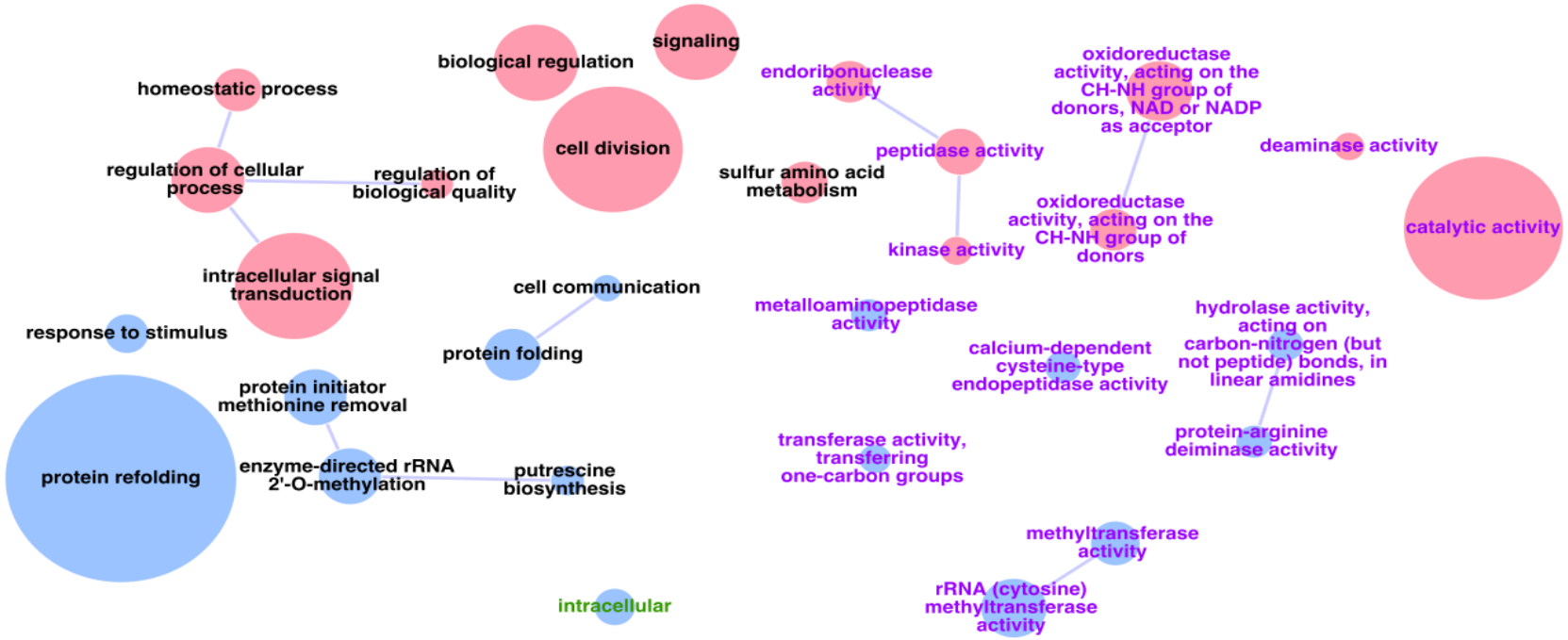


Figure 3.12 GO terms enriched in the up- and down regulated genes of Δ*ppad* mutant compared with the wild-type strain within the second steady-state.

All GO terms with enrichment $P < 0.05$ were summarized and visualized by REVIGO. The bubble size is inversely proportional to the enrichment *P* value. Red bubbles represent the up-regulated GO terms and blue bubbles represent down-regulated terms. Black label is for GO terms related to biological process, purple is for molecular function and green for cellular component. Highly similar GO terms are linked by edges in the graph.

3.9.3 Differentially expressed genes after removing pH control within each strain

In order to investigate the effect of environmental pH on the gene expression of *P. gingivalis* W83 and its Δ*ppad* mutant, the gene expression profiles were compared between different stages of continuous culture within each strain using DESeq2 R package based on the RNA sequencing data. Differential expression analysis revealed that within the wild-type strain there were 52, 68 and 82 genes differentially expressed (one-fold or more, adjusted *P* less than 0.01) at the early transition-stage, late transition-stage and second steady-state compared with the first steady-state in which the pH was controlled at 7.25 ± 0.05 (Figure 3.13a). Within the Δ*ppad* mutant, 11, 64 and 151 genes were identified as differentially expressed at different stages after removing the pH control (Figure 3.13b). Remarkably, no up-regulated genes were found in the transition-stages of the wild-type strain and only six were found in the second steady-state with the cut-off of one-fold or more, adjusted $P < 0.01$. In contrast, for the mutant, nine, 47 and 110 genes were identified as being significantly up-regulated at early-transition, late-transition and second steady-state, respectively.

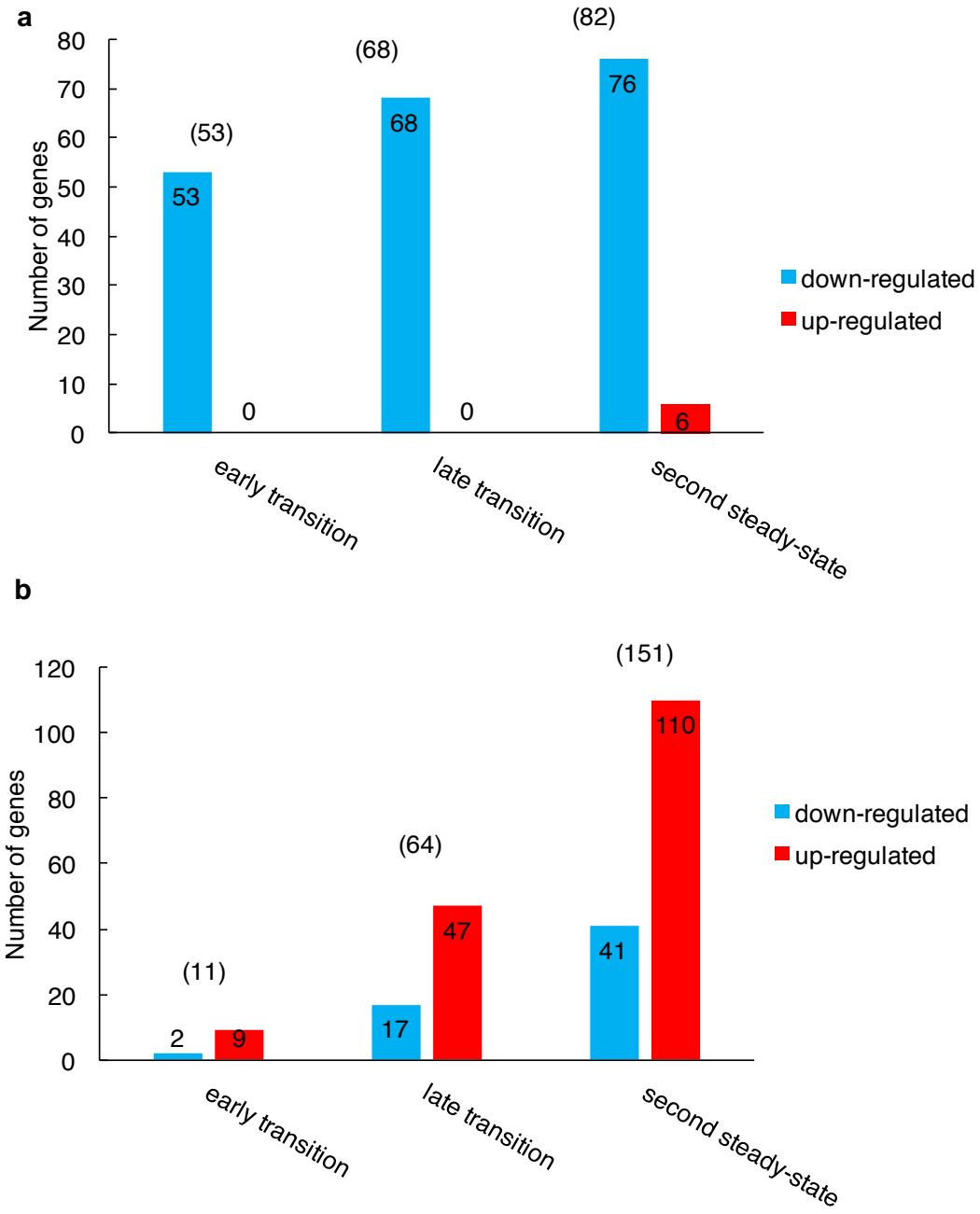


Figure 3.13 Summary of differentially expressed genes after removing pH control within each strain.
Differentially expressed genes (one-fold or more, adjusted $P < 0.01$) in the transition stages and second steady-state compared with the first steady-state (pH controlled at 7.25 ± 0.05) of the continuous culture were identified for (a) *P. gingivalis* W83 wild-type and (b) Δ*ppad* mutant using DESeq2 package. Numbers of up-and down-regulated genes are displayed. Totals of differentially expressed genes in each stage are shown in brackets.

To identify the genes that were differentially expressed throughout the continuous culture stages after removing pH control or only in a certain stage, overlap analysis

was performed on those genes differentially expressed in the early and late transition stages as well as the second steady-state, compared with the first steady-state. There were 39 genes in the wild-type strain that were found to be differentially expressed throughout the culture stages without pH control and they were all down-regulated (Figure 3.14a). Whereas the expression of four genes in the mutant were continuously altered after removing pH control including two up-regulated and two down-regulated genes (Figure 3.14b).

In both the wild-type strain and the mutant, PG_1837(*hagA)*, *nrd* and PG_0612, encoding haemagglutinin A, ribonucleotide reductase and a hypothetical protein, respectively, were up-regulated in second steady-state compared with the first steady-state. PG_0090 encoding stationary phase protection protein was down-regulated (Appendix C2&3). These genes were indicated to be mainly regulated by the changes of environmental pH in the present study and such regulation may possibly help *P. gingivalis* to adapt to the new environment conditions after removing pH control.

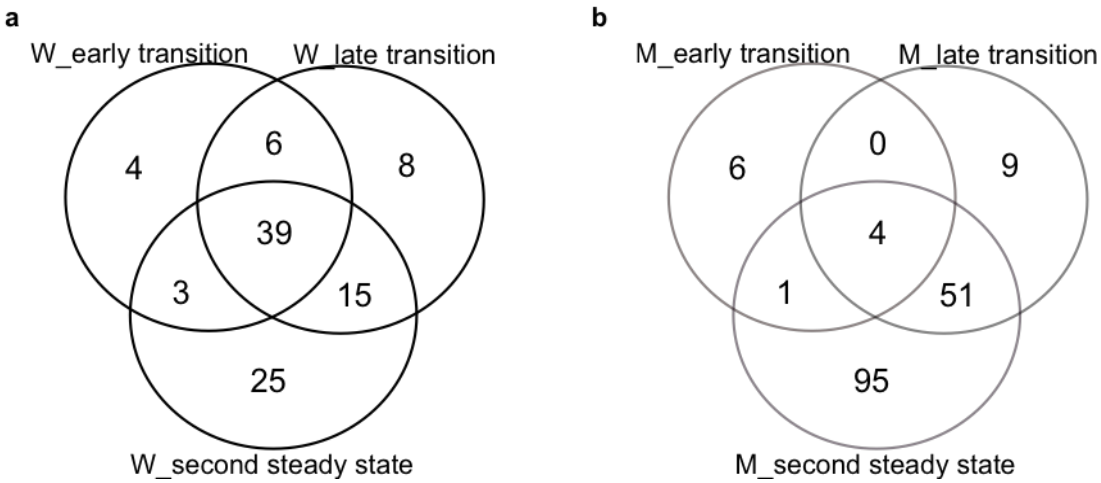


Figure 3.14 Overlap analysis of the differentially expressed genes after removing pH control within each strain
Overlap analysis (Venn diagram) was performed on the differentially expressed genes (one-fold or more, adjusted $P < 0.01$) in the early transition, late transition and second steady-state compared with the first steady-state, to find genes that were differentially expressed throughout the continuous culture stage after removing pH control or only in a certain stage for (a) *P. gingivalis* W83 (W) and (b) Δ*ppad* mutant (M).

The top 20 differentially expressed genes with highest log2-fold change (absolute value) in the second steady-state compared with the first steady-state of *P. gingivalis* W83 were all down-regulated. The details of these genes are listed in Table 3.3. For the Δ*ppad* mutant, there were only three down-regulated genes in the list of the top 20 differentially expressed gene comparing between the two steady-states (Table 3.4).

Table 3.3 The top 20 differentially expressed genes with highest log2 fold change (absolute value) in the second steady-state compared with the first steady-state of *P. gingivalis* W83.

Locus name	Product	log2FoldChange	Adjusted *P*
PG_0337	hypothetical protein	-4.2217	4.91E-07
PG_1489	relaxase	-3.5555	1.08E-11
PG_1482	DUF4133 domain-containing protein	-3.4436	5.62E-05
PG_2114	NA	-3.0610	3.76E-05
PG_1480	MULTISPECIES: DUF4141 domain-containing protein	-2.9307	3.06E-09
PG_1479	conjugative transposon protein TraJ	-2.8690	9.70E-08
PG_2063	hypothetical protein	-2.7277	4.54E-05
PG_1478	MULTISPECIES: conjugative transposon protein TraK	-2.7123	4.43E-06
PG_1485	DUF3408 domain-containing protein	-2.5879	6.23E-07
PG_1486	ParA family protein	-2.5554	1.56E-08
PG_0718	hypothetical protein	-2.5508	4.50E-12
PG_0732	hypothetical protein	-2.5498	1.46E-03
PG_1020	TonB-dependent receptor	-2.5457	2.31E-37
PG_1490	conjugal transfer protein TraG	-2.4133	3.48E-11
PG_0283	RND transporter	-2.3776	6.71E-11
PG_1022	hypothetical protein	-2.3482	2.70E-10
PG_1398	hypothetical protein	-2.3021	1.33E-03
PG_1975	hypothetical protein	-2.2371	3.54E-03
PG_0285	TolC family protein	-2.2360	2.51E-09
PG_1481	TraG family conjugative transposon ATPase	-2.2189	1.88E-09

Table 3.4 The top 20 differentially expressed genes with highest log2 fold change (absolute value) in the second steady-state compared with the first steady-state of *P. gingivalis* W83 Δ*ppad* mutant.

Locus name	Product	log2FoldChange	Adjusted *P*
PG_1660	sigma-70 family RNA polymerase sigma factor	-3.7961	1.78E-03
PG_1659	hypothetical protein	-2.5665	4.84E-03
PG_0173	DNA-binding protein	-2.3125	2.25E-16
PG_2013	CRISPR-associated endonuclease Cas2	2.3056	6.41E-04
PG_1982	CRISPR-associated endonuclease Cas1	2.3752	3.07E-16
PG_0646	ABC transporter ATP-binding protein	2.3799	9.22E-11
PG_0492	hypothetical protein	2.4021	1.69E-04
PG_0611	hypothetical protein	2.4939	7.71E-44
PG_1988	hypothetical protein	2.4972	5.03E-23
PG_0411	T9SS C-terminal target domain-containing protein	2.5447	1.40E-65
PG_1984	hypothetical protein	2.5449	2.99E-07
PG_1514	glycerol dehydrogenase	2.5594	1.01E-06
PG_1892	hypothetical protein	2.5650	8.51E-04
PG_1989	hypothetical protein	2.5861	1.22E-07
PG_0645	adenosylcobinamide amidohydrolase	2.6448	3.68E-08
PG_1983	type III-B CRISPR module RAMP protein Cmr6	2.8422	8.52E-16
PG_1837	hemagglutinin A	2.9884	5.42E-30
PG_0613	hypothetical protein	3.3050	2.00E-33
PG_0865	ISAs1 family transposase ISPg2	3.6163	5.32E-03
PG_0612	hypothetical protein	4.0687	8.58E-44

Differentially expressed genes in the second steady-state compared with the first steady-state were annotated to the GO terms by using R package topGO. For the wild-type strain, there were 17, 27 and 25 differentially expressed genes involved in the biological process, cellular component and molecular function respectively. For the Δ*ppad* mutant, more differentially expressed genes were mapped to the GO term where 54 genes for the biological process, 47 for cellular component and 74 for molecular function.

A Fisher exact test was used to define the enriched GO terms based on the differentially expressed genes between the two steady-states for each strain. GO terms enriched ($P < 0.05$) in up- or down regulated genes in the second steady-state of each strain were summarized and visualized by REVIGO. The enrichment analysis showed that genes related to the pathogenesis of the wild-type strain were up-regulated in the second steady-state when the bacteria were cultured without environmental pH control (Figure 3.15). Within the mutant, enrichment of GO terms of cell division, haemin binding and peptidase activity in the up-regulated genes was supportive for the evidence of increased OD_{600} and cell numbers in the second steady-state compared with the first steady-state. The overrepresented cell division related genes were only found in the mutant but not in the wildtype when comparing the two steady-states, which supports the result that OD_{600} of the mutant was significantly higher than that of the wild-type strains during the second steady-state (Table 3.1). Taken together, above data suggest that environmental pH of the culture can affect the growth of *P. gingivalis* W83 Δ*ppad* mutant via regulation of growth-related gene expression (Figure 3.16). Main findings of the transcriptomic analysis of chemostat cultre were summarized in Table 3.5.

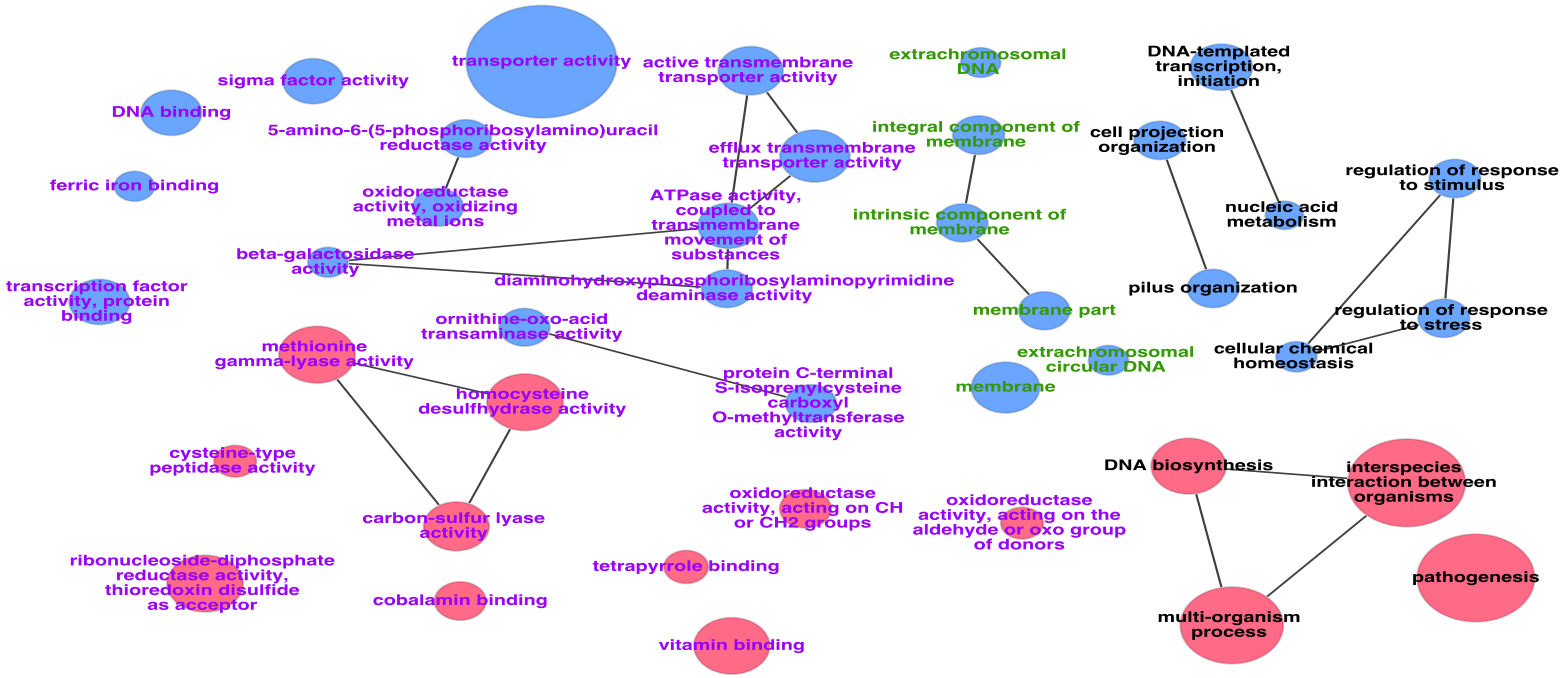


Figure 3.15 GO terms enriched in the up- and down regulated genes in the second steady-state (without pH control) compared with the first steady-state (pH controlled at 7.25 ± 0.05) within the continuous culture of *P. gingivalis* W83.
All GO terms with enrichment $P < 0.05$ were summarized and visualized by REVIGO. The bubble size is inversely proportional to the enrichment *P*. Red bubbles represent the up-regulated GO terms and blue bubbles represent down-regulated terms. Black label is for GO terms related to biological process, purple is for molecular function and green for cellular component. Highly similar GO terms are linked by edges in the graph.

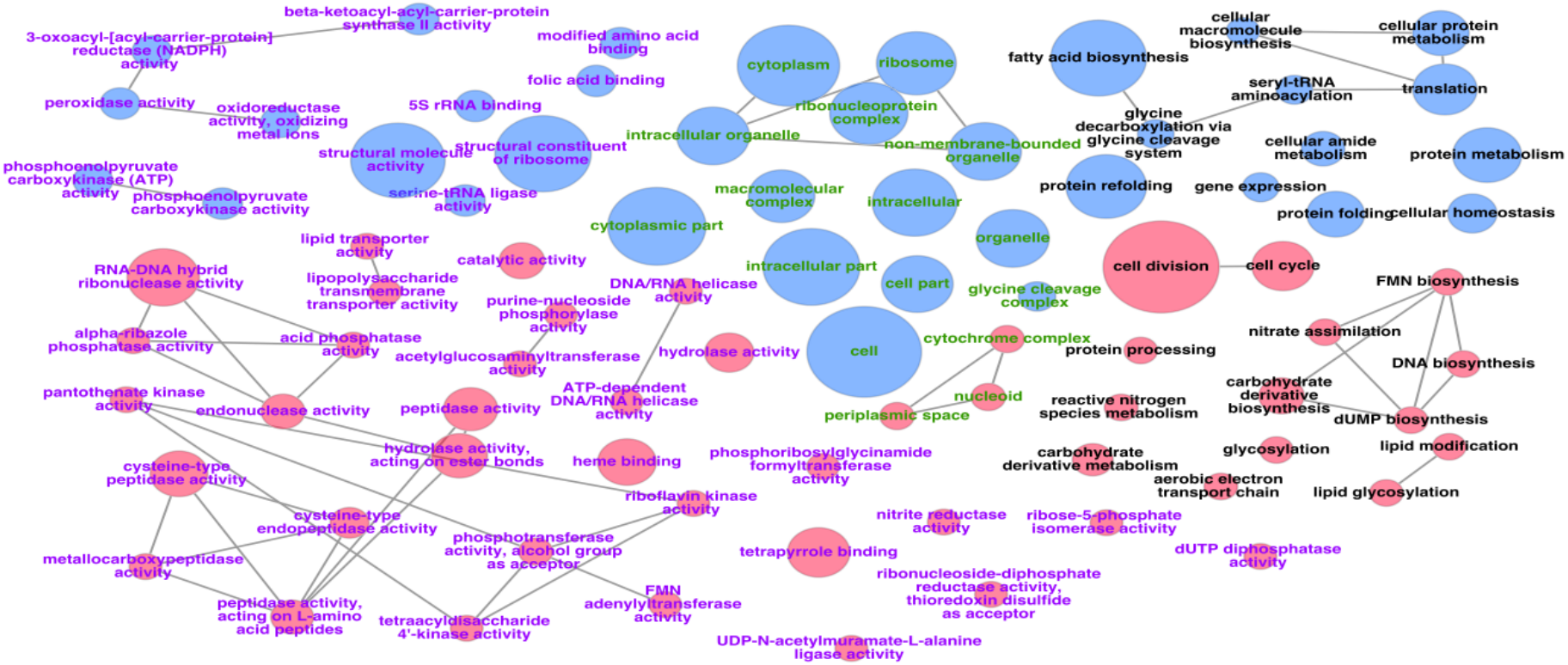


Figure 3.16 GO terms enriched in the up- and down regulated genes in the second steady-state (without pH control) compared with the first steady-state (pH controlled at 7.25 ± 0.05) within the continuous culture of W83 Δ*ppad* mutant.
All GO terms with enrichment $P < 0.05$ were summarized and visualized by REVIGO. The bubble size is inversely proportional to the enrichment *P*. Red bubbles represent the up-regulated GO terms and blue bubbles represent down-regulated terms. Black label is for GO terms related to biological process, purple is for molecular function and green for cellular component. Highly similar GO terms are linked by edges in the graph.

Table 3.5 Summary of the main findings of the transcriptomic analysis of chemostat cultre.

Transcriptomic analysis results	Description
Δ*ppad* mutant vs. wild-type strain	
PG_0275 encoding thiol reductase thioredoxin was over-expressed in the Δ*ppad* mutant compared with wild-type strain.	In accordance with the decrease in the redox potential of chemostat culture of the mutant.
The GO term of cell division was enriched in the up-regulated genes of Δ*ppad* mutant in the second steady-state compared with the wild-type strain.	This result is in accordance with the higher OD_{600} values found in the chemostat culture mutant compared with the wild-type strain, indicating that PPAD deficiency could upregulate the expression of cell division related genes in absence of additional pH control and then prompt the bacteria growth.
Second steady-state vs. first steady state	
HmuY and *hagA* of *P. gingivalis* W83 wild-type were significantly up-regulated in the second steady-state compared with the first steady-state	Regulation of those genes related with haem acquisition of *P. gingivalis* may help the bacteria to adapt to the environmental changes and facilitate bacterial growth.
CRISPR-associated genes of Δ*ppad* mutant were up-regulated in the second steady-state compared with the first steady-state	CRISPR systems can protect bacterial against the foreign genetic elements, thus environmental changes can upregulate the expression of CRISPR-associated genes in Δ*ppad* mutant, which may lead to an increase in the defence capability of the Δ*ppad* mutant.

Type IX secretion system (T9SS) genes of Δ*ppad* mutant were up-regulated in the second steady-state compared with the first steady-state.	Environmental changes can regulate the expression of some T9SS genes in the *P. gingivalis* Δ*ppad* mutant and may subsequently facilitate the secretion of virulence factors.

3.10 Metagenomic study of subgingival microbiome in relation to RA

3.10.1 Optimization of DNA extraction from subgingival plaque samples and validation of DNA library preparation with low-yield samples

Preliminary experiment of DNA extraction was conducted following the manufacturer's instructions of Microbial DNA Isolation Kit on eight additional subgingival plaque samples. The average DNA yields was 5.67 ng and three of the samples had less than 5 ng DNA which is the minimum input requirement of NEBNext® Ultra™ DNA Library Prep Kit according to its protocol. Thus, the normal protocol of the kit was modified to increase the yields of DNA extraction. The paper-points carrying the plaque were thoroughly vortexed in the Micro-bead solution from the kit to re-suspend the cells and the samples were additionally incubated with lysozyme (10 mg/mL) at 37 °C for 10 minutes to break the cell walls. Then the extraction continued with combined chemical and mechanical lysis as described in the protocol of the kit. The DNA yields of each sample were quantitated by using the PicoGreen kit, which is varying between different plaque samples (mean is 8.13 ng, range from 0 to 76.25 ng). There was no significant

difference of the DNA between four groups either in healthy site or diseased site samples (Figure 3.17) (Kruskal-Wallis test, $P > 0.05$). Among total 196 samples, there were 154 samples with less than 10 ng DNA extracted (Figure 3.18). For those samples, NEBNext® Ultra™ II DNA Library Prep Kit was used for library preparation which has a broader range of input amounts (500 pg -1 μg). After clean-up to remove unwanted adaptor dimers, primer-dimers and other contamintants, the average amount of DNA libraries is 888.98 ng (Figure 3.19). All the libraries obtained are ready for sequencing, which is also validated by checking the DNA size distribution on Tape-station (approximately 300 bp, data not shown).

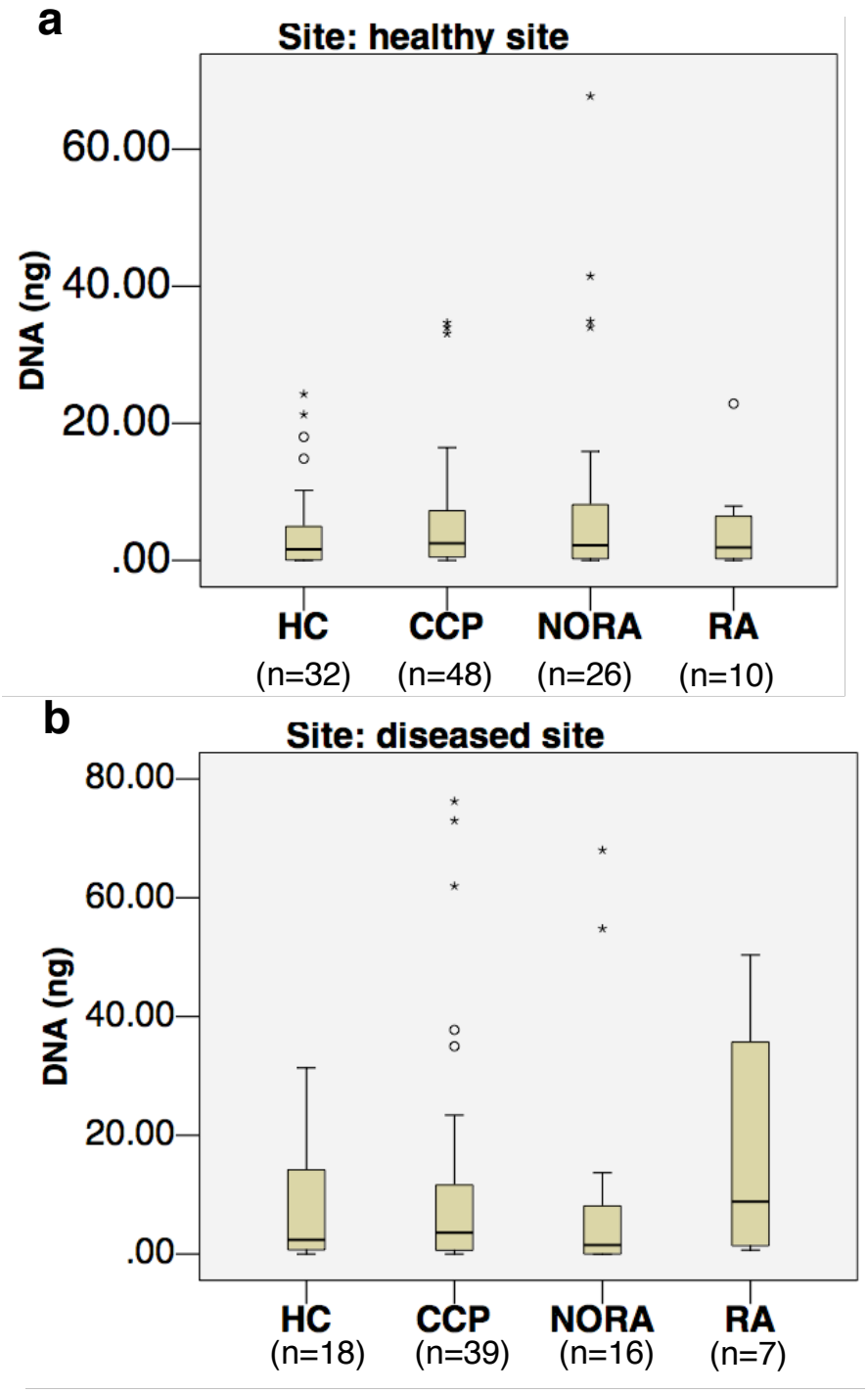


Figure 3.17 DNA extracted from subgingival plaque samples.
No significant difference of DNA quantity was found between four groups either in (a) healthy site or (b) diseased site samples. n: number of samples

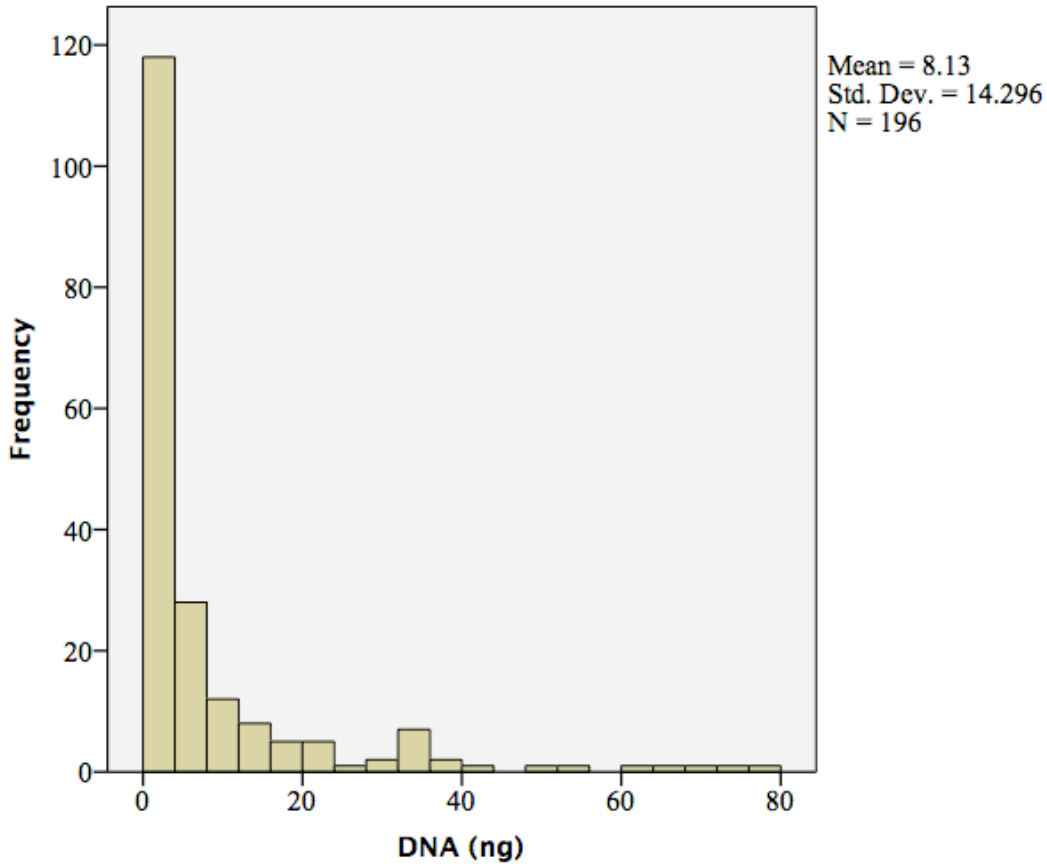


Figure 3.18 Histogram of DNA yields from subgingival plaque samples. Samples were thawed on ice from -80°C. The samples of diseased sites or healthy sites were pooled together for each participant. DNA was extracted from pooled samples using the modified UltraClean® Microbial DNA Isolation Kit and quantified by using PicoGreen ® dsDNA Reagent and Kits.

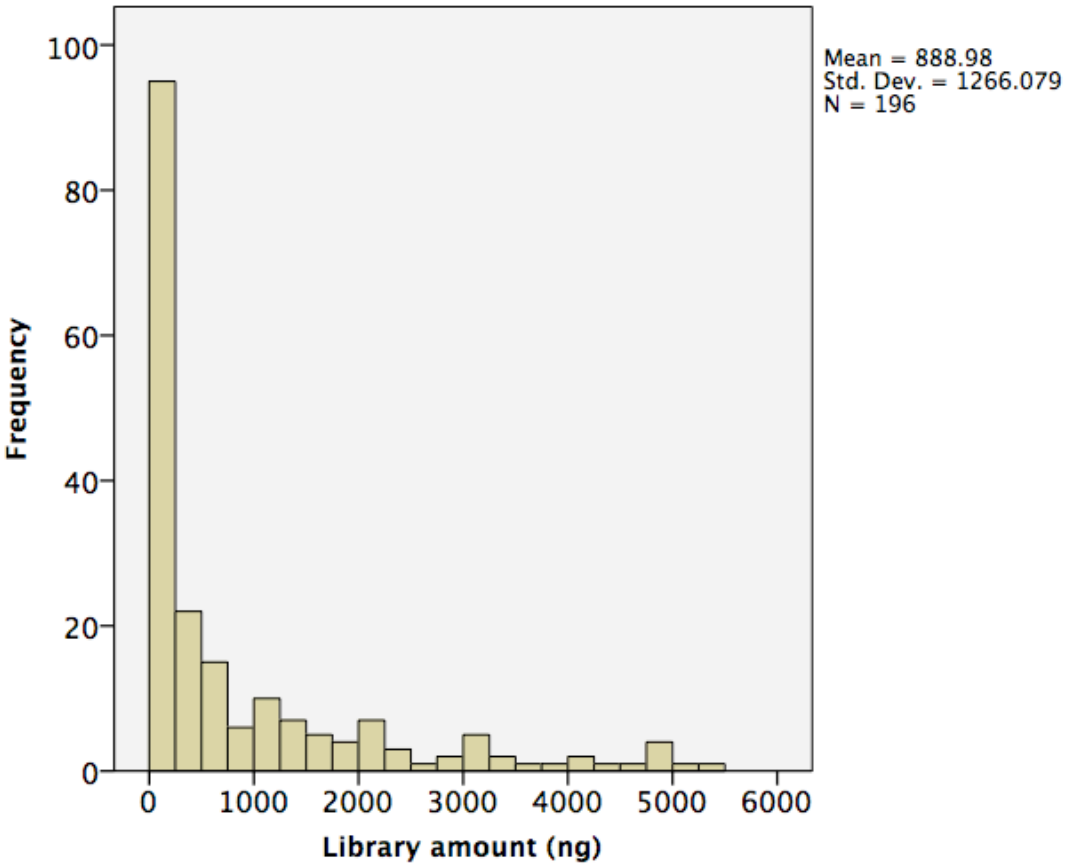


Figure 3.19 Histogram of amounts of DNA libraries.
Either NEBNext® Ultra™ DNA Library Prep Kit (for DNA input > 10 ng) or NEBNext® Ultra™ II DNA Library Prep Kit (for DNA input < 10 ng) was used for library constructions. AxyPrep™ Mag PCR Clean-up beads were used for the clean-up steps during and after the library preparation. The amount of purifed DNA library of each sample was quntified by using PicoGreen ® dsDNA Reagent and Kits.

3.10.2 General information of sequencing data

A total of 1178 gigabases of sequence reads were generated from 196 subgingival dental plaque samples by paired-end sequencing, resulting in an average of 39.8 ± 62.5 million reads per sample. In total 19.9 million reads passed quality control out of which 19.7 million reads were successfully assigned to NCBI taxonomy. Overall, 69.81% of all annotated reads were identified as coming from bacteria. The 116 samples from periodontally healthy sites and 80 from diseased sites were grouped

according to the classification of participants with respect of RA and periodontitis (Table 3.6).

Table 3.6 Description of subgingival plaque samples.

Group	**HC** [a]	**CCP** [b]	**NORA** [c]	**RA** [d]
(number of participants)	**(n=32)**	**(n=48)**	**(n=26)**	**(n=10)**
Age: mean (SD)	49.4 (15.3)	51.9 (11.4)	54.4 (16.7)	62.2 (14.6)
Female: n (%)	19 (59.4)	31 (64.6)	14 (53.8)	7 (70.0)
Ever smoker: n (%)	18 (56.3)	31 (64.6)	17 (65.4)	6 (60.0)
Periodontitis patients [e]**: n (%)**	13 (40.6)	35 (72.9)	14 (53.8)	2 (20.0)
Sample:				
Perio-healthy sites [f]**: n**	32	48	26	10
Perio-diseased sites [g]**: n**	18	39	16	7

[a] Healthy control.

[b] Anti-CCP positive individuals who are at risk of RA development.

[c] New onset RA patients (NORA) who are anti-CCP positive and within 3 months of commencing DMARD therapy.

[d] Patients with chronic RA, defined as ≥ 6 months DMARD therapy.

[e] Diagnosis of periodontitis was given by dentists based on their consensus decision.

[f] Periodontally healthy sites were defined as sites of 3 mm depth or less with no bleeding on probing.

[g] Periodontally diseased sites were defined as sites with pockets of 4 mm depth or more with bleeding on probing.

3.10.3 α-diversity

The α-diversity (observed species and Shannon index) of each sample was calculated at species level using raw count data within bacteria domain. I first compared the α-diversity between periodontally healthy sites and diseased sites irrespective of RA status. Both observed species and Shannon index were significantly increased in diseased sites compared with healthy sites (Figure 3.20) (Mann-Whitney test, $P < 0.05$).

Then I studied the impact of RA status on microbial diversity using samples from periodontally healthy sites and diseased sites separately, irrespective of periodontitis status of the patients (Figure 3.21). Within periodontally healthy sites, a significantly decreased species richness (observed species) was found in NORA group compared with CCP and HC groups (Kruskal-Wallis test, $P < 0.05$). When the Shannon index (which estimates the species evenness in a specific sample) was applied, a significant difference was only found between NORA and CCP groups ($P < 0.05$). No significant differences were found among the groups when analysing samples from diseased sites. Except for the RA group, the diversity in HC, CCP and NORA group displayed a decreasing trend in both healthy and diseased sites.

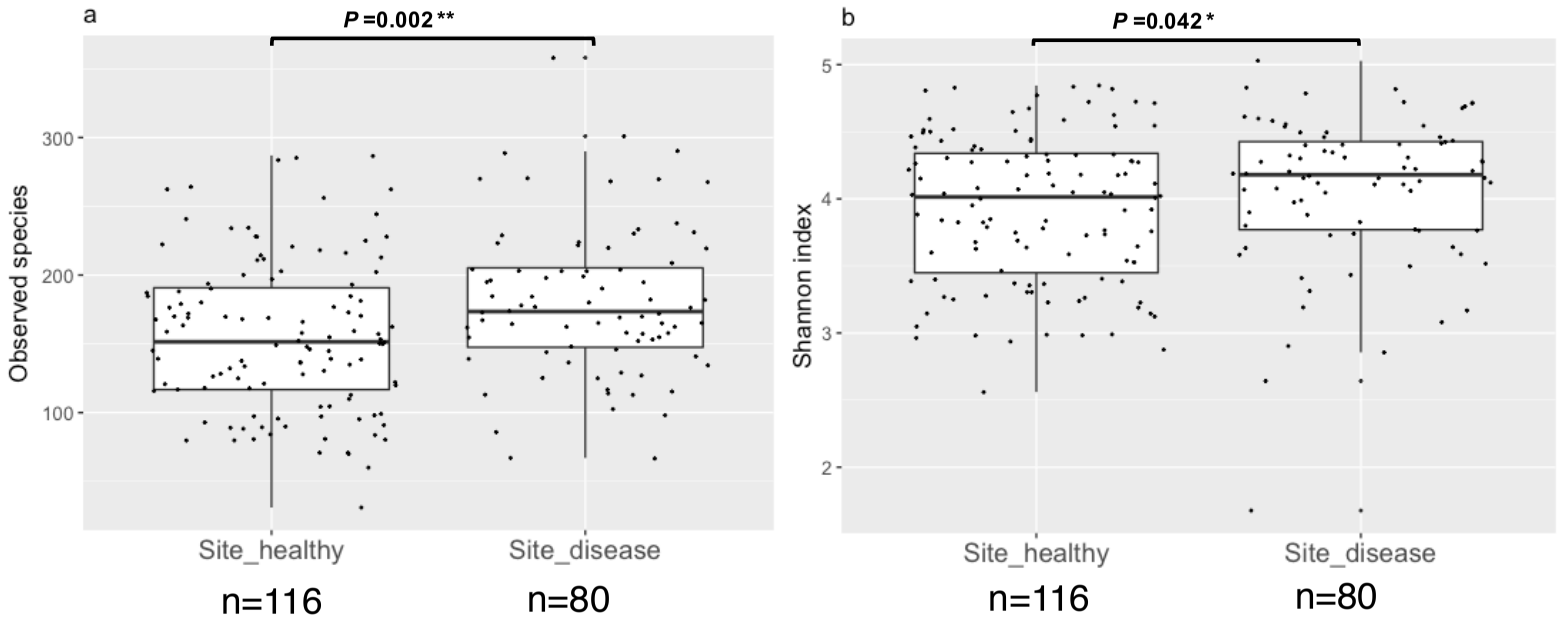


Figure 3.20 Comparison of α-diversity of samples from periodontally healthy sites and diseased sites.
A significantly increased α-diversity was found in periodontally diseased sites compared with healthy sites by Mann-Whitney test ($P < 0.05$). (a) Observed species as inferred from raw counts were compared between healthy sites and diseased sites. (b) Shannon index as inferred from raw counts were compared between sites. *: $P < 0.05$, **: $P < 0.01$. n: number of samples.

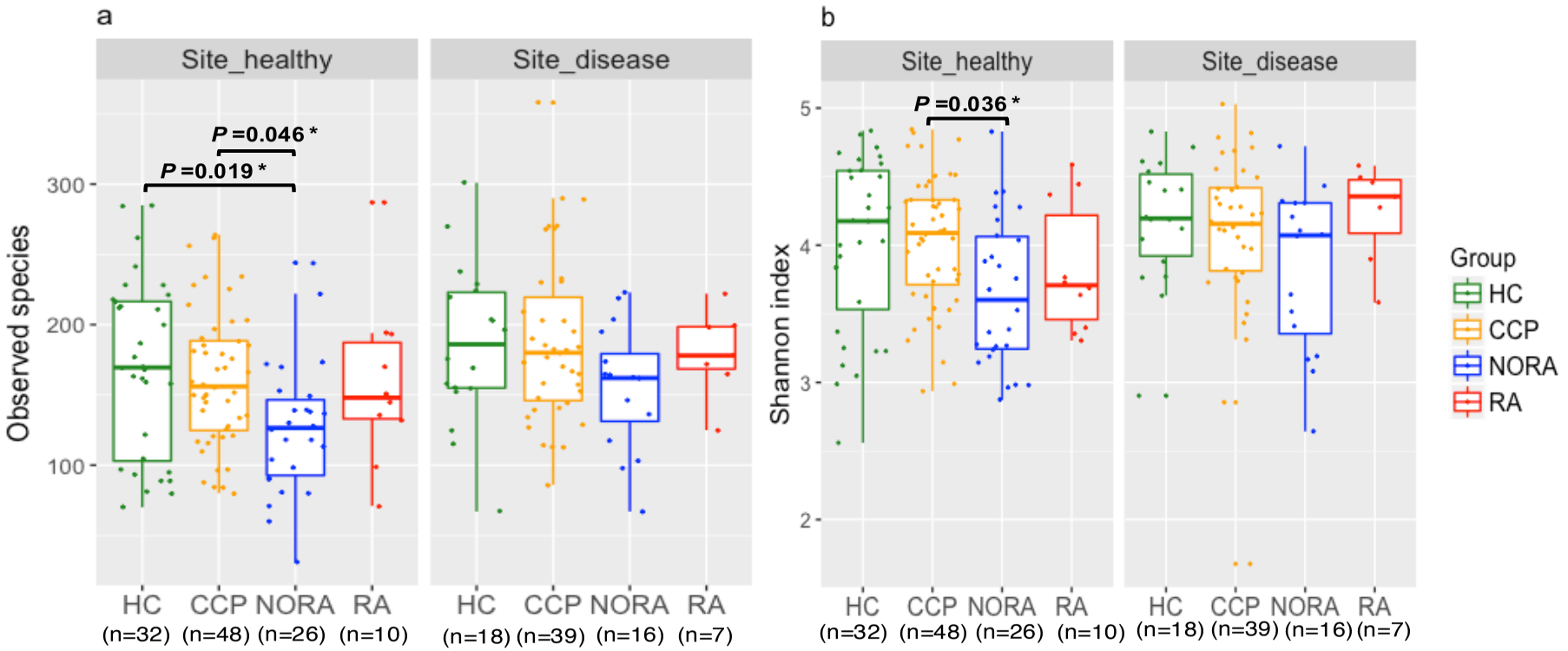


Figure 3.21 Comparison of α-diversity in different groups using samples from healthy sites and diseased sites.
(a) Significantly decreased α-diversity (observed species) was found in the NORA group compared with CCP and HC groups in periodontally healthy sites by Kruskal-Wallis Test ($P < 0.05$). However, no significant difference was found between the four groups in diseased site samples. (b) The Shannon index only showed a significant difference between CCP and NORA group in healthy site samples, Kruskal-Wallis test ($P < 0.05$). *: corrected $P < 0.05$ (Bonferroni correction). n: number of samples.

3.10.4 β-diversity

Filtering was performed to remove taxa accounting for less than ten sequences in total and observed in less than three samples. β-diversity at species level was determined by Bray-Curtis dissimilarity based on relative abundance and plotted using principal coordinates analysis (PCoA). Permutational multivariate analysis of variance (PERMANOVA) statistical tests (*Adonis* function, vegan package, R) showed a significantly different β-diversity between groups in periodontally healthy site samples (P =0.002, R^2 = 0.045) and in diseased site samples (P = 0.002, R^2 = 0.068) (Figure 3.22). Homogeneity of multivariate dispersions was tested by using the *betadisper* function to ensure PERMANOVA assumptions (significant result is not due to differences in group dispersions).

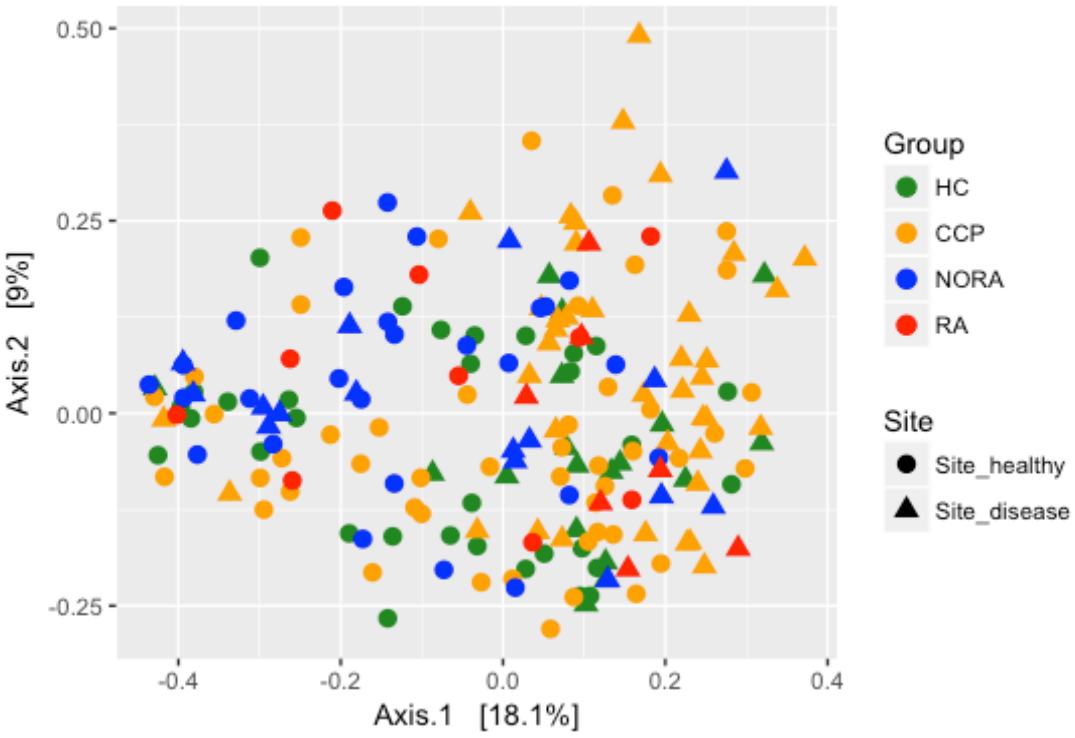


Figure 3.22 β-diversity determined by Bray-Curtis dissimilarity and plotted using PCoA.
Raw data were filtered to remove taxa accounting less than ten sequences in total and observed in less than three samples. A significantly different β-diversity was found between groups in periodontally healthy sites samples (P =0.002, R^2 = 0.045) and in diseased sites samples (P = 0.002, R^2 = 0.068) by PERMANOVA test.

3.10.5 Taxonomic profiles

Overall, 14 bacterial phyla, 195 genera and 772 species were identified in this study. In an attempt to discriminate among study groups, I analysed the bacterial community composition and structure at the various taxonomic levels (phylum, genus and species). According to the average phylum assignment result within each group, *Actinobacteria* was the most predominant phylum in all groups. *Bacteroidetes* were the second followed by *Firmicutes* (Figure 3.23). *Candidatus Kryptonia* was only found in diseased sites. The permutation test (one-sided *signassoc* function, R) was used to test the significant differences of relative

abundance of each phylum between different groups in periodontally heathy sites or diseased sites. *Spirochaetes* was found with significantly higher relative abundance in the CCP group compared with other groups in healthy sites as well as diseased sites. *Actinobacteria* were more abundant in the NORA group compared with other groups in diseased sites ($P < 0.05$).

The top 20 predominant bacterial genera were identified for healthy sites and diseased sites separately, based on relative abundance (Figure 3.24). Comparing the list of top 20 predominant bacterial genera between healthy and diseased sites, 19 genera were in common. The differences of the two lists were that *Methylobacterium* was only in the list of healthy sites and *Alloprevotella* was only in diseased sites. *Actinomyces* was the most abundant genus in all participant groups, followed by *Prevotella* and *Streptococcus*, except within the RA group in diseased site samples where *Streptococcus* was more abundant than *Prevotella.* Within healthy sites or diseased sites, genera with significantly higher relative abundance between groups were investigated by using the permutation test (one-sided *signassoc* function, R). *Capnocytophaga* and *Treponema* were with significantly more abundant in periodontally healthy sites in the CCP group compared with other groups. *Methylobacterium* and *Pseudopropinobacterium* were more abundant in NORA. The RA group had more *Bradyrhizobium* than other groups ($P < 0.05$). Within periodontally diseased sites, *Porphyromonas* and *Treponema* were significantly higher in the CCP compared with other groups, while *Corynebacterium* was found to be more abundant in samples from NORA patients. Apart from these predominant genera, other genera were also found with significantly higher relative abundance between groups (Appendix C4). Within healthy sites, *Peptostreptococcus* was more abundant in the HC group than in other groups. *Bacillus* and *Vibrio* were

more abundant in the RA group. For diseased sites, Neisseria, *Cardiobacterium*, *Haemophilus* and *Klebsiella* were found with higher relative abundance in the NORA group. *Stomatobaculum* was significantly more abundant in the RA group.

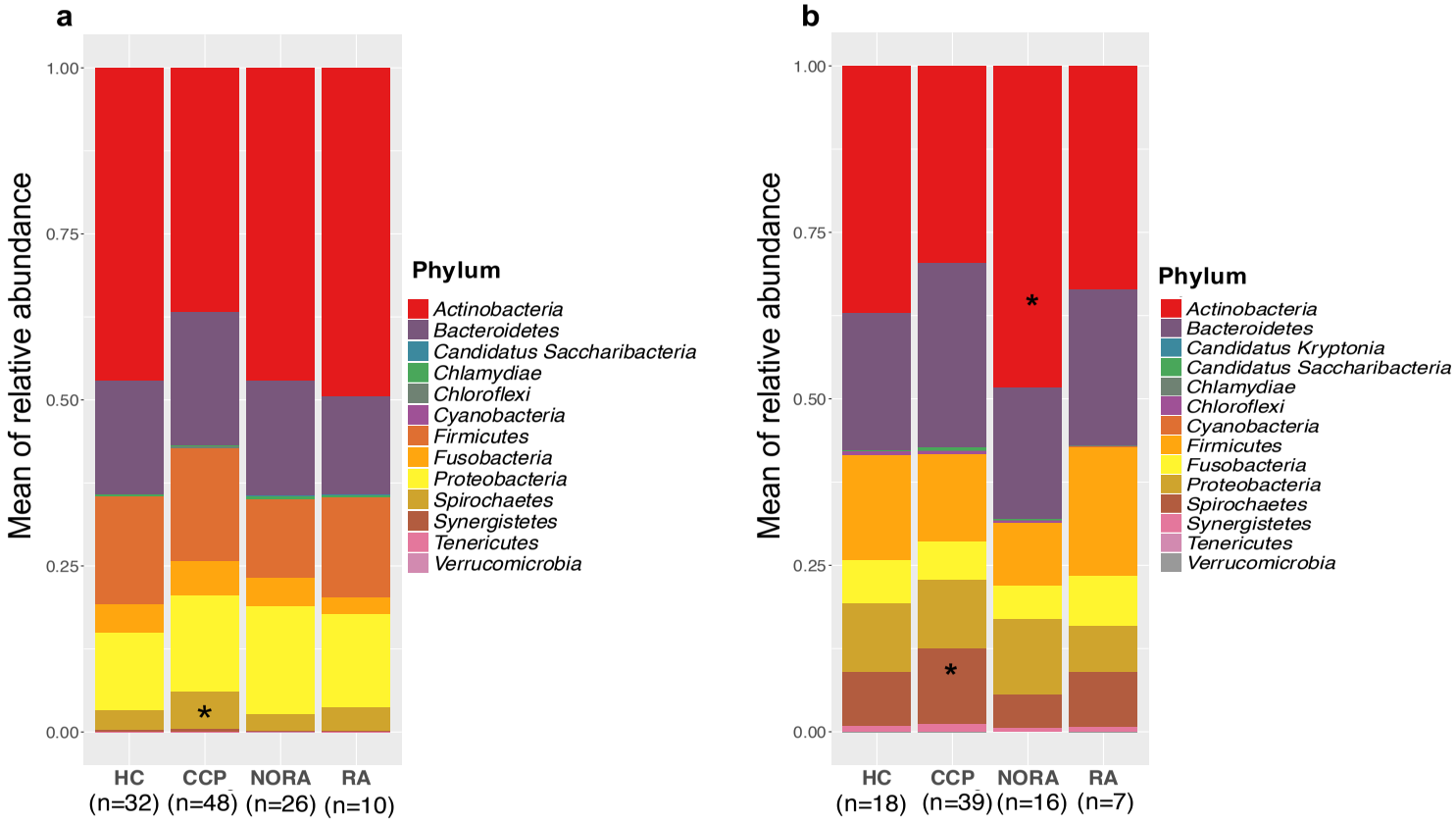


Figure 3.23 Phylum composition of different groups.
The relative abundance of the oral microbiota (phylum level) was compared between different groups in (a) periodontally healthy sites and (b) diseased sites. The permutation test (one-sided *signassoc* function, *indicspecies* R-package) was used to find the phyla with significantly different relative abundances between groups. *: corrected $P < 0.05$ (Sidak's correction). n: number of samples.

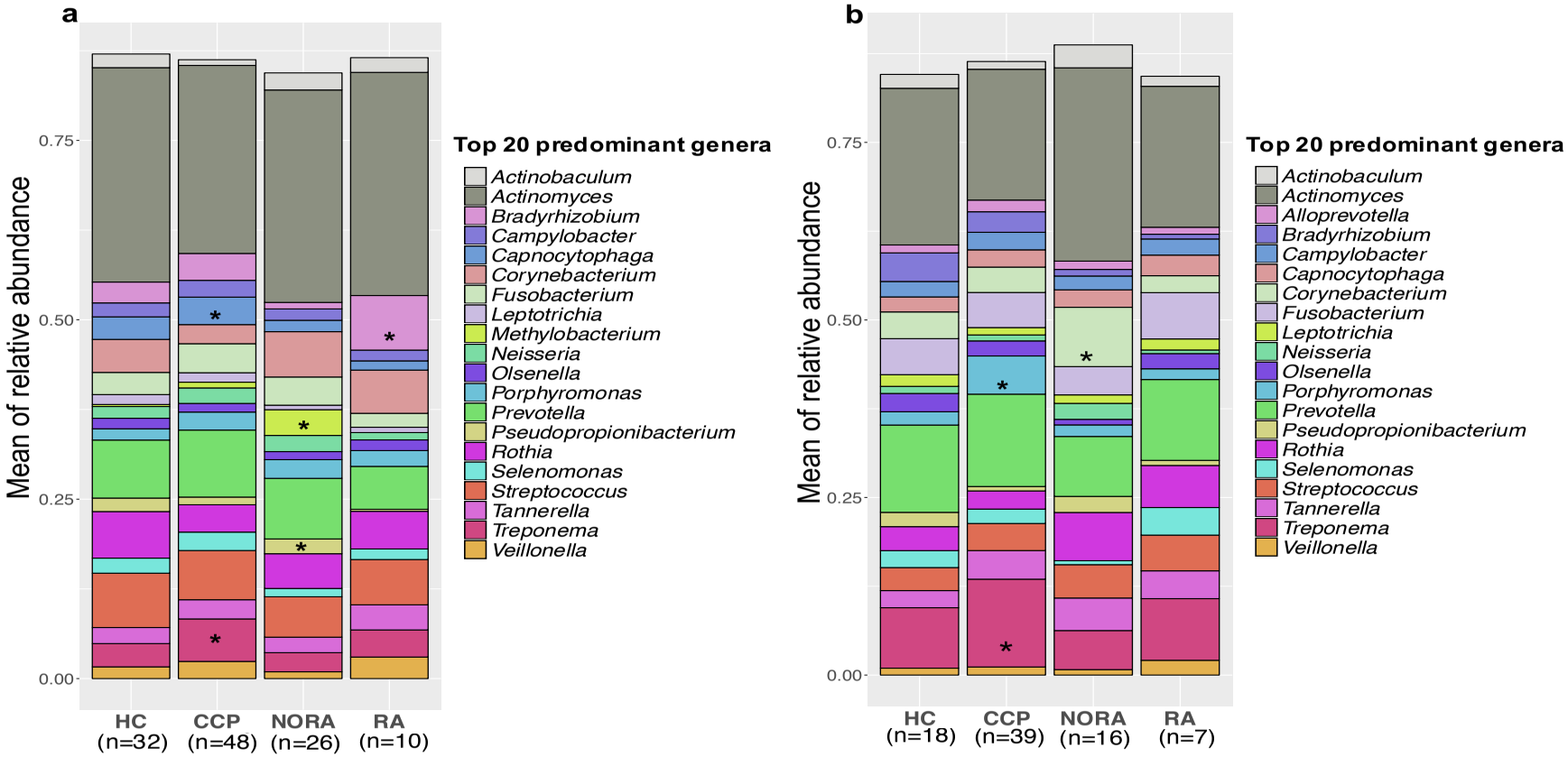


Figure 3.24 Taxonomic profiles for the 20 most abundant genera in healthy sites and diseased sites.
Relative abundance of the 20 most abundant genera was plotted for each group within periodontally healthy sites (a) and diseased sites (b). The permutation test (one-sided *signassoc* function, *indicspecies* R-package) was used to find the genera with significantly different relative abundances between groups. *: corrected $P < 0.05$ (Sidak's correction). n: number of samples.

3.10.6 Bacterial species associated with different groups

To identify bacterial species with significantly different relative abundance between groups, the permutation test (one-sided *signassoc* function) from indicspecies R-package was used to test the statistical significance of species-group association. Within periodontally healthy sites, 19 CCP group-associated, four NORA-associated, and two RA-associated bacterial species were identified based on relative abundance of assigned DNA reads ($P < 0.05$). No HC-associated species were found (Figure 3.25). However, in periodontally diseased site samples, there were six HC-associated, five CCP group-associated, 11 NORA-associated, and one RA-associated bacterial species identified (Figure 3.26). *P. gingivalis* and *T. denticola* were among the species associated with CCP, this would be consistent with the increased presence of periodontitis in the CCP group. These species of healthy or diseased sites were listed in Table 3.7 by their significance (corrected *P* value).

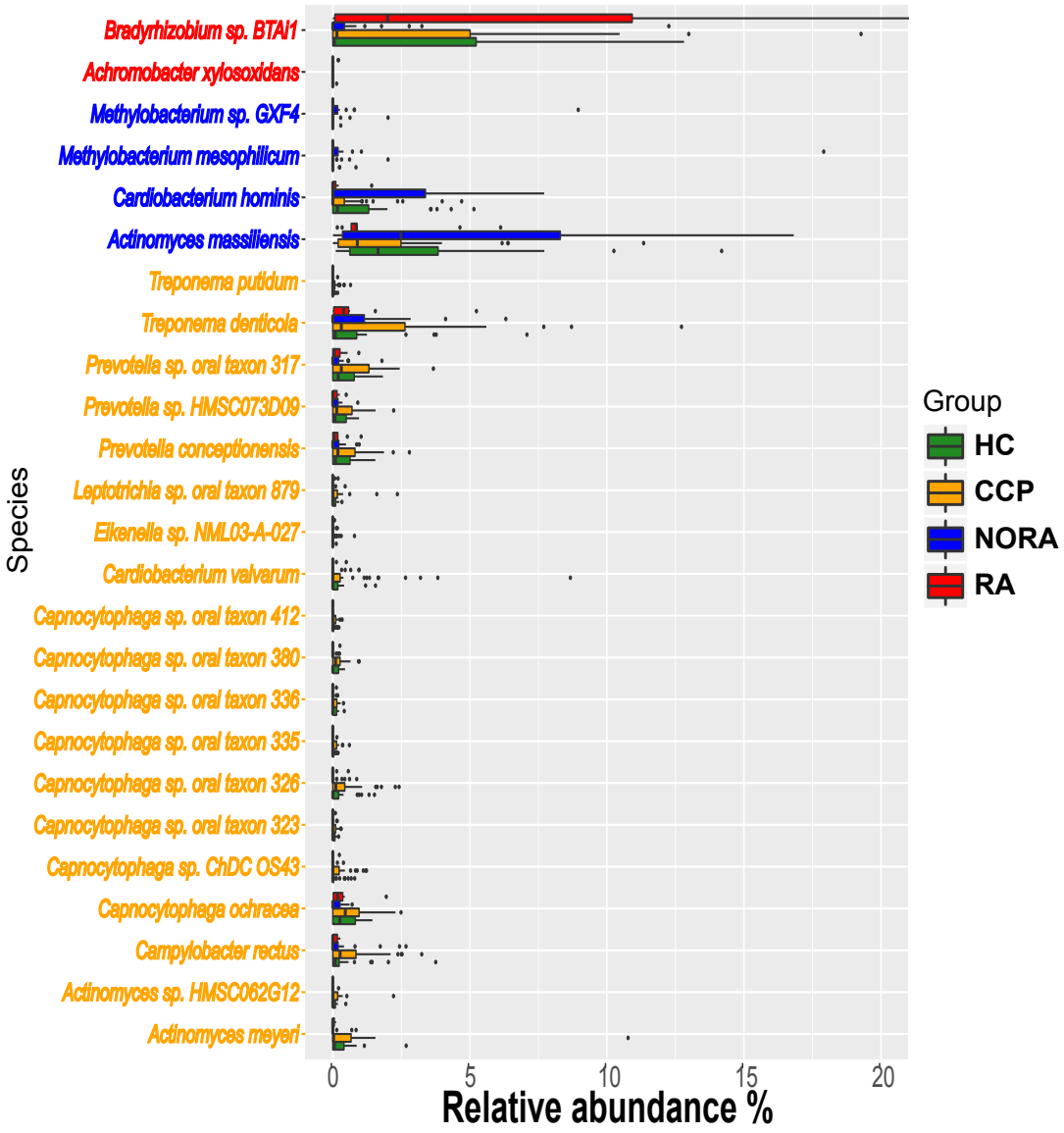


Figure 3.25 Bacterial species with significantly higher relative abundance in HC, CCP, NORA and RA groups in periodontally healthy site samples. Species-group association was tested by using one-sided *signassoc* function (indicspecies R-package) based on the relative abundance. Sidak's correction was applied for multiple testing. Species with significantly higher relative abundance were selected for each group. Significantly associated species: green: HC, yellow: CCP, blue: NORA, red: RA.

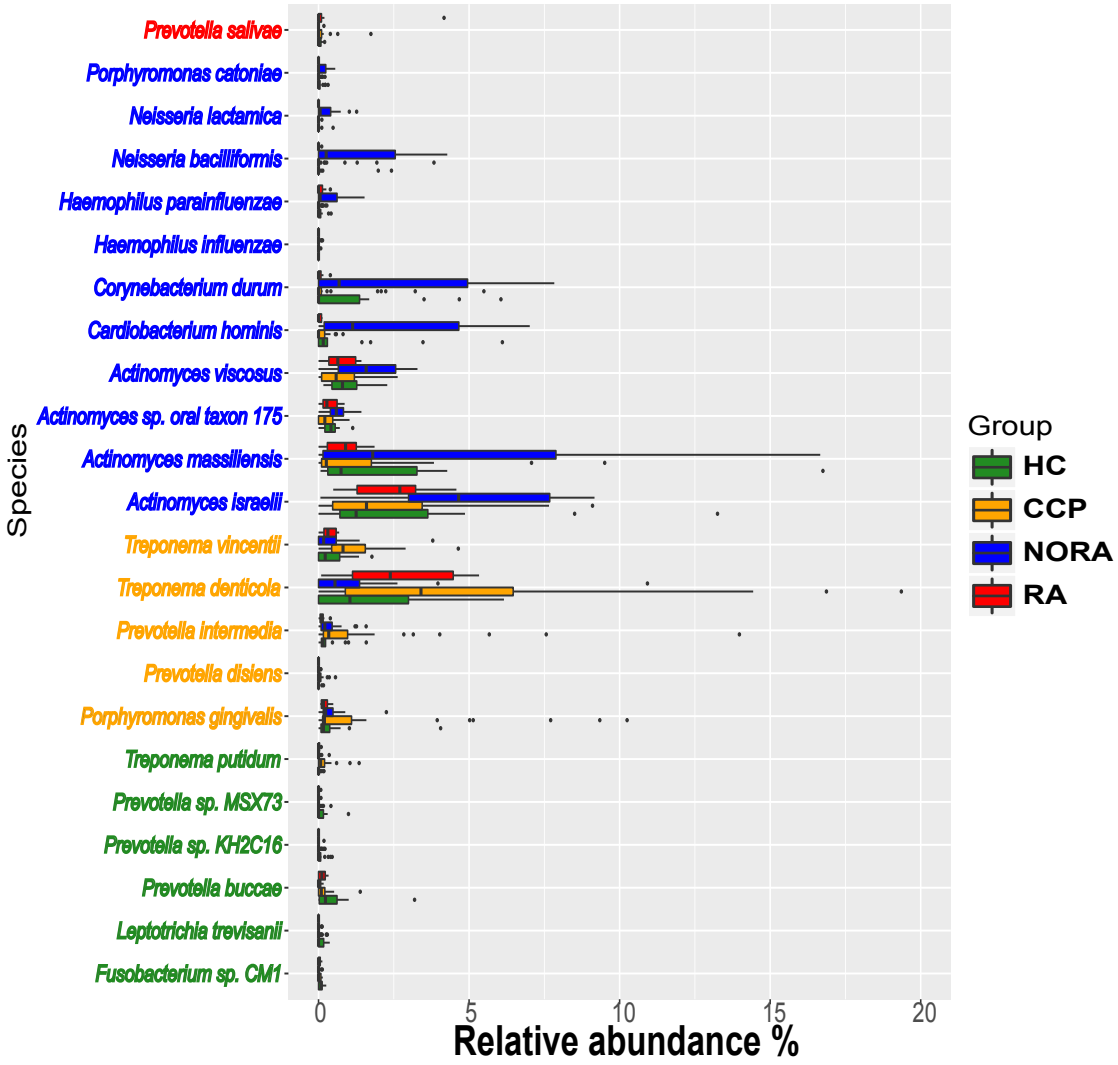


Figure 3.26 Bacterial species with significantly higher relative abundance in HC, CCP, NORA and RA groups in periodontally diseased site samples. Species-group association was tested by using one-sided *signassoc* function (indicspecies R-package) based on the relative abundace. Sidak's correction was applied for multiple testing. Species with significantly higher relative abundance were selected for each group. Significantly associated species: green: HC, yellow: CCP, blue: NORA, red: RA.

Table 3.7 Bacterial species with significantly higher relative abundance in HC, CCP, NORA and RA groups in periodontally healthy or diseased sites.

Species	Group	Corrected *P**
Healthy site		
Capnocytophaga ochracea	CCP	0.0040
Capnocytophaga sp. ChDC OS43	CCP	0.0040
Capnocytophaga sp. oral taxon 326	CCP	0.0040
Capnocytophaga sp. oral taxon 335	CCP	0.0040
Capnocytophaga sp. oral taxon 380	CCP	0.0040
Capnocytophaga sp. oral taxon 412	CCP	0.0040
Prevotella conceptionensis	CCP	0.0040
Prevotella sp. HMSC073D09	CCP	0.0080
Prevotella sp. oral taxon 317	CCP	0.0080
Treponema putidum	CCP	0.0080
Cardiobacterium hominis	NORA	0.0119
Actinomyces sp. HMSC062G12	CCP	0.0119
Capnocytophaga sp. oral taxon 336	CCP	0.0159
Cardiobacterium valvarum	CCP	0.0159
Methylobacterium mesophilicum	NORA	0.0199
Actinomyces meyeri	CCP	0.0199
Methylobacterium sp. GXF4	NORA	0.0238
Capnocytophaga sp. oral taxon 323	CCP	0.0238
Eikenella sp. NML03-A-027	CCP	0.0238
Achromobacter xylosoxidans	RA	0.0277
Campylobacter rectus	CCP	0.0316
Leptotrichia sp. oral taxon 879	CCP	0.0433
Bradyrhizobium sp. BTAi1	RA	0.0471
Actinomyces massiliensis	NORA	0.0471
Treponema denticola	CCP	0.0471
Diseased site		
Treponema putidum	HC	0.004
Cardiobacterium hominis	NORA	0.004
Neisseria bacilliformis	NORA	0.004
Neisseria lactamica	NORA	0.004
Porphyromonas gingivalis	CCP	0.008
Treponema vincentii	CCP	0.008
Prevotella buccae	HC	0.008
Porphyromonas catoniae	NORA	0.008
Treponema denticola	CCP	0.0119
Leptotrichia trevisanii	HC	0.0119
Actinomyces israelii	NORA	0.0119
Actinomyces viscosus	NORA	0.0119

Corynebacterium durum	NORA	0.0119
Prevotella sp. KH2C16	HC	0.0159
Prevotella intermedia	CCP	0.0199
Prevotella disiens	CCP	0.0238
Actinomyces massiliensis	NORA	0.0238
Prevotella sp. MSX73	HC	0.0277
Haemophilus parainfluenzae	NORA	0.0277
Haemophilus influenzae	NORA	0.0316
Fusobacterium sp. CM1	HC	0.0355
Actinomyces sp. oral taxon 175	NORA	0.0471
Prevotella salivae	RA	0.0471

*: Sidak's correction was applied for multiple testing.

3.10.7 Common and unique species in different groups

Overlap analysis on species level was performed with Venn diagrams to provide the number of species for one, two, three or all investigated groups. By analysing samples from periodontally healthy sites, a total of 616 species were found, of which 234 species were common to all groups. Every group was found to have its own unique species and CCP had the highest number while HC had the lowest (Figure 3.27a). In diseased site samples, 609 species were found in total, of which 221 species were shared between the four groups. Similarly, CCP had the highest number at 166 and HC had the lowest at 14 (Figure 3.27b). For the details of uniquely detected species in each group see Appendix C5 and C6.

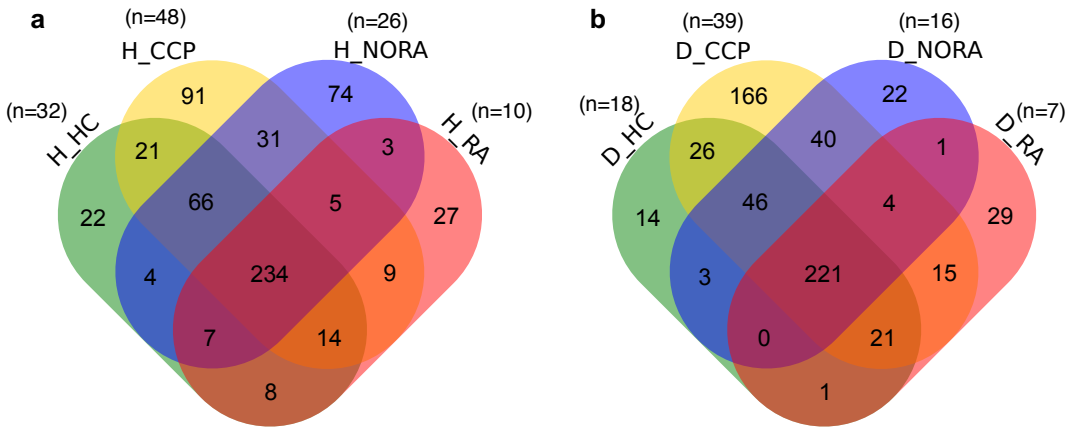


Figure 3.27 Overlap analysis of the group specific and shared species. samples from healthy sites; (b) samples from diseased sites. Group denotation: green: HC, yellow: CCP, blue: NORA, red: RA. n: number of samples.

3.10.8 Co-occurrence networks of bacterial species

Co-occurrence networks of the species with occurrence in at least 20% of samples in each group were constructed to explore the topological and taxonomic characteristics of microbial co-occurrence patterns. Based on Spearman's rank correlation analysis, 603, 571, 187 and 130 edges were identified as the strong and significant pairwise correlations between species ($q < -0.5$ or > 0.5 and adjusted $P < 0.05$), in HC, CCP, NORA and RA groups in periodontally heathy site samples (Figure 3.28-3.31). In periodontally diseased site samples, there were 235, 720, 296 and 65 edges identified in HC, CCP, NORA and RA groups (Figure 3.32-3.35). The edge/node ratio (density) of the network of CCP group is higher than that of other groups in periodontally diseased sites, reflecting a higher number of co-occurrence instances in the CCP group. Other network-level topological features were listed in Table 3.8 for each group, including the clustering coefficient (CC) and average shortest path length (APL).

In all groups, majority of the edges (> 87.5%) were identified as positive correlation ($q > 0.5$) (Table 3.8). There were 49 negative pairwise correlation between the species in HC group in periodontally heathy sites and *Actinomyces* species were involved in 63% of these correlation (Figure 3.28). Analogously, 57.8% of negative correlation in CCP group were between *Actinomyces* species and other species (Figure 3.29). In both groups, a negative correlation was found between *Actinomyces naeslundii* and *P. gingivalis*. In periodontally diseased site samples, NORA group had the highest number of negative pairwise correlation where 64.9% were between *Actinomyces* species and other species (Figure 3.34). In CCP group of the periodontally diseased site samples, *P. gingivalis* was negatively correlated with *Actinomyces oris* and three *Veillonella* species (Figure 3.33). Topological features of each node were also calculated for the networks (data not shown). *Treponema socranskii* has the highest degree (number of edges connected to node) and betweenness centrality (number of shortest paths going through a node) in the network of HC group in periodontally healthy sites, indicated a core location in the network. In the network of CCP group, *T. socranskii* has the highest degree while *Campylobacter showae* exhibits the highest betweenness centrality.

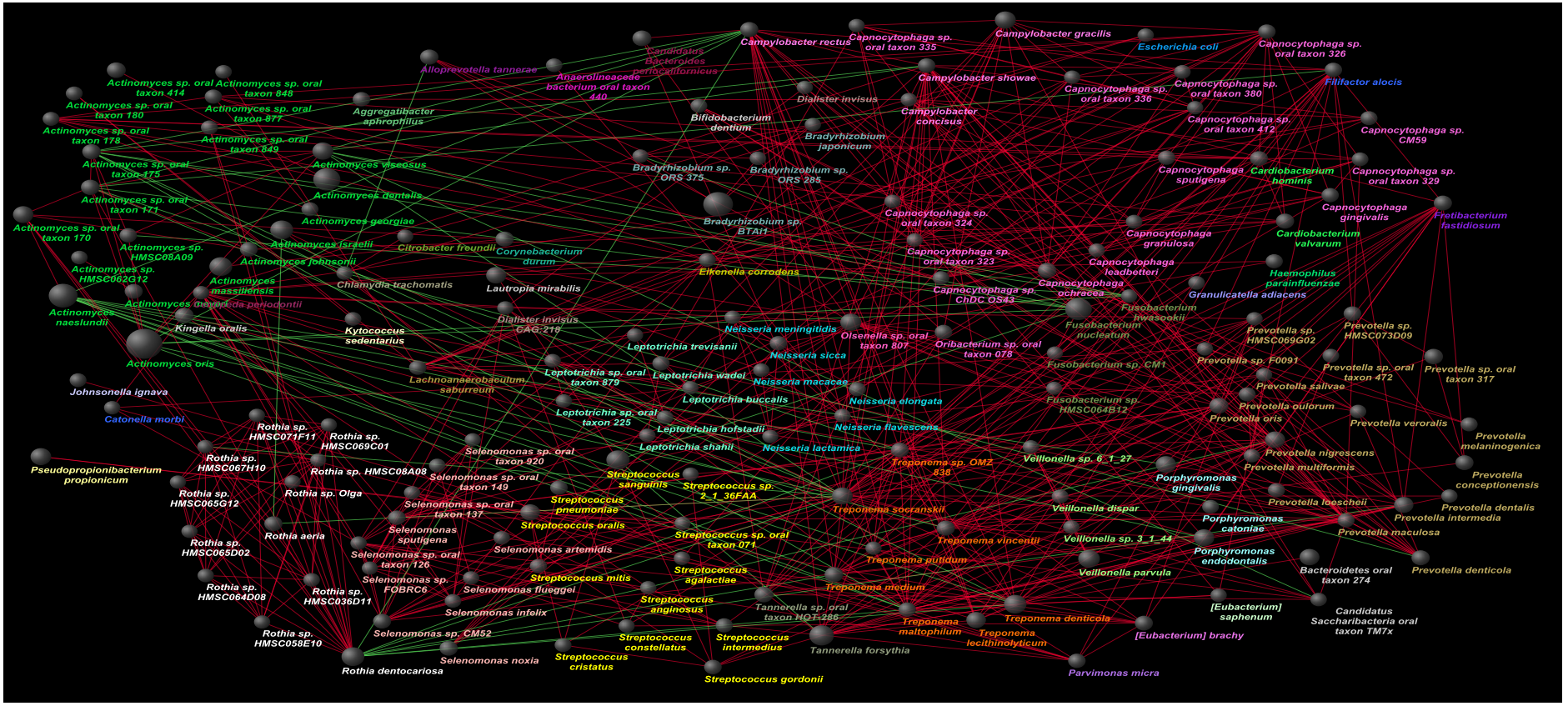


Figure 3.28 Co-occurrence networks of the species in HC group from periodontally healthy site samples.
Spearman's rank coefficients (q) between species with occurrence in at least 20% of samples in the group were calculated pairwise. Edge stands for a strong ($q > 0.5$ or $q < -0.5$) and significant (adjusted $P < 0.05$) correlation. Red edge is for the positive correlation and green for the negative. The size of each node is proportional to the relative abundance of the species. The species were coloured by the genus-level taxonomy.

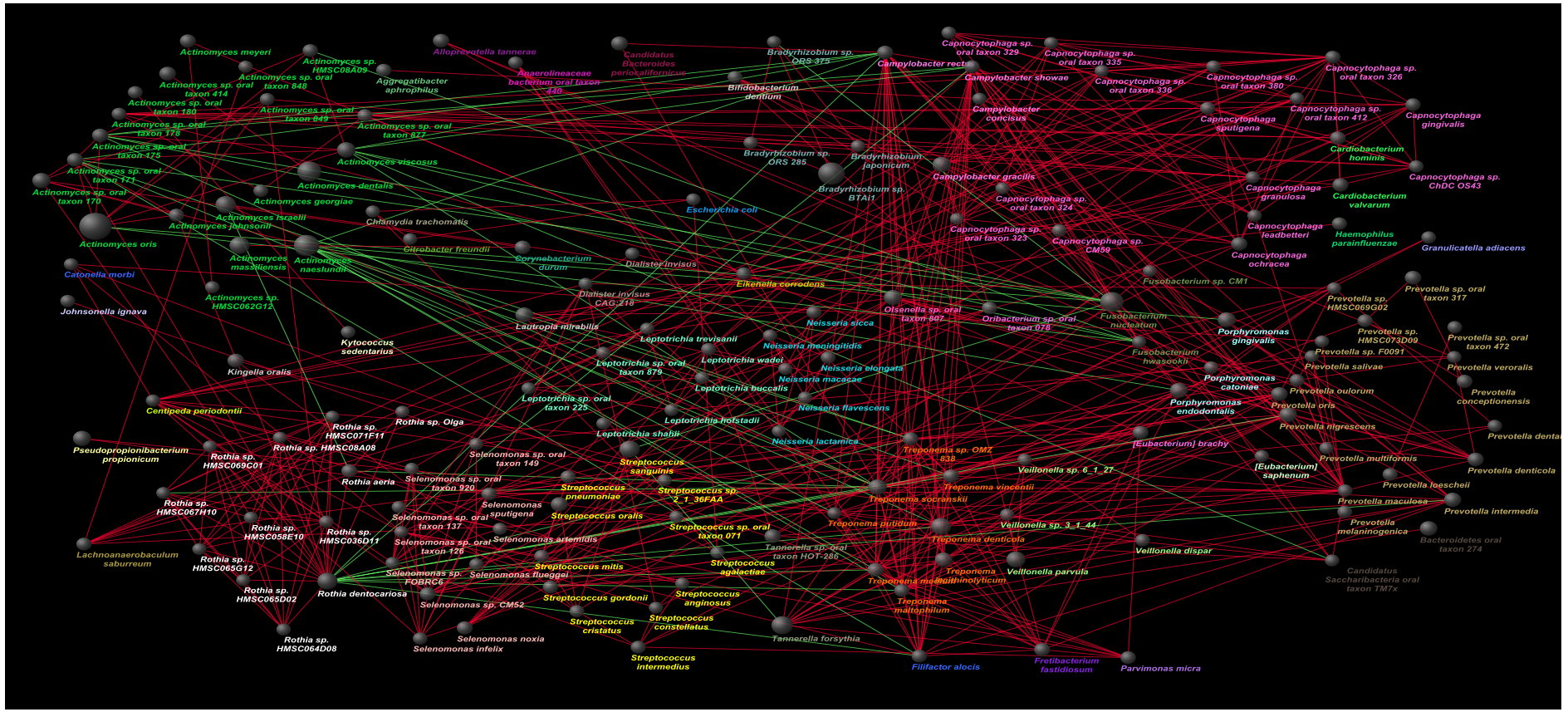


Figure 3.29 Co-occurrence networks of the species in CCP group from periodontally healthy site samples.
Spearman's rank coefficients (q) between species with occurrence in at least 20% of samples in the group were calculated pairwise. Edge stands for a strong ($q > 0.5$ or $q < -0.5$) and significant (adjusted $P < 0.05$) correlation. Red edge is for the positive correlation and green for the negative. The size of each node is proportional to the relative abundance of the species. The species were coloured by the genus-level taxonomy.

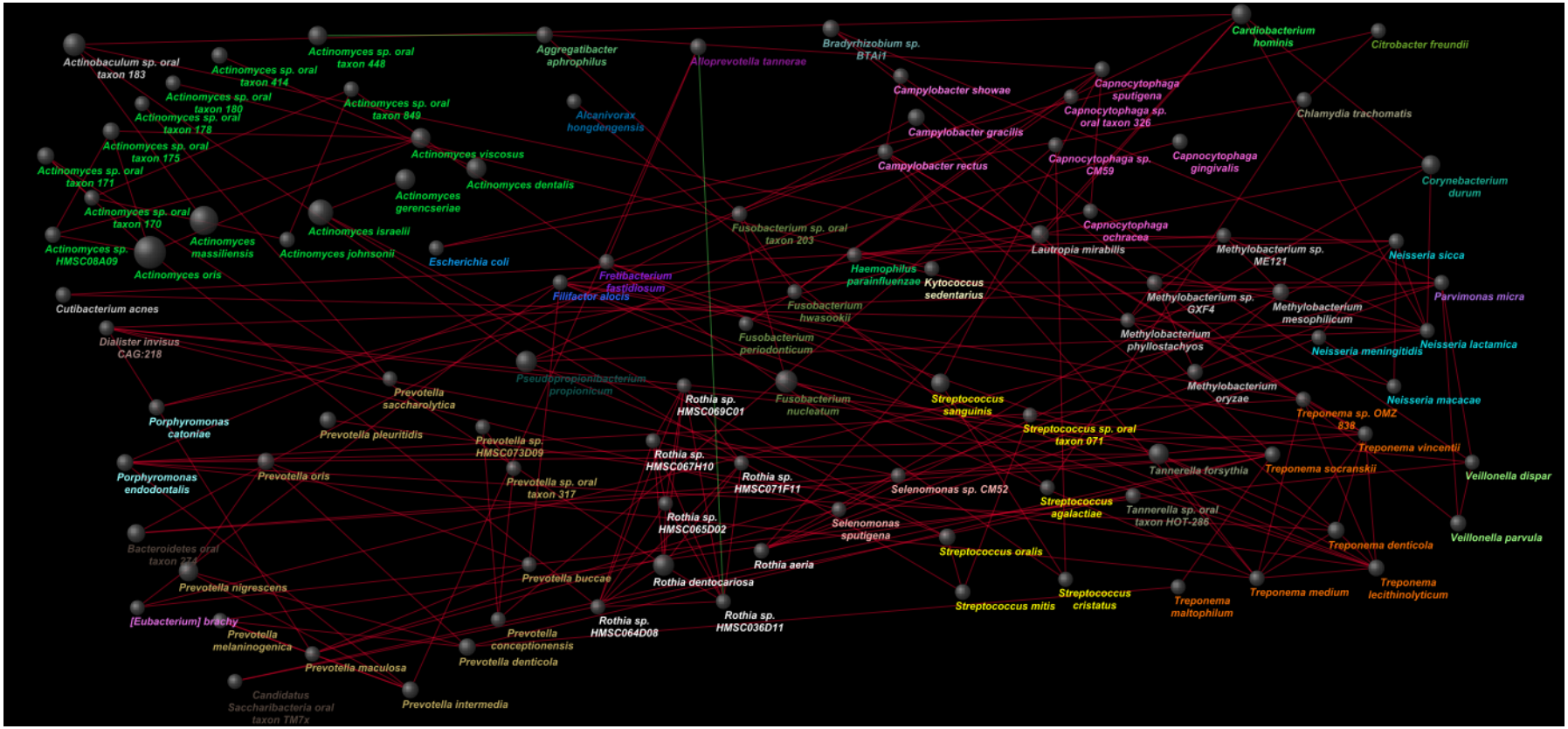


Figure 3.30 Co-occurrence networks of the species in NORA group from periodontally healthy site samples.
Spearman's rank coefficients (q) between species with occurrence in at least 20% of samples in the group were calculated pairwise. Edge stands for a strong ($q > 0.5$ or $q < -0.5$) and significant (adjusted $P < 0.05$) correlation. Red edge is for the positive correlation and green for the negative. The size of each node is proportional to the relative abundance of the species. The species were coloured by the genus-level taxonomy.

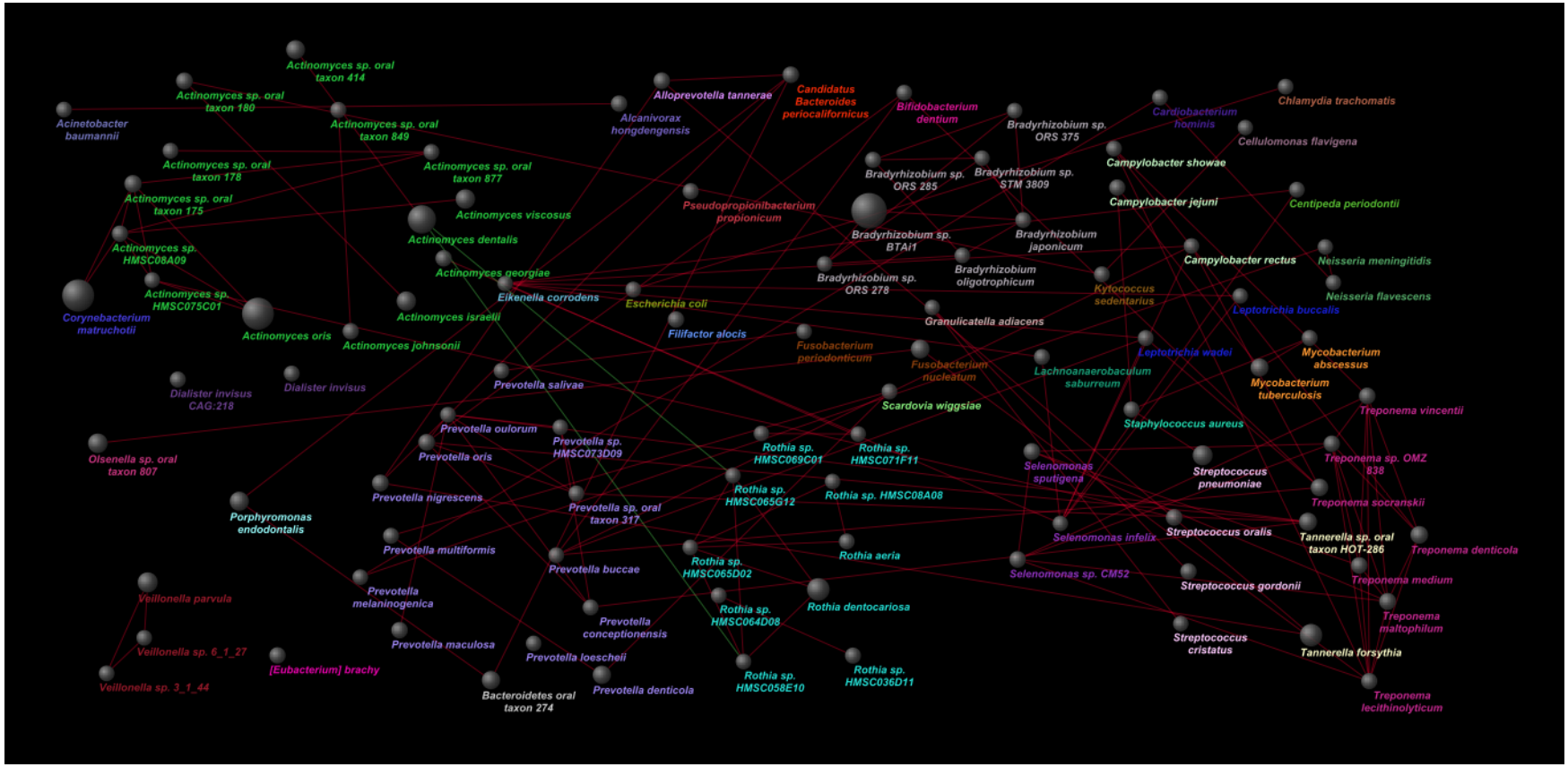


Figure 3.31 Co-occurrence networks of the species in RA group from periodontally healthy site samples.
Spearman's rank coefficients (q) between species with occurrence in at least 20% of samples in the group were calculated pairwise. Edge stands for a strong ($q > 0.5$ or $q < -0.5$) and significant (adjusted $P < 0.05$) correlation. Red edge is for the positive correlation and green for the negative. The size of each node is proportional to the relative abundance of the species. The species were coloured by the genus-level taxonomy.

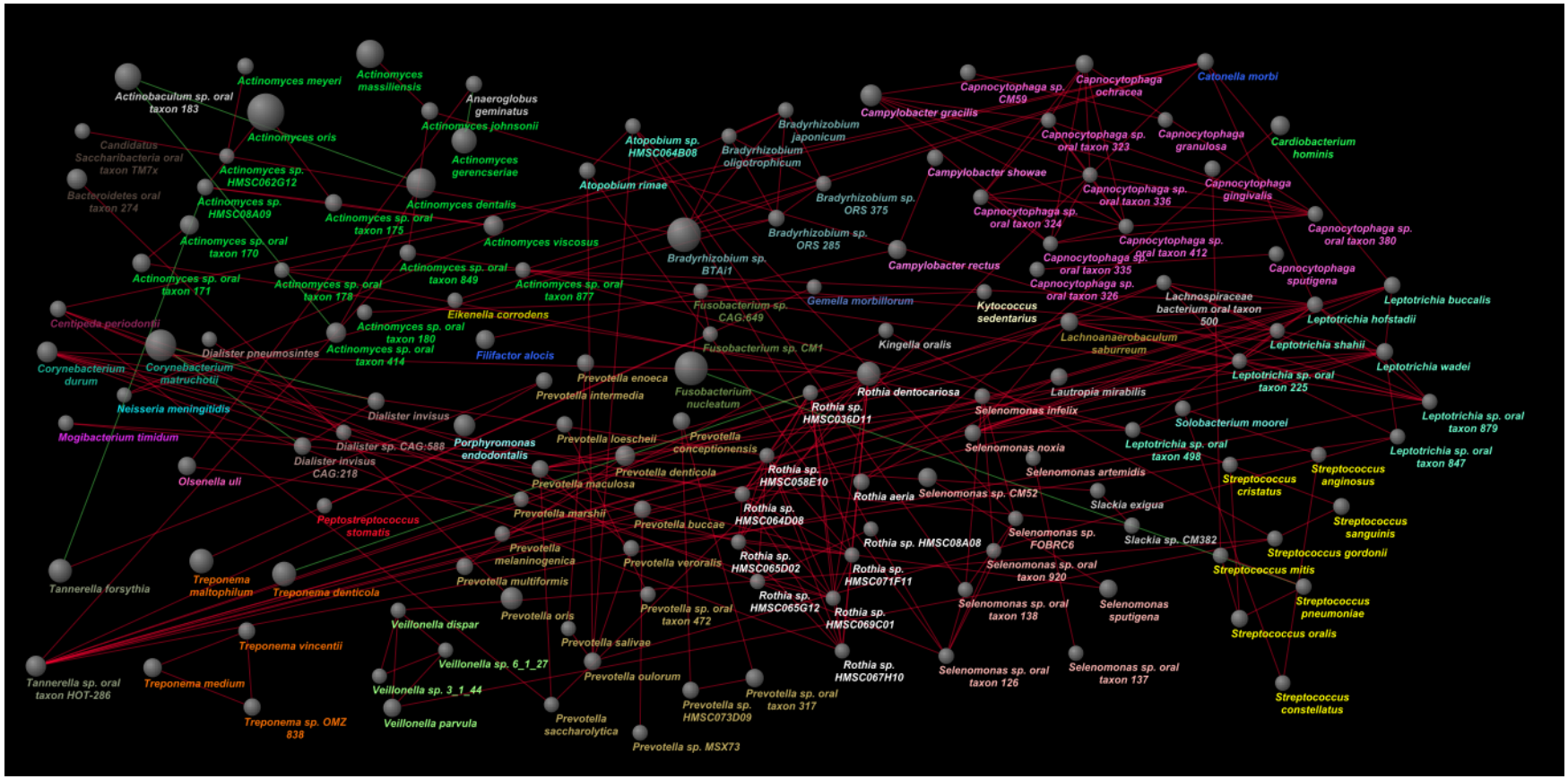


Figure 3.32 Co-occurrence networks of the species in HC group from periodontally diseased site samples.
Spearman's rank coefficients (q) between species with occurrence in at least 20% of samples in the group were calculated pairwise. Edge stands for a strong ($q > 0.5$ or $q < -0.5$) and significant (adjusted $P < 0.05$) correlation. Red edge is for the positive correlation and green for the negative. The size of each node is proportional to the relative abundance of the species. The species were coloured by the genus-level taxonomy.

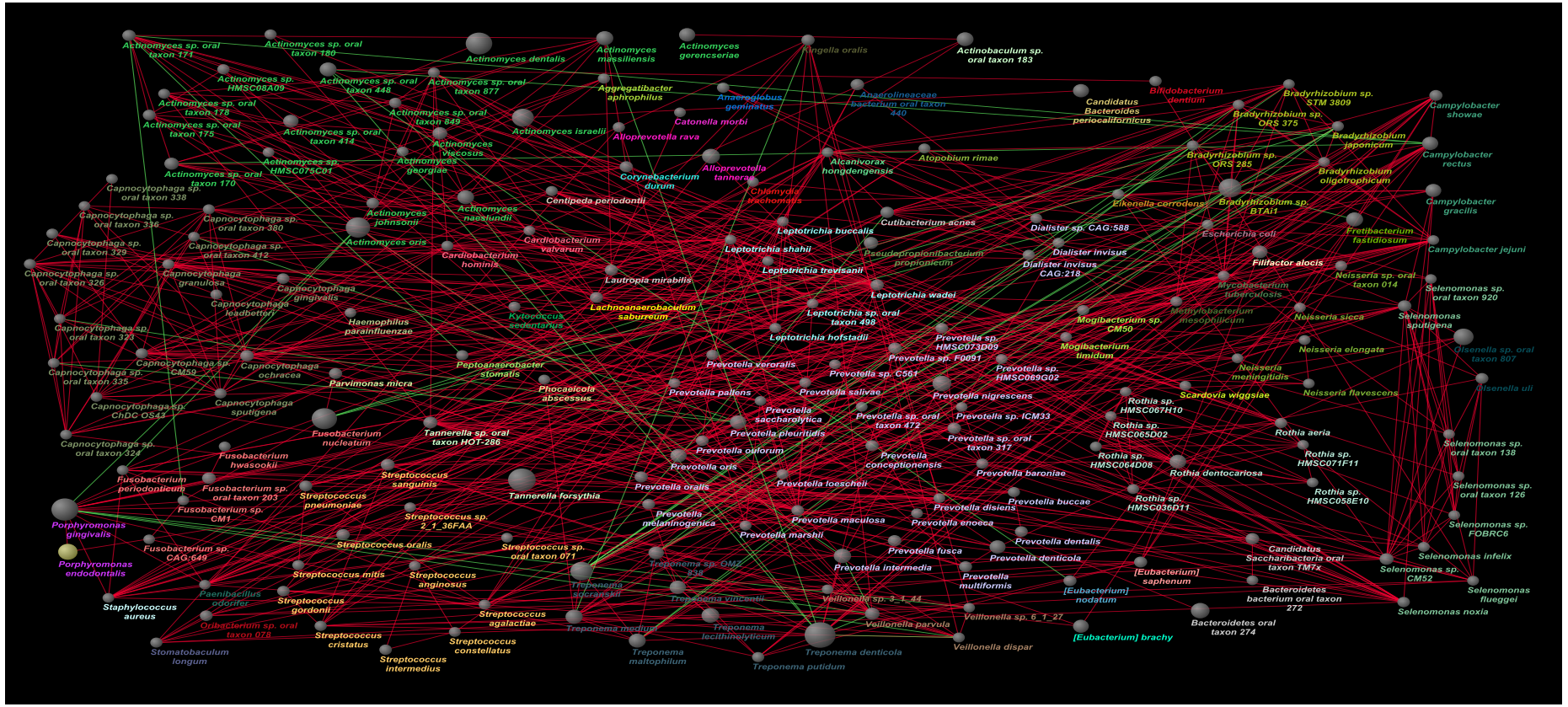


Figure 3.33 Co-occurrence networks of the species in CCP group from periodontally diseased site samples.
Spearman's rank coefficients (q) between species with occurrence in at least 20% of samples in the group were calculated pairwise. Edge stands for a strong ($q > 0.5$ or $q < -0.5$) and significant (adjusted $P < 0.05$) correlation. Red edge is for the positive correlation and green for the negative. The size of each node is proportional to the relative abundance of the species. The species were coloured by the genus-level taxonomy.

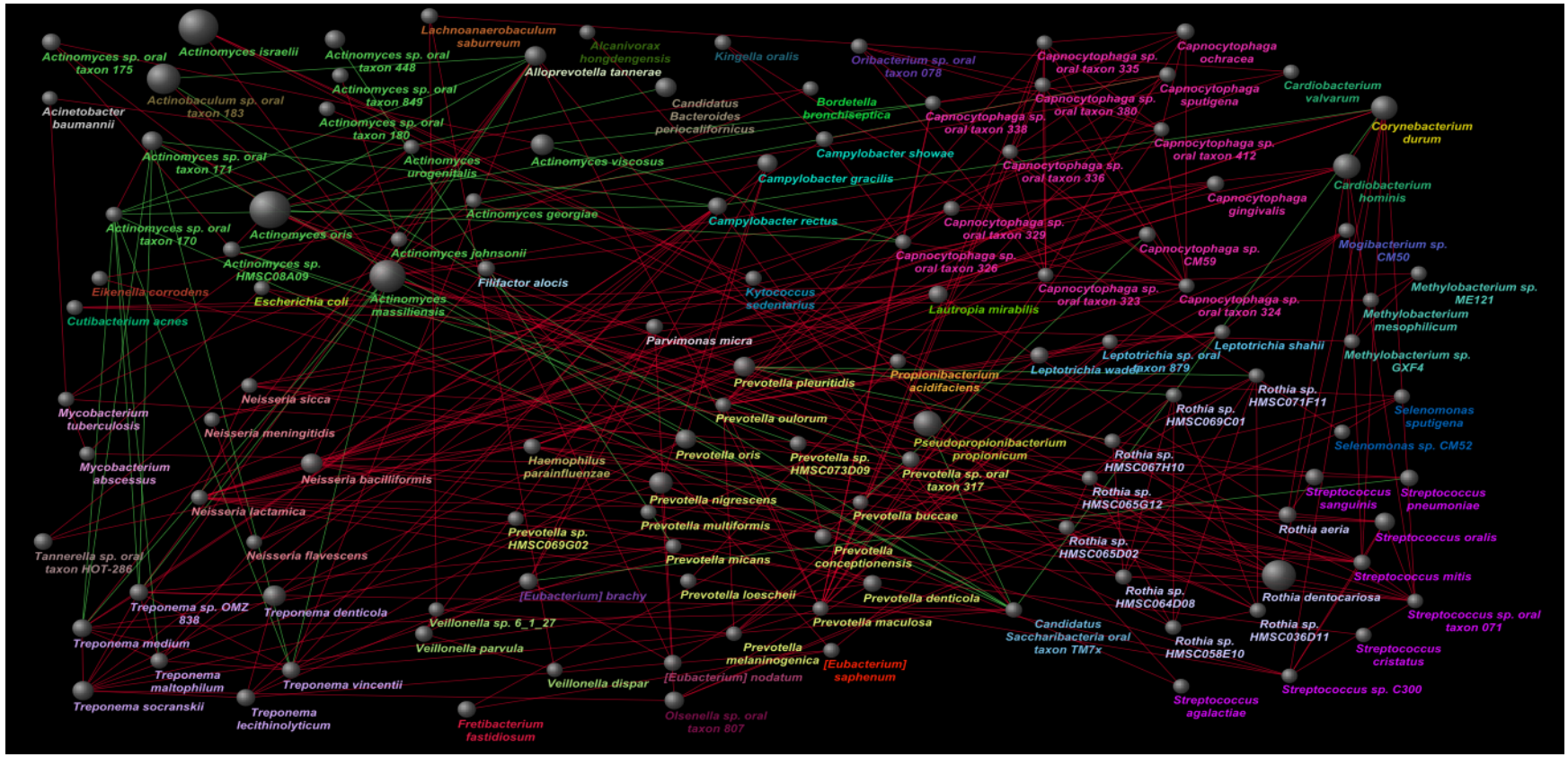


Figure 3.34 Co-occurrence networks of the species in NORA group from periodontally diseased site samples.
Spearman's rank coefficients (q) between species with occurrence in at least 20% of samples in the group were calculated pairwise. Edge stands for a strong ($q > 0.5$ or $q < -0.5$) and significant (adjusted $P < 0.05$) correlation. Red edge is for the positive correlation and green for the negative. The size of each node is proportional to the relative abundance of the species. The species were coloured by the genus-level taxonomy.

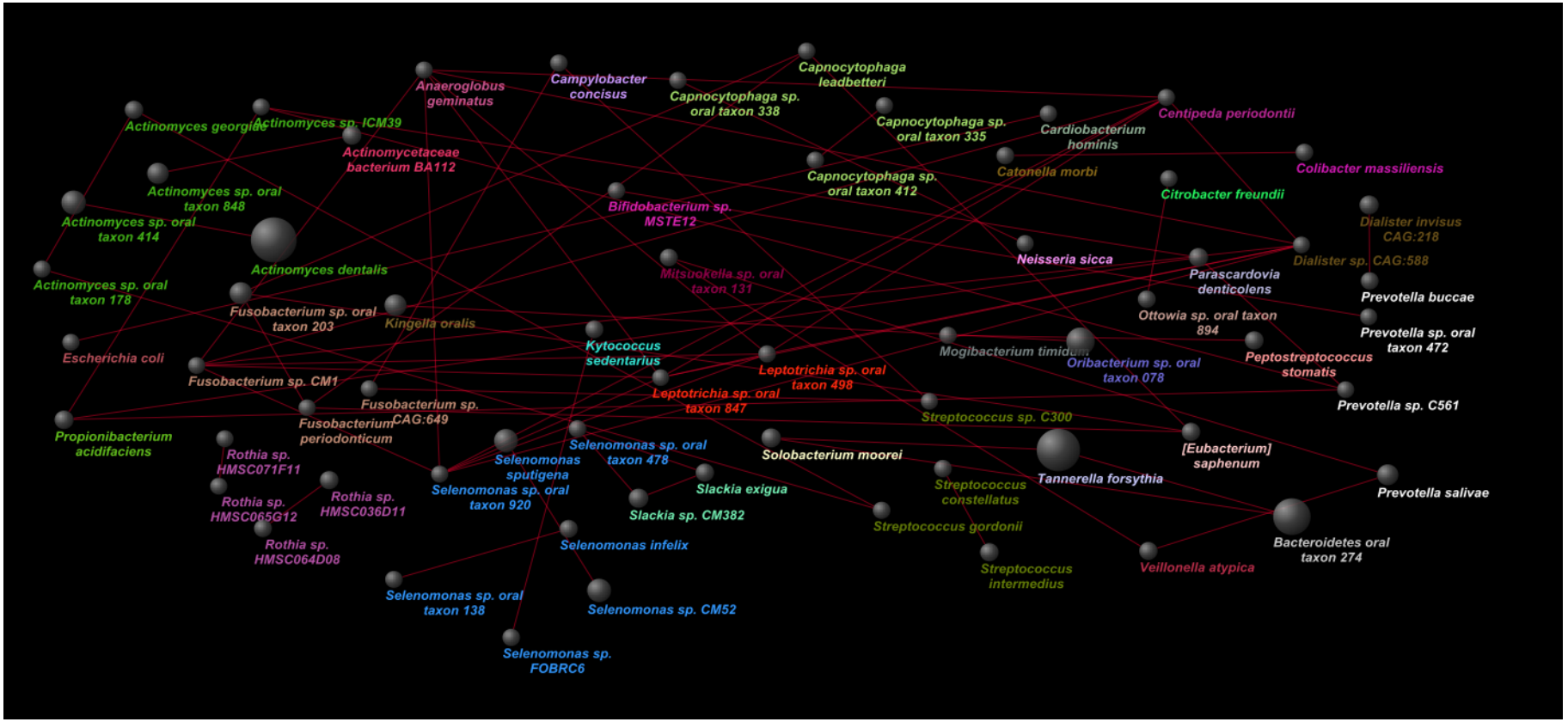


Figure 3.35 Co-occurrence networks of the species in RA group from periodontally diseased site samples.
Spearman's rank coefficients (q) between species with occurrence in at least 20% of samples in the group were calculated pairwise. Edge stands for a strong ($q > 0.5$ or $q < -0.5$) and significant (adjusted $P < 0.05$) correlation. Red edge is for the positive correlation and green for the negative. The size of each node is proportional to the relative abundance of the species. The species were coloured by the genus-level taxonomy.

Table 3.8 Topological properties of co-occurrence networks of species in each group.

Group	E [a] (positive%)	N [b]	Density [c]	CC [d]	APL [e]
Heathy site					
HC	603 (91.9)	163	3.70	0.47	3.65
CCP	571 (92.1)	162	3.52	0.49	3.84
NORA	187 (98.9)	100	1.87	0.51	4.77
RA	130 (98.5)	99	1.31	0.68	2.25
Diseased site					
HC	235 (96.6)	138	1.70	0.68	4.46
CCP	720 (94.4)	188	3.83	0.50	3.86
NORA	296 (87.5)	118	2.51	0.53	4.85
RA	65 (100)	64	1.02	1.00	1.00

[a] E, number of edges;

[b] N, number of nodes;

[c] Density, ratio of edges to nodes;

[d] CC, average clustering coefficient;

[e] APL, average shortest path length.

3.10.9 Core microbiota of each group

The core microbiota, defined as the set of taxa that are detected in a remarkable fraction of the population above a given abundance threshold was identified for each group based on the threshold of 50% prevalence, 0.2% relative abundance. In healthy site samples, 36, 53, 22 and 33 species were described as core bacterial species for HC, CCP, NORA and RA groups, respectively. Similarly, 48, 44, 37 and 48 species were identified as core species for HC, CCP, NORA and RA groups in diseased site samples. By overlap analysis, 17 species were found to be the core species commonly for all groups in periodontally healthy site samples and 25 in diseased site samples (Figure 3.36). There were three, five, one and four species identified as the core species only for HC, CCP, NORA and RA groups respectively, in healthy site samples. Three, two, six and four species were identified as the core species only for HC, CCP, NORA and RA groups respectively, in diseased site samples. Different to the overlap analysis in section 3.10.7, the rare species with low relative abundance were removed from each group in this analysis. The species found to be core species only for one group (species exceeding a given detection threshold in one group but below the threshold in other groups) in healthy site samples or diseased site samples are listed in Table 3.9. The full lists of core species of each group are in Appendix C7 and C8.

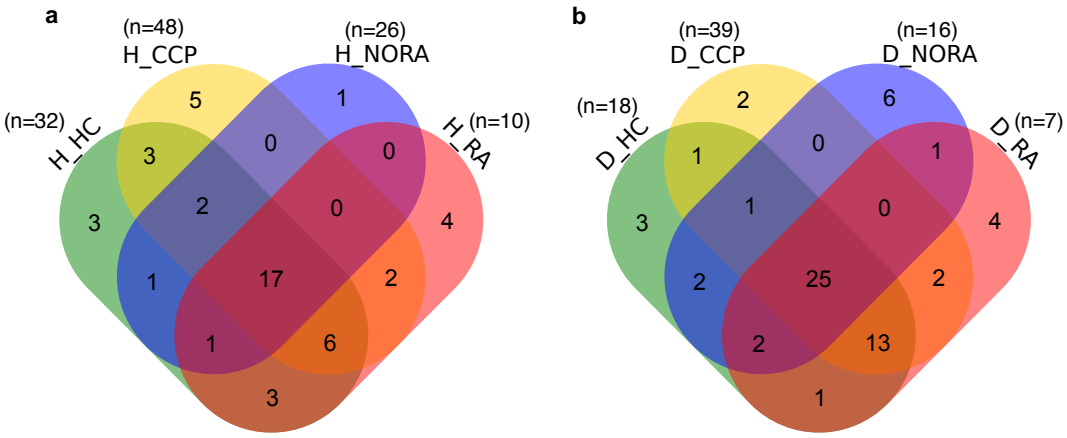


Figure 3.36 Overlap analysis of of the group specific and shared core species.
Core species in each group of healthy site and diseased site samples were identified, respectively (> 50% prevalence, > 0.2% relative abundance). Number of group specific and shared core specie were visulized in (a) healthy site samples; (b) diseased site samples. Group denotation: green: HC, yellow: CCP, blue: NORA, red: RA. n: number of samples.

Table 3.9 List of core species specific for each group in healthy site samples or diseased site samples.

Site	Species	* Core of
Healthy site	*Actinomyces sp. oral taxon 849*	HC
	Prevotella denticola	HC
	Streptococcus mitis	HC
	Campylobacter rectus	CCP
	Campylobacter showae	CCP
	Prevotella sp. oral taxon 317	CCP
	Selenomonas sp. CM52	CCP
	Treponema vincentii	CCP
	Porphyromonas gingivalis	NORA
	Alloprevotella tannerae	RA
	Bradyrhizobium sp. BTAi1	RA
	Candidatus Bacteroides periocalifornicus	RA
	Porphyromonas endodontalis	RA
Diseased site	*Mogibacterium sp. CM50*	HC
	Prevotella conceptionensis	HC
	Prevotella sp. HMSC073D09	HC
	Bradyrhizobium sp. BTAi1	CCP
	Prevotella intermedia	CCP
	Actinomyces sp. HMSC08A09	NORA
	Cardiobacterium hominis	NORA
	Porphyromonas gingivalis	NORA
	Streptococcus mitis	NORA
	Streptococcus pneumoniae	NORA
	Streptococcus sanguinis	NORA
	Kingella oralis	RA

Leptotrichia wadei	RA
Oribacterium sp. oral taxon 078	RA
Selenomonas noxia	RA

***** Core species in each group of healthy site and diseased site samples were identified, respectively (> 50% prevalence, > 0.2% relative abundance).

In order to characterize the functional capability of the microbiome, aligned reads were also assigned to the EggNOG database using the functional identifier mapping as implemented in MEGAN6. Overall, 8253 functional units were identified for healthy site samples and 7964 for diseased sites. Abundances of functional units were normalized and compared between groups using DESeq2 for healthy sites and diseased sites. Twenty-nine functional units were significantly under-represented in the NORA group compared with the CCP group (Table 3.10) (adjusted $P < 0.05$, Wald test, FDR adjusted), out of which 14 were involved in metabolism, seven in information storage and processing, and eight in cellular processes and signalling. The three most under-represented functional units (log2FoldChange < -1) were COG0399, COG3842 and COG0402, which are involved in amino acid and nucleotide transport and metabolism. No significant difference was found in diseased site samples. Those functional units with significant differences were mapped to Kyoto Encyclopaedia of Genes and Genomes (KEGG) metabolic pathway maps by using Interactive Pathways Explorer (iPath3) for visualization (Figure 3.37). Pathways are sized by the log fold changes of the functional units calculated by DESeq2.

Table 3.10 Functional units that were significantly under-represented in the NORA group compared with the CCP group in periodontally healthy site samples (adjusted $P < 0.05$, Wald test, FDR adjusted).

Functional category	EggNOG functional unit	log2FoldChange
[E] Amino acid transport and metabolism	COG0399 DegT DnrJ EryC1 StrS aminotransferase	-1.25
	COG3842 Part of the ABC transporter complex PotABCD involved in spermidine putrescine import. Responsible for energy coupling to the transport system (By similarity)	-1.21
	COG0028 acetolactate synthase	-0.68
[F] Nucleotide transport and metabolism	COG0402 deaminase	-1.12
[G] Carbohydrate transport and metabolism	COG0395 Binding-protein-dependent transport systems inner membrane component	-0.98
	COG1621 Hydrolase	-0.79
	COG1653 transporter activity	-0.76
	COG2017 converts alpha-aldose to the beta-anomer. It is active on D-glucose, L-arabinose, D-xylose, D-galactose, maltose and lactose (By similarity)	-0.97
[H] Coenzyme transport and metabolism	COG1154 Catalyzes the acyloin condensation reaction between C atoms 2 and 3 of pyruvate and glyceraldehyde 3-phosphate to yield 1-deoxy-D-xylulose-5-phosphate (DXP) (By similarity)	-0.82
[I] Lipid transport and metabolism	COG0304 Catalyzes the condensation reaction of fatty acid synthesis by the addition to an acyl acceptor of two carbons from malonyl-ACP (By similarity)	-0.78
[J] Translation, ribosomal structure and biogenesis	COG0024 Removes the N-terminal methionine from nascent proteins (By similarity)	-0.81
	COG0532 One of the essential components for the initiation of protein synthesis. Protects formylmethionyl-tRNA from spontaneous hydrolysis and promotes its binding to the 30S ribosomal subunits. Also involved in the hydrolysis of GTP during the formation of the 70S ribosomal complex (By similarity)	-0.70

	COG0621 Catalyzes the methylthiolation of N6- (dimethylallyl)adenosine (i(6)A), leading to the formation of 2- methylthio-N6- (dimethylallyl)adenosine (ms(2)i(6)A) at position 37 in tRNAs that read codons beginning with uridine (By similarity)	-0.78
[K] Transcription	COG1309 Transcriptional regulator	-1.00
[L] Replication, recombination and repair	COG0210 helicase	-0.69
	COG0322 The UvrABC repair system catalyzes the recognition and processing of DNA lesions. UvrC both incises the 5' and 3' sides of the lesion. The N-terminal half is responsible for the 3' incision and the C-terminal half is responsible for the 5' incision (By similarity)	-0.77
	COG0587 DNA polymerase III (alpha subunit)	-0.63
[M] Cell wall/membrane/envelope biogenesis	COG0859 heptosyltransferase	-1.08
	COG0463 Glycosyl transferase, family 2	-0.75
[O] Posttranslational modification, protein turnover, chaperones	COG1404 peptidase (S8 and S53, subtilisin, kexin, sedolisin	-0.70
[P] Inorganic ion transport and metabolism	COG0601 Binding-protein-dependent transport systems inner membrane component	-0.86
	COG0370 Ferrous iron transport protein b	-0.82
	COG0841 acriflavin resistance protein	-0.90
	COG0168 Low-affinity potassium transport system. Interacts with Trk system potassium uptake protein TrkA (By similarity)	-0.93
[T] Signal transduction mechanisms	ENOG410XNMH Histidine kinase	-0.97
	COG0664 transcriptional regulator, crp fnr family	-0.75
	COG0745 regulatoR	-0.70
	COG1217 gtp-binding protein typa	-0.71
[V] Defense mechanisms	COG1680 Beta-lactamase	-0.96

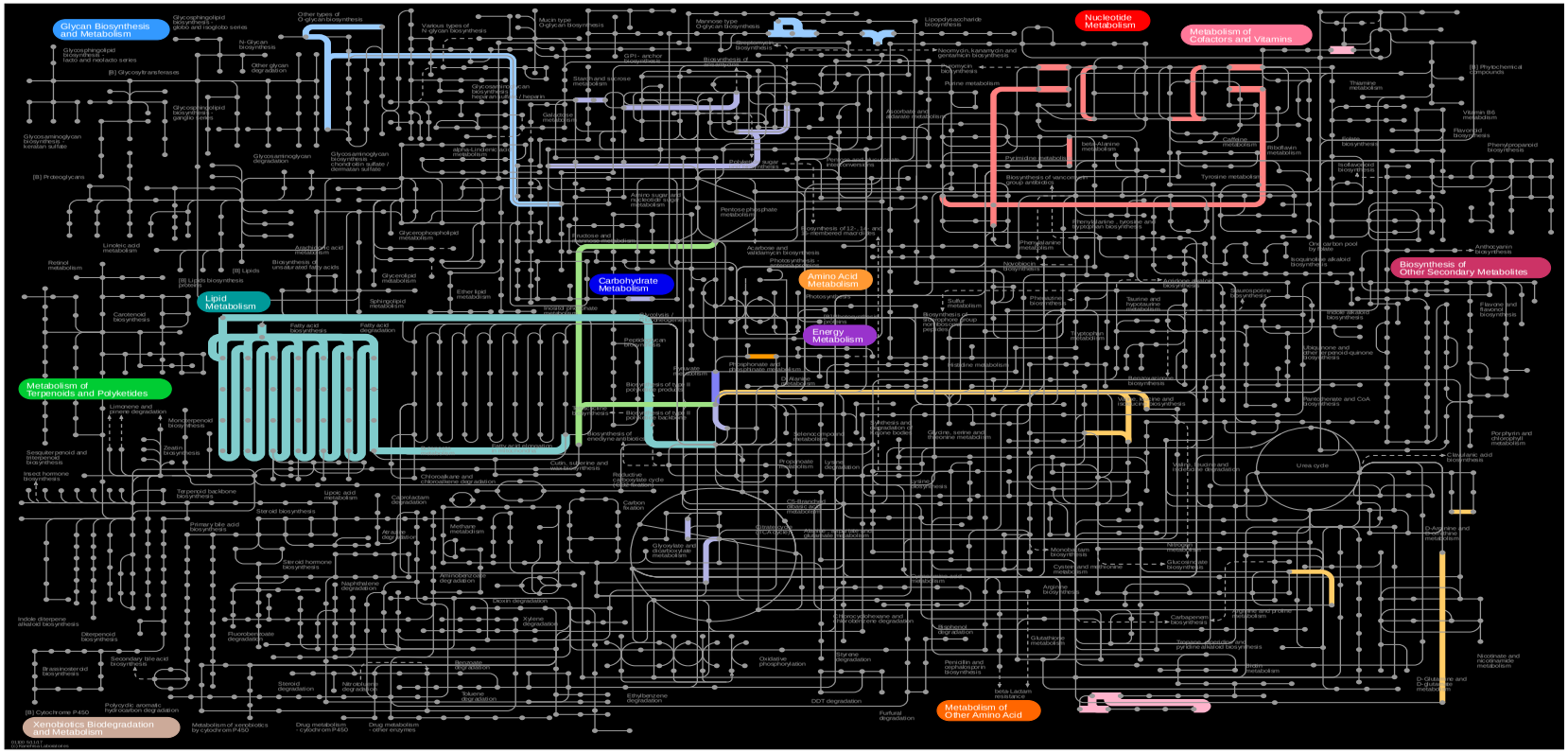


Figure 3.37 Metabolic pathway maps of significantly different functional units between the CCP and NORA groups in healthy site samples.

Mapped pathways are highlighted by different colours and sized by the log fold changes of the functional unit that have been identified as significantly different by Wald test using DEseq2. The map was generated using Interactive Pathways Explorer (iPath3).

3.10.10 Analysis of periodontally healthy site samples from individuals without periodontitis

To avoid the potential influence of periodontitis on the subgingival microbiome, I only used the samples from healthy sites from individuals without periodontitis to identify the possible impact of RA status. The α-diversity was compared between groups and no significant difference was found either by observed species or Shannon index (ANOVA, $P > 0.05$) (Figure 3.38). β-diversity at species level was determined by Bray-Curtis dissimilarity based on relative abundance after taxa filtering (prevalence equal or more than three, total counts more than ten). The results were plotted using PCoA. PERMANOVA statistical tests (*Adonis* function, vegan package, R) showed a significantly different β-diversity between groups ($P = 0.0039$, $R^2 = 0.078$) (Figure 3.39). Homogeneity of multivariate dispersions was tested by using the *betadisper* function to ensure PERMANOVA assumptions (significant result is not due to differences in group dispersions).

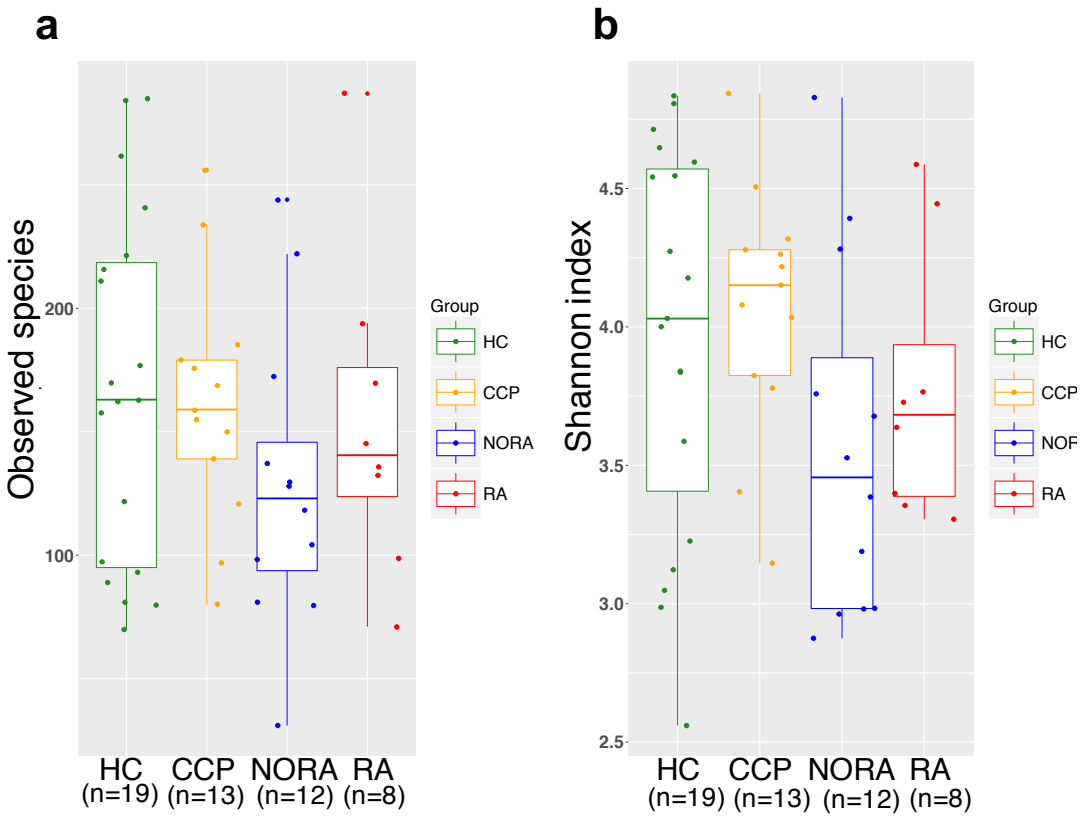


Figure 3.38 Comparison of α-diversity in the different groups using healthy site samples from individuals without periodontitis.
No significant difference was found between groups either in (a) Observed species or (b) Shannon index (ANOVA, $P > 0.05$). n: number of samples.

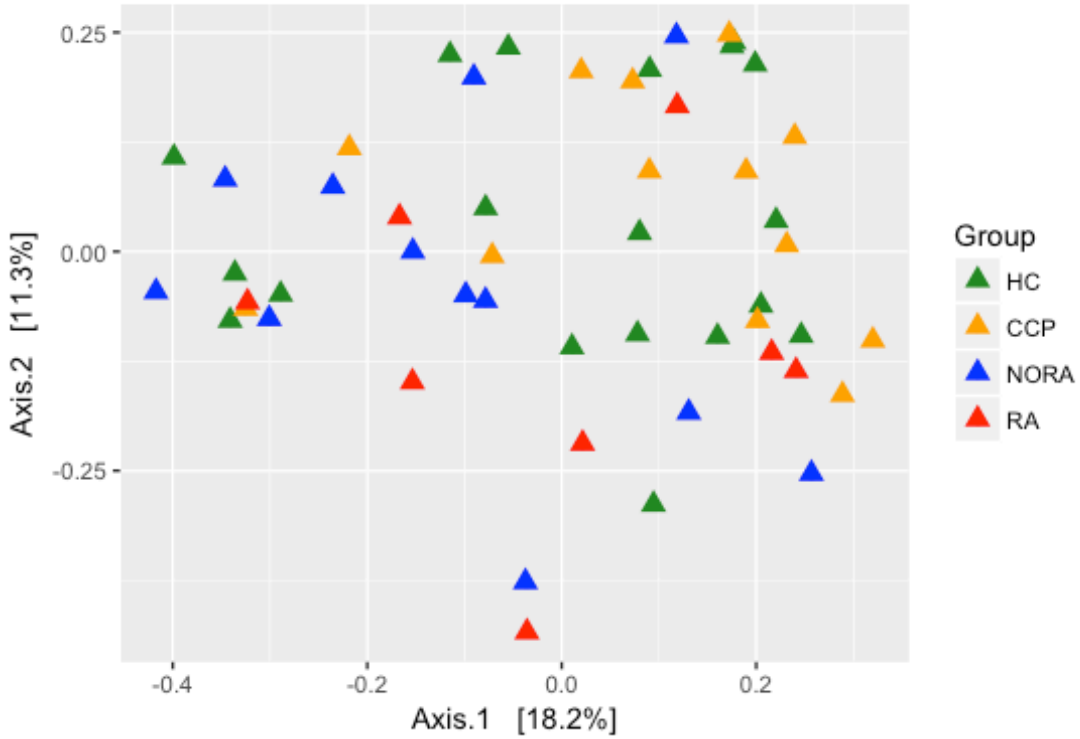


Figure 3.39 β-diversity in healthy site samples from individuals without periodontitis.
β-diversity was determined by Bray-Curtis dissimilarity and plotted using PCoA. Raw data were filtered to remove taxa accounting for less than ten sequences in total and observed in less than three samples. A significantly different β-diversity was found between groups ($P = 0.0039$, $R^2 = 0.078$) by PERMANOVA test.

At phylum level (Figure 3.40), only *Spirochaetes* were found with significantly higher relative abundance in the CCP group compared with other groups, this was also the case when I tested healthy site samples regardless of periodontal disease status ($P < 0.05$, permutation test, one-sided *signassoc* function, *indicspecies* R-package). Within the top 20 most abundant genera (Figure 3.41) in healthy site samples from individuals without periodontitis, *Methylobacterium* exhibited significantly higher relative abundance in NORA patients and *Bradyrhizobium* was more enriched in the RA group compared with other groups ($P < 0.05$). Other less abundant genera were also significantly different when stratified by RA status (Appendix C9). Relative abundance of *Mogibacterium* was significantly higher in the CCP group compared with other groups. *Cardiobacterium* and *Sphingomonas*

were higher in NORA patients. Relative abundance of *Staphylococcus, Bacillus, Escherichia, Achromobacter, Delftia* and *Stenotrophomonas* were higher in the RA group compared to other groups ($P < 0.05$).

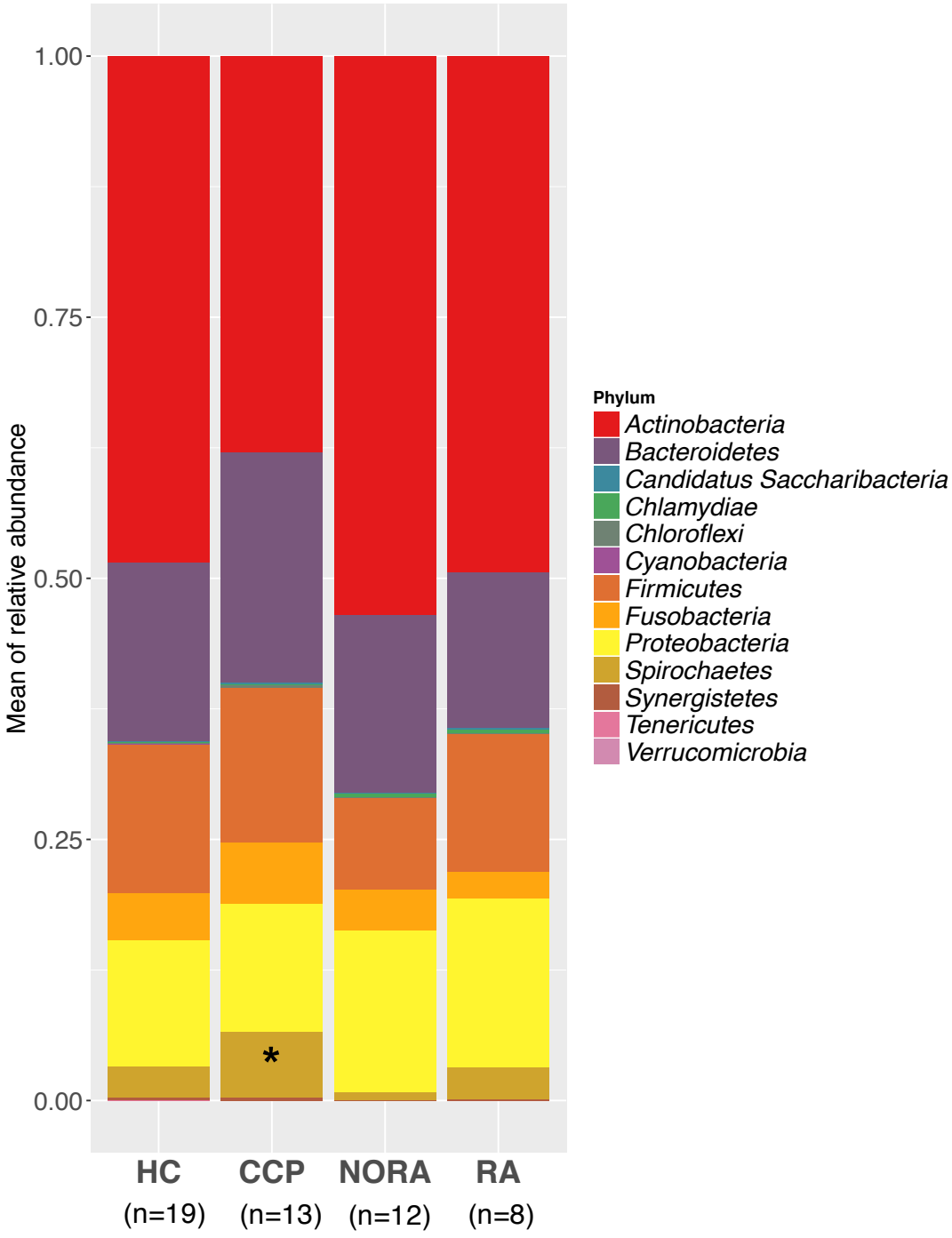


Figure 3.40 Phylum composition of different groups in periodontally healthy site samples from individuals without periodontitis.
Phyla with significantly higher relative abundance were determined using the permutation test (one-sided *signassoc* function, *indicspecies* R-package) between groups. *: corrected $P < 0.05$ (Sidak's correction). n: number of samples.

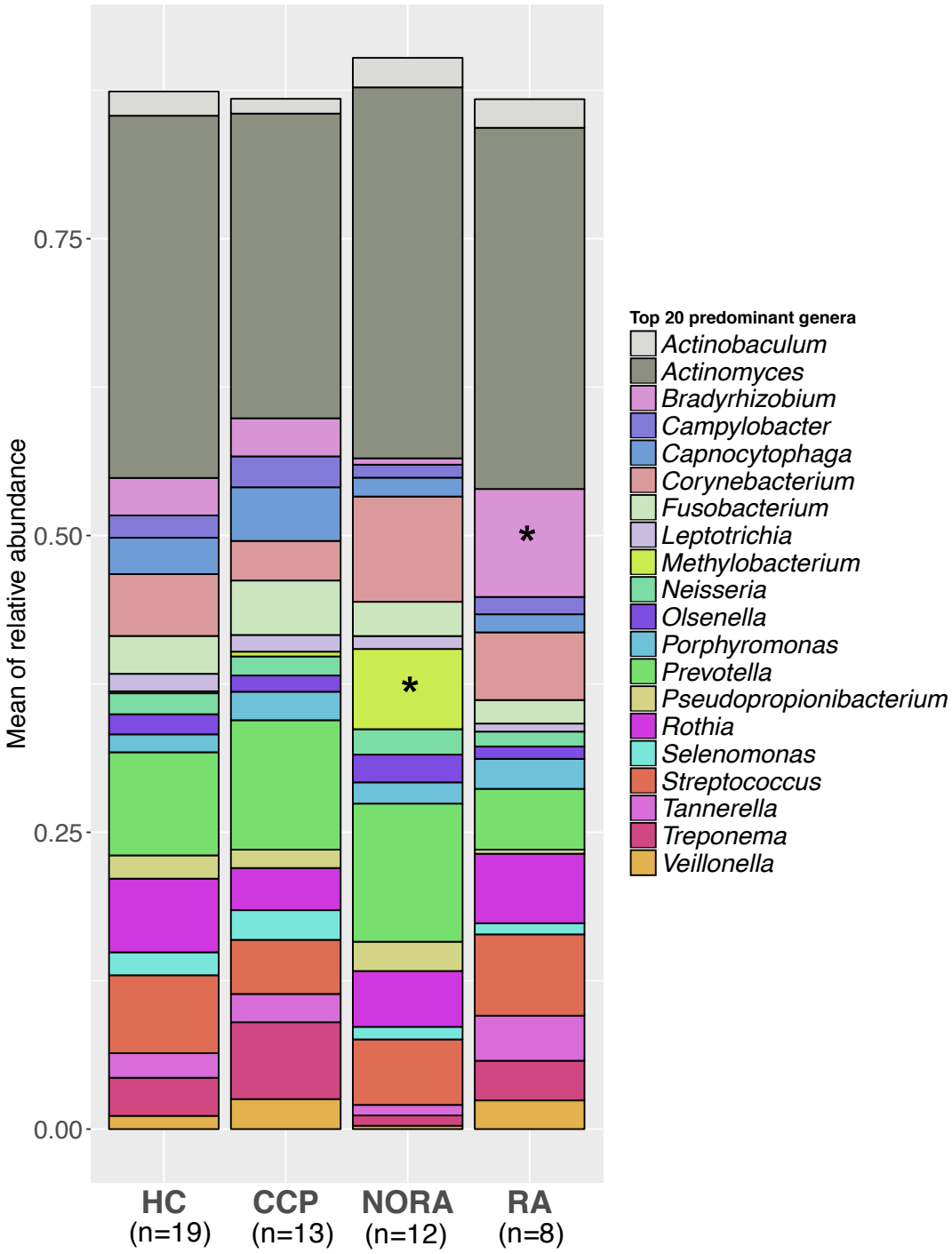


Figure 3.41 Taxonomic profiles for the 20 most abundant genera in healthy site samples from individuals without periodontitis.
Genera with significantly higher relative abundance were determined using the permutation test (one-sided *signassoc* function, *indicspecies* R-package) between groups. *: corrected $P < 0.05$ (Sidak's correction). n: number of samples.

At species level, 26 species were identified with significantly different relative abundance between groups, out of which only one species, *Leptotrichia wadei,* was

associated with the HC group ($P < 0.05$) (Figure 3.42). Fourteen species were significantly associated with the CCP group. Among these species, six *Capnocytophaga* species, three *Prevotella* species, *Leptotrichia sp.* oral taxon 879 and *Treponema putidum* were identified as associated with the CCP group when I tested healthy sites samples regardless of periodontal disease status. *F. nucleatum*, *Fusobacterium sp. HMSC064B12* and *T. socranskii* only emerged in this test based on the restricted criteria of individuals without periodontitis. Among four species significantly associated with NORA, only *Methylobacterium oryzae* was new in this test. Seven species were identified for the RA group, of which five species were restricted to this test.

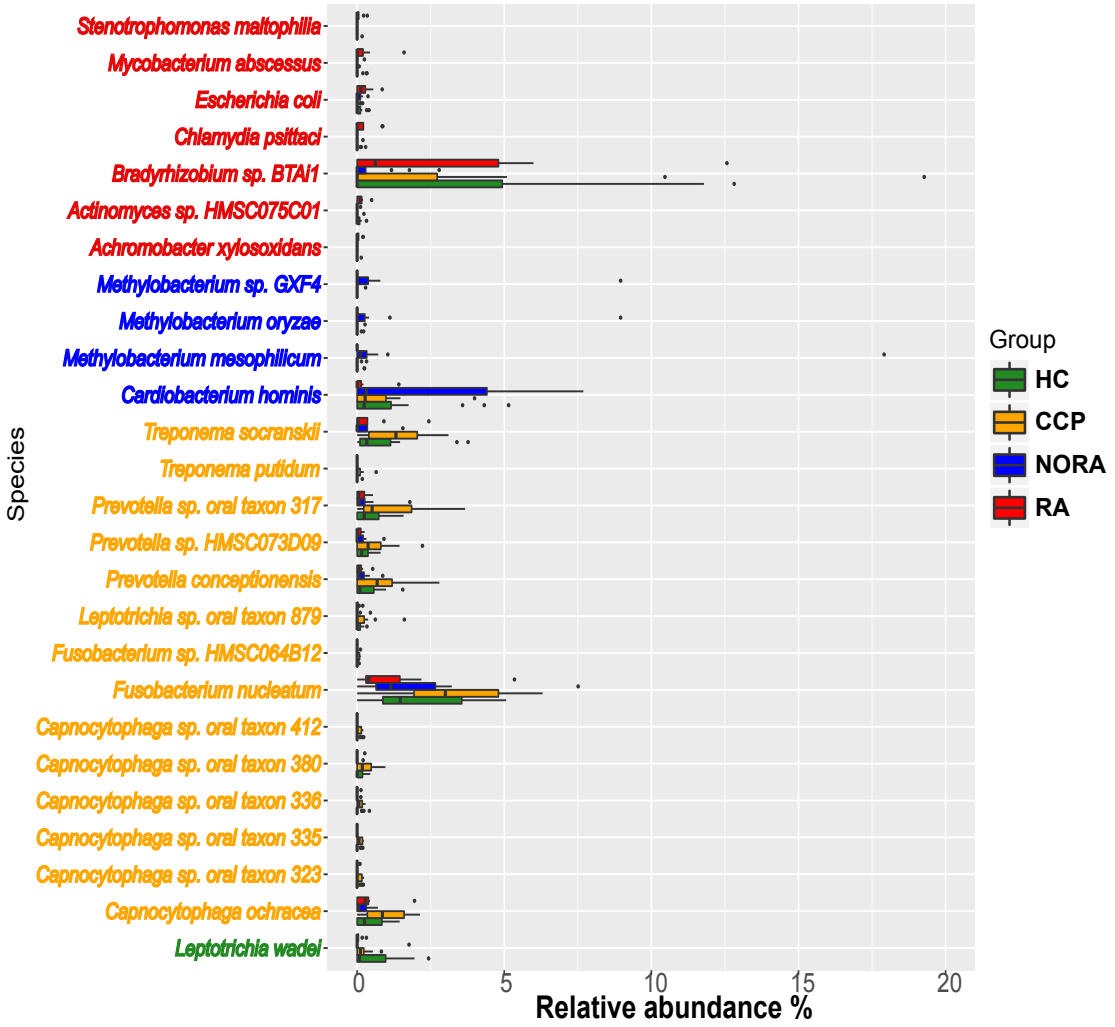


Figure 3.42 Bacterial species with significantly higher relative abundance in HC, CCP, NORA and RA groups in periodontally healthy site samples from individuals without periodontitis.
Species-group association was tested by using one-sided *signassoc* function (indicspecies R-package) based on the relative abundaces. Sidak's correction was applied for multiple testing. Species with significantly higher relative abundances was selected for each group. Group denotation: green: HC, yellow: CCP, blue: NORA, red: RA.

By overlap analysis on species level, 118 species were found to be common for all groups in the in periodontally healthy site samples from individuals without periodontitis. There were 23, 60, 33, and 30 species found to be unique for HC, CCP, NORA and RA group, respectively (Figure 3.43, Appendix C10).

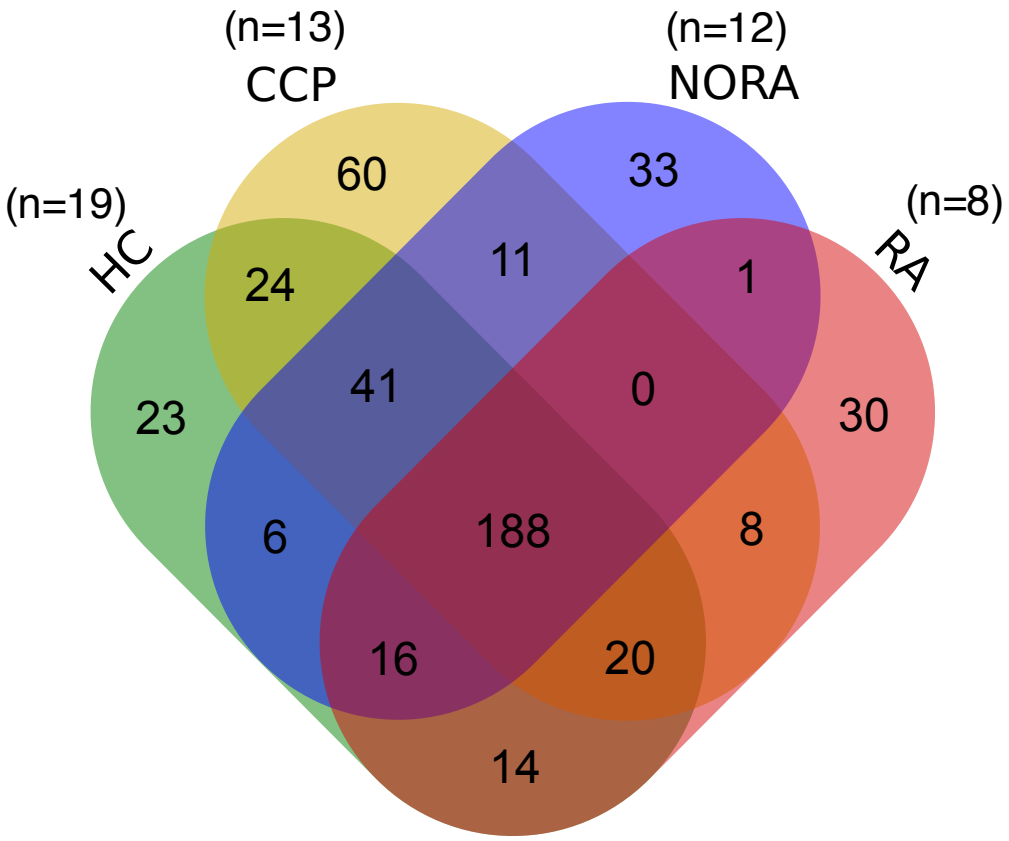


Figure 3.43 Overlap analysis of group specific and shared species in periodontally healthy sites samples from individuals without periodontitis. Group denotation: green: HC, yellow: CCP, blue: NORA, red: RA. n: number of samples.

A core microbiota of each group was identified based on the threshold of 50% prevalence, 0.2% relative abundance. By overlap analysis, specific and shared core species were identified for each group (Figure 3.44, Table 3.11).

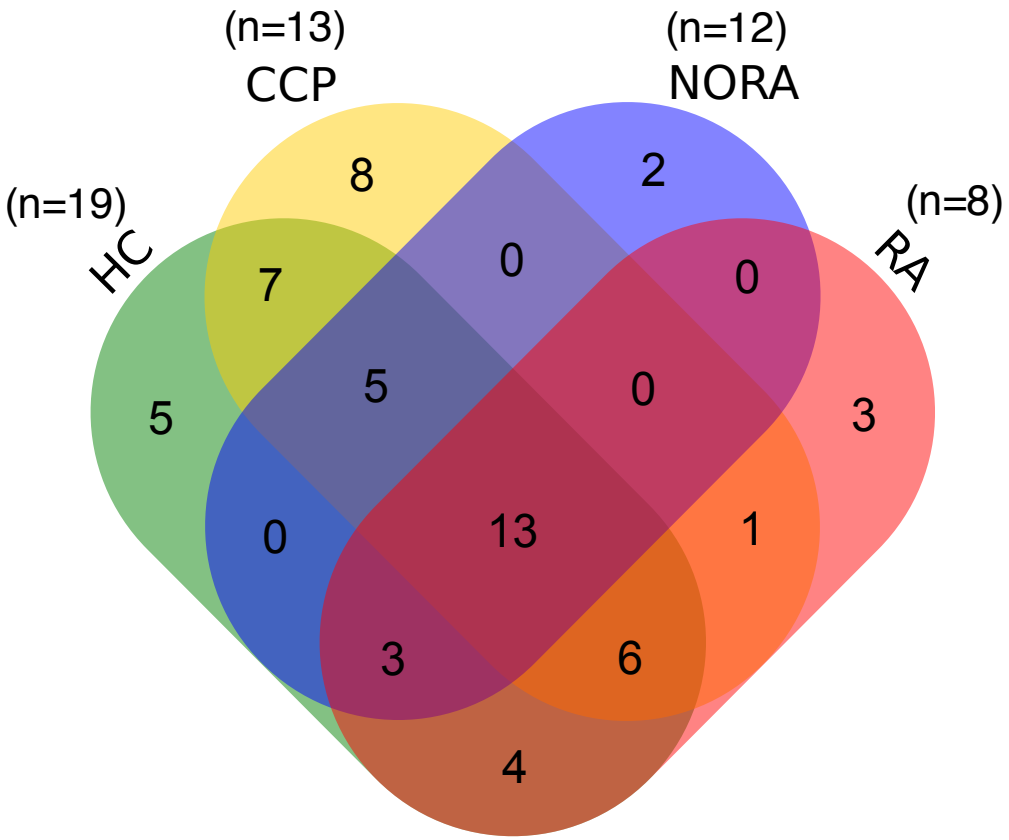


Figure 3.44 Overlap analysis of group specific and shared core species. Core species in each group of periodontally healthy site samples from non-periodontitis individuals were identified, respectively (> 50% prevalence, > 0.2% relative abundance). Group denotation: green: HC, yellow: CCP, blue: NORA, red: RA. n: number of samples.

Table 3.11 List of core species specific for each group in periodontally healthy site samples from individuals without periodontitis (> 50% prevalence, > 0.2% relative abundance).

Group	* Species
HC	*Prevotella denticola*
	Corynebacterium durum
	Actinomyces meyeri
	Prevotella pleuritidis
	Rothia aeria
CCP	*Campylobacter showae*
	Campylobacter rectus
	Prevotella sp. HMSC073D09
	Selenomonas sputigena
	Treponema denticola
	Selenomonas sp. CM52
	Veillonella dispar
	Prevotella conceptionensis
NORA	*Actinomyces sp. HMSC08A09*
	Porphyromonas gingivalis
RA	*Bradyrhizobium sp. BTAi1*
	Alloprevotella tannerae
	Porphyromonas endodontalis

*: Core species in each group was identified based on the threshold of 50% prevalence, 0.2% relative abundance.

3.10.11 Detection of PAD in the subgingival plaque samples using the shotgun sequencing data.

By aligning the DNA sequencing reads to NCBI NR protein database with at least 70% identity, PAD from different bacterial species as described in NCBI database were identified in my sequencing data (Table 3.12). Apart from *P. gingivalis*, PAD from *Campylobacter concisus, Porphyromonas gulae*, *Prevotella melaninogenica* and *Prevotella veroralis* were also found. There were also sequences matched with hypothetical proteins related to PAD, PAD-like proteins and human PAD

Table 3.12 Annotated PAD and related proteins in subgingival plaque samples based on NCBI protein database.

Protein name	Organism	aa length	Accession number
PAD	*Porphyromonas gingivalis*	556	AKV57251.1
PAD	*Porphyromonas gingivalis*	556	WP_061156921.1
PAD	*Porphyromonas gingivalis*	556	WP_077112131.1
PAD	*Porphyromonas gingivalis*	556	WP_004585430.1
PAD	*Porphyromonas gingivalis*	556	WP_005873463.1
PAD	*Porphyromonas gingivalis*	556	WP_013816581.1
PAD	*Porphyromonas gingivalis*	556	WP_021662880.1
PAD	*Porphyromonas gingivalis*	556	WP_021665004.1
PAD	*Porphyromonas gingivalis*	556	WP_023847729.1
PAD	*Porphyromonas gingivalis*	556	WP_053444041.1
PAD	*Campylobacter concisus*	265	WP_021084474.1
PAD	*Porphyromonas gulae*	556	WP_018964406.1
PAD	*Prevotella melaninogenica*	511	WP_013265317.1
PAD	*Prevotella veroralis*	512	WP_004383428.1

hypothetical protein A343_2192; PAD_porph; pfam04371	*Porphyromonas gingivalis JCVI SC001*	353	EOA10175.1
hypothetical protein HMPREF0649_00225; Porphyromonas-type PAD; pfam04371	*Prevotella buccae D17*	353	EFC76913.1
PAD-like protein	*Homo sapiens*	694	AAS07634.1
PAD family protein	*Campylobacter concisus UNSW2*	223	ERJ31051.1
PAD family protein	*Prevotella oris C735*	349	EFI48643.1
PAD family protein	*Prevotella sp. CAG:891*	378	CDE86775.1
PAD type 6	*Homo sapiens*	694	AAR38850.1
PAD type I	*Homo sapiens*	663	BAA85771.1
PAD-like enzyme	*Porphyromonas gingivalis*	353	AKV63383.1
protein-arginine deiminase type-1	*Homo sapiens*	663	NP_037490.2
protein-arginine deiminase type-3	*Homo sapiens*	664	NP_057317.2
protein-arginine deiminase type-4	*Homo sapiens*	663	NP_036519.2
putative arginine deiminase	*Porphyromonas gingivalis TDC60*	312	BAK24590.1

3.11 Potential citrullination activity of PAD/PAD-like enzyme in *Prevotella* species

To further examine the presence of PAD or PAD-like enzyme in the bacterial species other than *P. gingivalis*, the citrullination activity was measured in the cells of a range of *Prevotella* spp. using colorimetric assays with different substrates and with varying calcium concentrations. The synthetic arginine substrate BAEE, which is a standard substrate for both PPAD and human PAD, was used in the presence or absence of calcium to detect any enzyme activity of citrullination and assess the dependence of any enzyme activity upon calcium ions. The activity in *P. gingivalis* W83, its Δ*ppad* mutant, *P. gingivalis* W50 and its Δ*rgps* mutant were also measured, and the data included for comparative purposes. Although the levels of activity were low, some citrullination was observed in all *Prevotella* spp. However, only the genome of *P. intermedia* ATCC25611 (https://www.ncbi.nlm.nih.gov/protein/1132728393) and *P. melaninogenic* NCTC 12963 (https://www.ncbi.nlm.nih.gov/protein/WP_013265317.1) are available on NCBI and these contain sequences annotated as PAD. Citrullination activities in the cellular samples of *P. gingivalis* W83 Δ*ppad* mutant and *Prevotella* spp. were observed only in the assay using buffer without calcium (Figure 3.45) and they were significantly lower than those of *P. gingivalis* W83, *P. gingivalis* W50 and its Δ*rgps* mutant (ANOVA test). No significant difference was found comparing the activities of the W83 Δ*ppad* mutant and *Prevotella* spp. Integrated with the findings of PAD or PAD-like enzyme sequences in *P. gingivalis* and a range of *Prevotella* spp. as described in the Results section 3.10.11, these citrullination activities indicate that *Prevotella* spp. can express a type of active PAD or PAD-like enzyme, albeit with a low level of activity compared with PPAD, and *P. gingivalis* also harbours a type of

PAD or PAD-like enzyme which provides a low level of citrullination activity even when the PPAD activity is depleted by mutation. Higher activity was observed in the reaction without calcium compared with that containing calcium for *P. gingivalis* W83, *P. nigrescens* OMZ227 and *P. intermedia* OMZ326 ($P < 0.05$, t test), while other differences observed did not reach the statistical significance (Figure 3.45) ($P > 0.05$, t test). This result suggests that unlike human PAD, the potential PAD or PAD-like enzyme in these bacterial species is more like PPAD, in that the enzyme activity is not dependent on the presence of calcium. More interestingly, the presence of calcium with the concentration of 10 mM in this study impaired the PPAD activity of *P. gingivalis* W83 (activity decreased 46.3%).

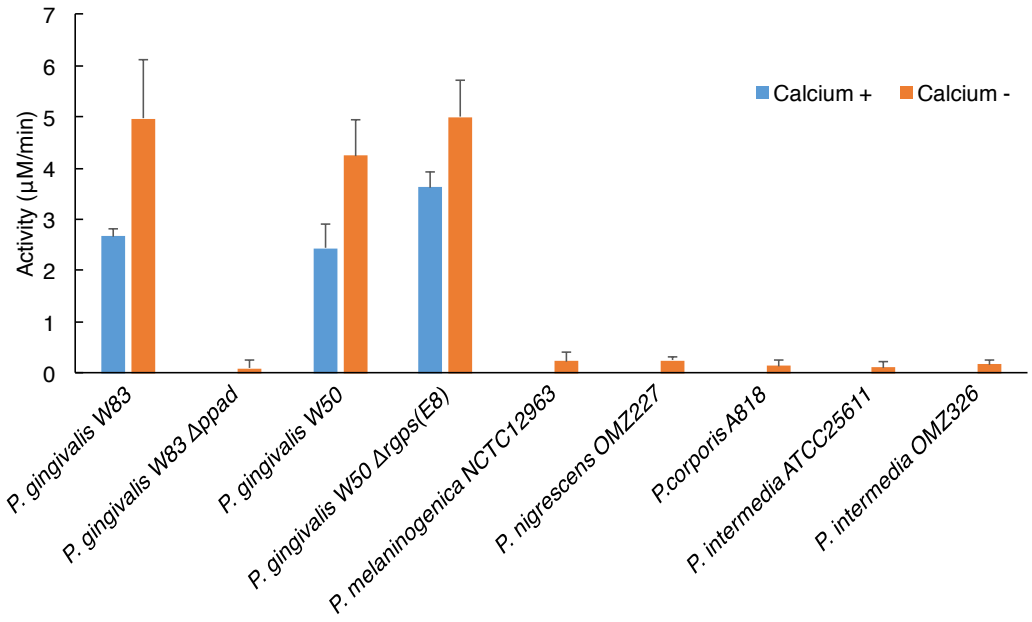


Figure 3.45 Detection of citrullination activity in *P. gingivalis* and *Prevotella* species using BAEE as substrate.
Each species was grown in BHI until the late log phase and the OD_{600} of the culture was adjusted to 1.0 using BHI broth. Bacterial cells were re-suspended in activity buffer with or without calcium for the colorimetric assay for citrullination activity. Results are expressed as means ± standard deviations (n=3). Blue bar: buffer with calcium; orange bar: buffer without calcium.

To examine above results in a more physiologically relevant context, citrullination activity against protein substrate was assessed by using bovine serum albumin (BSA) as substrate. Citrullination was detected in all species regardless of the presence of calcium in the buffer. There was no significant difference in activity between the assays with and without calcium for each species ($P > 0.05$, t test) (Figure 3.46). The activity in *P. gingivalis* W83 and *P. gingivalis* W50 was significantly higher than that of other species in both reaction conditions (calcium+/-) ($P < 0.05$, ANOVA test followed up with Bonferroni correction). The significant decreased activity in the Δ*rgps* mutant compared with *P. gingivalis* W83 or *P. gingivalis* W50 indicated that, with the complex substrate such as BSA, PPAD activity was dependent on the presence of active Rgps. Similar to BAEE, no

significant difference in activity was found between the *Prevotella* spp. and the two *P. gingivalis* mutants ($P > 0.05$, ANOVA test).

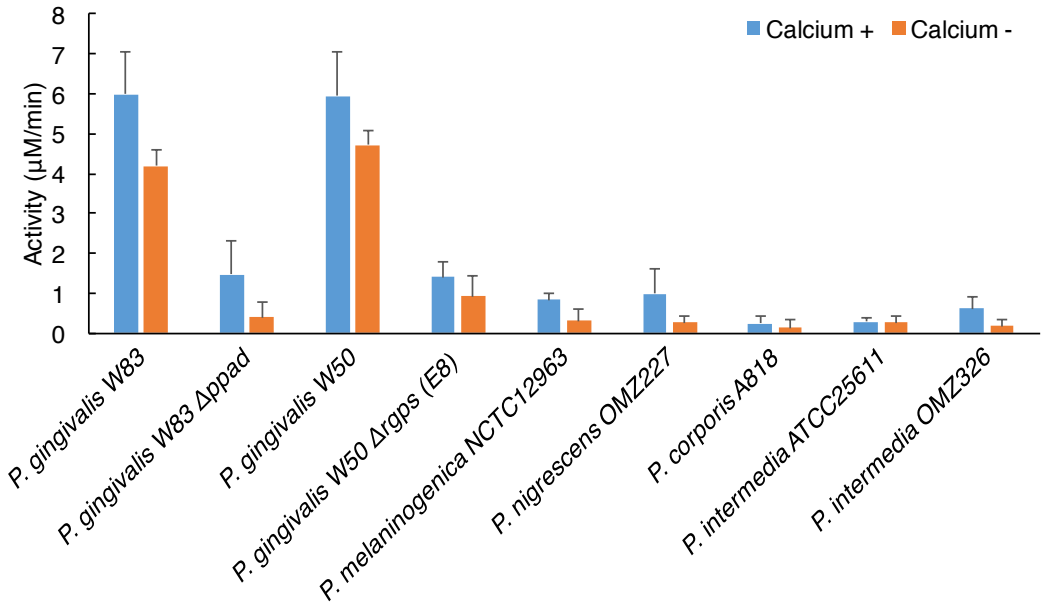


Figure 3.46 Detection of citrullination activity in *P. gingivalis* and *Prevotella* species using BSA as substrate.
Each species was grown in BHI until the late log phase and OD_{600} of the culture was adjusted to 1.0 using BHI broth. Cell pellets of the bacteria were re-suspended in the activity buffer with or without calcium for the colorimetric assay for citrullination activity. Results are expressed as means ± standard deviations (n=3). Blue bar: buffer with calcium; orange bar: buffer without calcium.

To examine the substrate specificity of potential PAD/PAD-like enzyme in *Prevotella* spp., compared with PPAD, colorimetric detection of citrullination was performed with three short peptides containing arginine at different positions. Statistically, an ANOVA test followed up with Bonferroni adjustment was used for the multiple comparisons and the significance level was set at 0.05. Activities were observed in all species, albeit at low levels for *Prevotella* spp., but varied with arginine residue position. *P. nigrescens* OMZ227 and *P. intermedia* ATCC25611showed no preference for any arginine residue positions (Figure 3.47),

while all other species including *P. melaninogenica* NCTC 12963, *P. corporis* A818 and *P. intermedia* OMZ326 had higher citrullination activity for C-terminal arginine peptides (Gly-Arg) than that for internal (Met-Arg-Phe) and N-terminal (Arg-Gly-Glu) arginine peptides (ANOVA test). This result showed that, besides the independency on calcium, the substrate specificity of the potential PAD or PAD-like enzyme in *Prevotella* spp. is also more like that of PPAD which has a preference for C-terminal arginine residues. Again, there was no significant difference in the citrullination activity comparing *Prevotella* spp. and the *P. gingivalis* Δ*ppad* mutant for any of the substrates I used. Rgps deficiency in the *P. gingivalis* W50 had no effect on its PPAD activity, irrespective of the arginine residue position in the substrate.

In the assay using C-terminal arginine peptides as the substrate, the activities were significantly lower in the *Prevotella* spp. compared with that in *P. gingivalis* W83, *P. gingivalis* W50 and its Δ*rgps* mutant. For the internal arginine peptide, the activity in *P. gingivalis* W83 was significantly higher than that in *P. corporis* A818 and *P. intermedia* ATCC25611. No significant difference was found in the activities when comparing *P. gingivalis* W83 with other *Prevotella* spp. In contrast, there was no significant difference in activity when comparing *P. gingivalis* W83 with any of the *Prevotella* spp. investigated for the N-terminal arginine peptides.

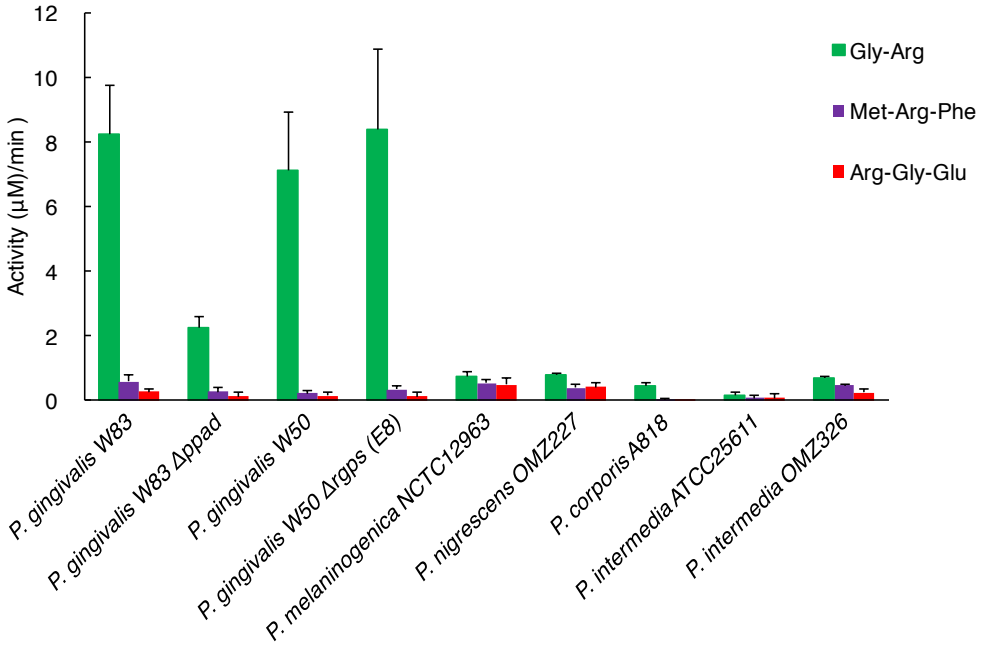


Figure 3.47 Detection of citrullination activity in *P. gingivalis* and *Prevotella* species using substrates with different positions of arginine residues.

Each strain was grown in BHI until the late log phase and OD_{600} of the culture was adjusted to 1.0 using BHI broth. Cell pellets of the bacteria were re-suspended in the activity buffer with calcium for the colorimetric assay for citrullination activity. Synthetic peptides containing C-terminal (green), internal (purple), N-terminal (red) arginine residue were used as substrate. Results are expressed as means ± standard deviations (n=3).

Chapter 4

Discussion

4.1 *In vitro* study of PPAD in *P. gingivalis* physiology

4.1.1 Effect of *P. gingivalis* and PPAD on the local environmental pH

During the development of periodontal disease, the local environmental parameters change such as pH and redox potential. In health the pH of the gingival sulci is 6.90, whereas in disease that can rise to between 7.20 and 7.40 (Eggert et al., 1991, Marsh and Devine, 2011). The changes in the local environmental pH has an effect on the growth as well as the metabolic properties of the microbial community. *P. gingivalis* can be grown stably in pure culture in a chemostat, with an optimum growth between pH 7.00 and 8.00 (McKee et al., 1988). Significant virulence factors of *P. gingivalis,* such as gingipains (Rangarajan et al., 1997, Scott et al., 1993) and PPAD (Abdullah et al., 2013) also showed a preference for slightly alkaline conditions. However, the influence of *P. gingivalis* and its virulence factors on the local environmental pH is less known. *P. gingivalis* was shown to be able to raise the growth pH slightly in batch culture, although the mechanism was not clear (Takahashi and Schachtele, 1990). Because the process of citrullination produces ammonia, PPAD was hypothesized to be able to assist in the maintenance of the alkaline conditions required for *P. gingivalis*. The hypothesis was disproved in the present study as the *P. gingivalis* W83 PPAD deficient mutant exhibited a significantly higher growth pH compared with the wild-type strain in the chemostat culture without additional pH control, although the pH following growth of both strains was below 7.00 (Figure 3.4a, Table 3.1). The present data suggest that there must be an alternative strategy *P. gingivalis* utilizes to adjust the environmental pH and PPAD may have a negative effect on the maintenance of an alkaline

environment, despite the fact that it can generate the ammonia through citrullination. Takahashi *et al*. have shown that cell suspensions of *P. gingivalis* in the presence of amino acid or dipeptide raised the pH from 5.50 to 6.55～6.70 after a six-hour incubation and this acid-neutralizing activity was due to base generated from amino acid degradation (Takahashi, 2003). In the present study, *P. gingivalis* W83 wild-type, the identical strain used in their study (Takahashi, 2003), showed a more acidic growth pH compared with the mutant without any PPAD activity, indicating that PPAD may impair the acid-neutralizing activity or alkali promoting activity of *P. gingivalis* by affecting amino acid fermentation, which is possibly due to the citrullination of related enzymes by PPAD.

4.1.2 Effect of *P. gingivalis* on the local environmental redox potential

Redox potential is another vital environmental parameter affecting the growth and properties of *P. gingivalis* (Leke et al., 1999). During the development of periodontitis, the formation of periodontal pockets is associated with a significant decrease in redox potential, thus, resulting in a reduced environment favourable for the multiplication of anaerobic periodontal pathogens such as *P. gingivalis*. From healthy gingival sulci to periodontal pockets, the redox potential decreased from approximately +74 mV to -157~ +14 mV(Kenney and Ash, 1969), although lower values have also been claimed (ca. -300 mV) (Marsh, 2016). Similar to the pH, oral bacteria are also able to modify the redox potential of the environment. Previous study has shown that *P. gingivalis* decreased the redox potential from -74 mV to -360~ -333 mV, in batch culture under anaerobic conditions (Leke et al., 1999), although the mechanism of such modification is not clear.

In the present study, the redox potential was well controlled for both strains at approximately -350 mV and anaerobic conditions were maintained during the first steady-state by gassing the culture vessel with a mixture of N_2 and CO_2. However, in the second steady state without pH control, the redox potential of the medium following growth of the mutant was significantly lower than that following growth of the wild-type, indicating a more reduced environmental condition (Figure 3.4b, Table 3.1). It is possible that the redox potential was affected by the environmental pH as the pH can modify the driving force of reactions involving protons/hydroxide ions and subsequently influence the redox potential. Furthermore, *P. gingivalis* is likely able to modify the local redox potential by regulation of related gene expression. By comparing gene expression between the two strains using the chemostat culture samples, PG_0275 encoding thiol reductase thioredoxin was found over-expressed in the mutant, which may cause a decrease in redox potential as described above. The thioredoxin family of proteins constitutes a disulfide-reducing system present in a variety of bacteria, in which they are involved in oxidative stress protection (Lewis et al., 2009). Although the mechanism for the regulation of this gene is still not clear, the present data indicated that PPAD deficiency can affect the expression of the gene encoding thioredoxin family proteins, which may subsequently influence the redox potential of the local environment.

4.1.3 Effect of environmental pH and redox potential on Rgp activity

P. gingivalis is an asaccharolytic organism that utilizes amino acids as sources of energy, carbon and nitrogen. Gingipains are major *P. gingivalis* virulence factors which can degrade many human proteins (Potempa and Travis, 1996). In my study, the Rgp activity of both wild-type and mutant strains significantly decreased in the chemostat cultures in which pH was not controlled (Figure 3.6), suggesting that Rgp activity can be affected by the growth pH and the alkaline culture condition leads to a higher Rgp activity. In agreement with the data presented here, Rgp activity decreased with the reduction of growth pH in a previous study (McKee et al., 1988). By comparing the wild-type and the mutant, the significant difference of the Rgp activity was only identified when the growth pH was also significantly different between the two strains and the effect of growth pH on Rgp activity accords with the previous data.

It has been illustrated that a positive redox potential (1.5 mM) can inhibit Rgp activity (Leke et al., 1999), but it is not clear whether the different redox potential in the present study can affect the Rgp activity as they were all negative (-338.42 mV~ -419.60 mV).

4.1.4 Effect of PPAD on Rgp and DPP activity

Apart from the effect of those environmental parameters on Rgp activity, PPAD may decrease Rgp activity via citrullination. It has been shown that PPAD is able to citrullinate endogenous proteins of *P. gingivalis* (Wegner et al., 2010, Bickel and Cimasoni, 1985). Intriguingly, by mass spectrometry analysis, a recent study

demonstrated that six proteins of *P. gingivalis*, especially the RgpA of *P. gingivalis* W83, were citrullinated dependent on the presence of PPAD (Stobernack et al., 2016). Although it is unknown whether the activity of the citrullinated proteins were changed in their study, citrullination, as a post-translational modification of proteins, can alter the charge and structure of the respective target proteins and may consequently change their functions (Gyorgy et al., 2006). Of note, gingipains account for 85% of the total proteolytic activity of *P. gingivalis* (Potempa et al., 1997) and they are crucial in amino-acid utilization by *P. gingivalis*, as nutritional extracellular proteins are initially degraded to oligopeptides by the gingipains for further metabolism (Nemoto and Ohara-Nemoto, 2016). Therefore, it is conceivable that citrullination of Rgp may lead to a decrease in proteolytic activity of *P. gingivalis* and subsequently affect the generation of alkali end products during amino acid fermentation, which contributes to lowering the environmental pH. Kgp, which also contributes to the total proteolytic activity of *P. gingivalis* although to a lesser extent (Potempa et al., 1997), was not included in my study.

DPP is another enzyme that is important for the growth of *P. gingivalis*. Among four DPPs tested in my study, only DPP 11 showed a significant reduction in activity against its major substrate (Leu-Asp-X) in the *P. gingivalis* W83 wild-type compared with its PPAD deficient mutant (Figure 3.7). Although this result was derived from the study on batch culture samples instead of chemostat culture of which all environmental parameters were measured, the two strains were grown and processed at same conditions to make the data comparable. This difference found in the DPP activity between the two strains was confirmed not to be due to changes in gene expression of DPP 11 (Figure 3.8). Interestingly, there are studies demonstrating that the substrate preference of *P. gingivalis* DPP11 is primarily

mediated by its Arg_{673} residue (Sakamoto et al., 2015). Therefore, it is highly possible that PPAD can citrullinate the DPP11 at Arg_{673} and alter the substrate preference of DPP11, which may subsequently cause the reduction of DPP11 activity against the specific substrate.

Taken together, my study suggests that PPAD may affect the Rgp and DPP activities through citrullination, although it is unclear why *P. gingivalis* would reduce the activity of its important peptidase by PPAD and further studies are needed to confirm the citrullination of these enzymes with the help of western blot or mass spectrometry. With respect to the development of RA, these citrullinated proteins of *P. gingivalis* may lift the total citrullination level in the human host *in vivo* and contribute to the aetiology of RA.

4.1.5 Effect of environmental parameters on PPAD activity

In the present study, PPAD activity in the cells of *P. gingivalis* W83 wild-type decreased when growth pH was not externally controlled and became slightly acidic (pH 6.69 ± 0.03) compared with activity from chemostat cultures grown in controlled conditions (pH 7.25 ± 0.05) (Figure 3.5); thus growth pH may affect PPAD activity and my results suggest that the optimal growth pH for cellular PPAD activity is neutral-alkaline rather than acidic. My findings are consistent with the results of Abdullah *et al.*, who reported that optimal pH of assay buffer for cellular PPAD activity was alkaline (Abdullah et al., 2013). Using purified PPAD, a pH optimum for activity was also found to be 9.3 (McGraw et al., 1999).

4.1.6 Effect of environmental parameters on the growth and gene expression of *P. gingivalis*

Viable cell numbers and OD_{600} of both strains increased in the chemostat cultures when the environmental pH decreased after removing pH control of 7.25 ± 0.05, and the difference in the wild-type strain was statistically significant (Table 3.1). My findings, differed from those of McDermid *et al*., who reported that the maximum yields were obtained between pH 7.00 and 8.00 by comparing the growth of *P. gingivalis* W50 in separate continuous culture experiments with differing environmental pH (McKee et al., 1988). In my study the bacteria were grown continuously and the external pH control was used and then removed, leading to different environmental conditions. Thus, it is possible that *P. gingivalis* in the present study may have had the opportunity to adapt to the changes in the environment and subsequently even promote its growth as found in my results. In the second steady-state the OD_{600} of the Δ*ppad* mutant samples was significantly higher than that of the wild-type strain indicating a better growth of the bacteria, which is in consistent with the findings of higher environmental pH and lower redox potential in the mutant culture, as those environmental conditions favours *P. gingivalis* growth. The OD_{600} of the Δ*ppad* mutant samples in the second steady state without external pH control was significantly higher than that in the first steady state in which the environmental pH was controlled at 7.25 ± 0.05, but the difference of viable cell number did not reach the significance, although it is also higher in the second steady state. It is not determined whether the cell size of the mutant in the present study was different between the two steady states, but it is possible that the mutant cells were larger in the second steady state, which may lead to the significant higher OD_{600} even though the viable cell number was not significantly increased.

Previous study has shown that atypical colony forms were observed during prolonged growth of *P. gingivalis* W50 in a chemostat, which had a reduced virulence (McKee et al., 1988). It is plausible that the Δ*ppad* mutant used in my study had a similar tendency to produce variants, and this might be a strategy by which *P. gingivalis* could persist adapting to the changes of environment. The cell sizes of the Δ*ppad* mutant at different culture conditions could be determined and compared by using atomic force microscopy in fulture study (Osiro et al., 2012).

By analysing the RNA sequencing data, differentially expressed genes were identified in both *P. gingivalis* W83 wild-type strain and its Δ*ppad* mutant after removing pH control, which indicated that environmental pH can regulate the gene expression of both *P. gingivalis* strains. By annotating those up/down-regulated genes to GO terms, more GO terms were identified in the mutant (Figure 3.16) compared with the wild-type strain (Figure 3.15) and most of the up-regulated genes in the mutant were involved in molecular function. In agreement with the results of PCA analysis (Figure 3.9), the data presented here suggested that the removal of pH control had less influence on the gene expression in *P. gingivalis* W83 wild-type strain than its Δ*ppad* mutant.

The expression of *HmuY* and PG_1837 (*hagA)*, encoding HmuY protein and haemagglutinin A, respectively, were found significantly up-regulated in the non-pH controlled chemostat culture of *P. gingivalis* W83 wild-type, compared with the culture under the controlled pH of 7.25 ± 0.05 (Appendix C2). HmuY is a haem-binding protein which is important in haem utilisation by *P. gingivalis*. HmuY captures free haem and delivers it to an outer-membrane transporter, the TonB-dependent receptor HmuR, which transports haem into the bacterial cell (Wojtowicz

et al., 2009). A *P. gingivalis* hmuY deficient mutant grew slower compared with the parent strain in haem-limited conditions (Olczak et al., 2015). Haemagglutinin A also take parts in haem acquisition through haemagglutination of red blood cells (Han et al., 1996, Olczak et al., 2005). It is well established that growth and colonization of *P. gingivalis* are dependent on the ability to acquire haem (Schifferle et al., 1996) and *P. gingivalis* itself is not able to synthesize protoporphyrin IX, the precursor of haem (Roper et al., 2000). Therefore, despite the fact that in the present study the bacteria were grown in haem excess, the data present here suggest that environmental pH can regulate gene expression by *P. gingivalis* and up-regulated genes such as *hmuY* and *hagA* may help the bacteria to adapt to the decreased pH condition and facilitate bacterial growth.

Within the Δ*ppad* mutant, clustered regularly interspaced short palindromic repeats (CRISPR) -associated genes such as PG_1982, PG_1983 and PG_2013, were up-regulated in the second steady-state compared with the first steady-state (Table 3.4, Appendix C3). CRISPR systems can protect bacteria against the foreign genetic elements, such as viruses, plasmids, and transposons (Burmistrz et al., 2017). Therefore, the changes of environmental parameters in the second steady state can upregulate the expression of CRISPR-associated genes in Δ*ppad* mutant, which may lead to an increase in the defence capability of the Δ*ppad* mutant. Moreover, genes encoding PorT family protein and TonB-dependent receptor, which belongs to the type IX secretion system (T9SS), were up-regulated in the second steady-state chemostat culture of Δ*ppad* mutant, compared with the first steady-state (Appendix C3). T9SS is a robust secretion system facilitating secretion of up to 35 proteins bearing the CTD in *P. gingivalis*, many of which are potent virulence factors, such as gingipains and PPAD (Lasica et al., 2017). A two-component system

(PorX/PorY) and extracytoplasmic function sigma factor have been identified to exhibit regulatory effect on the essential genes of T9SS (Kadowaki et al., 2016). However, the information about T9SS regulation by any environmental signal is still limited. Intriguingly, the present data indicated that environmental parameters such as pH and redox potential can possibly regulate the expression of some T9SS genes in the *P. gingivalis* Δ*ppad* mutant and may subsequently facilitate the secretion of virulence factors. The mechanism of the regulation of those genes requires further study.

4.1.7 Strengths and limitations

Despite the difficulties associated with the setting up of chemostat cultures, they provided clear advantages over batch culturing. Single parameters can indeed be manipulated while others are held constant, and the steady-state can be maintained over long periods of time. More importantly, chemostat culture is able to provide us with reliable biological samples for analysis by RNA sequencing (Hoskisson and Hobbs, 2005). Combination of continuous culture with next generation sequencing techniques is a powerful means for exploration of microbial physiology and pathogenicity.

In the current study, the effect of PPAD on the local environmental pH was determined as well as the influence of environmental parameters such as pH and redox potential on the growth and enzyme activity of *P. gingivalis*. However, oral bacteria such as *P. gingivalis* are part of a diverse, rich and complex microbial community and other bacteria may impact or contribute to the change in environmental conditions *in vivo*. The interactions between bacteria may help *P. gingivalis* to adapt to the changes of the environment. For example, *P. intermedia*

was identified be able to change the acidic subgingival environment toward more neutral pH levels which is suitable for the growth of *P. gingivalis* (Takahashi and Schachtele, 1990). *F. nucleatum*, *Peptostreptococcus micros* and *S. mutans* could decrease the redox potential, which may thus be beneficial for *P. gingivalis* (Leke et al., 1999). Therefore, clinical samples of subgingival biofilm or biofilm models including multiple species might be needed for future study to elucidate the role of PPAD for *P. gingivalis* and other bacterial species.

It is also important to notice that the altered gene expression between different pH conditions or strains does not necessarily translate to differences in protein synthesis or activity. qRT-PCR analysis should be used to confirm the differences in gene expression found by RNA sequencing. Further proteomic analysis, using mass spectrometry or western blot would be helpful for elucidating whether up- or down-regulated gene expression is actually changing the protein expression.

4.2 Metagenomic study of subgingival microbiome in relation to RA

Following the epidemiologic evidence of the link between periodontitis and RA, an accumulating body of studies has been conducted on the mechanism underlying the link. Especially, the role of the subgingival microbiota in these two diseases has been investigated with the help of NGS techniques. However, up to now it has been unclear whether or not local development of periodontitis and the alterations in the local microbial community precede the development of RA, although it has been hypothesized that the initiation of RA may occur at mucosal sites such as the periodontium. The present study targeting prospective cohorts of individuals at risk

for RA development (CCP) provides important insights into the pathogenesis of RA and lays the groundwork for finding potential markers for early diagnosis and prevention.

4.2.1 Periodontitis-related subgingival microbial dysbiosis in the individuals at-risk of RA development.

Alterations in the composition of the subgingival microbiome were observed in the CCP group, compared with the other groups including HC, when analysing subgingival plaque from deep periodontal pockets (4 mm or more) with bleeding on probing. At the genus level, *Porphyromonas* and *Treponema* were significantly higher in the CCP group compared with other groups and within top 20 most abundant genera in the CCP group (Figure 3.24b). Accordingly, *P. gingivalis, Treponema vincentii*, *T. denticola*, as well as *P. intermedia* and *Prevotella disiens* were found significantly associated with the CCP group (Figure 3.26). The increase of these red-complex organisms, such as *P. gingivalis* and *T. denticola*, and other periodontal pathogens in the CCP group is compatible with my clinical findings, which showed that the prevalence of periodontitis increased in anti-CCP positive individuals at-risk of RA development (Table 3.6). The present data indicate that periodontitis and relatively increased abundance of the major periodontal pathogens precedes the development of RA and may represent a risk factor for RA development. However, the mechanisms still remain to be further elucidated by which periodontitis or related pathogens may trigger or facilitate the development of RA. The prevailing speculation is that *P. gingivalis* acts as an environmental trigger for RA as it was shown to be the only organism capable of citrullinating both bacterial and host proteins, such as fibrinogen and α-enolase through its PPAD (Wegner et al., 2010). PPAD was shown to be omnipresent in *P. gingivalis*

(Gabarrini et al., 2015). PPAD genes were also identified in my data with the highest prevalence in the CCP group (data not shown). But they were not detected in all samples with presence of *P. gingivalis*, which could be due to the limitation of the shotgun sequencing method used in my study. The sample is just a fraction of the original environment and the limited sequencing reads may not be able to capture the PPAD gene in all samples in the present study. In addition, it is not known in the present data whether there is any difference in the expression or activity of PPAD between different groups, which may influence autoimmunity in RA. Analysis of PPAD activity in the clinical samples should be included in future studies to further determine if efficiency or amount of citrullination by PPAD is associated with the development of RA.

Increased abundance and prevalence of *P. gingivalis* have been observed in RA patients compared with healthy controls in a different study (Hitchon et al., 2010), although there were inconsistent results which may be due to the different types of techniques and samples used for the measurements (de Smit et al., 2012). In a recent comprehensive study, *P. gingivalis* was found enriched in healthy controls rather than RA patients when analysing the microbiome of saliva and supra-gingival dental plaque using shotgun sequencing (Zhang et al., 2015). Although *P. gingivalis* can grow anaerobically as well as when exposed to low levels of oxygen (Diaz and Rogers, 2004), the influence of environment conditions such as redox potential, should be taken into consideration when interpreting the data of samples from different sources, such as sub- and supra-gingival environments (Daniluk et al., 2006). In agreement with the present study, Mikuls *et al.* showed an increased level of anti-*P. gingivalis* antibodies in individuals at genetic risk of developing RA (Mikuls Ted et al., 2012). However, periodontal examination was not included in

their study and many other periodontal pathogens were not examined. Thus, the possibility cannot be excluded that RA related autoantibody production is a consequence of periodontal inflammation rather than of *P. gingivalis* activity alone.

4.2.2 Subgingival microbial dysbiosis in the individuals without periodontitis but at-risk of RA development

In the present study, I aimed to further determine if specific periodontal pathogen such as *P. gingivalis*, rather than the periodontitis status, is implicated in the development of RA. To preclude the effect of established periodontitis on the subgingival microbiome, metagenomics analysis was performed only on the subgingival plaque samples from shallow gingival sulci (3 mm depth or less) with no bleeding on probing from individuals without periodontitis. Alterations were also found in the composition of the subgingival microbiome, which distinguished the CCP group from others. At the phylum level, *Spirochaetes* were also found with significantly higher relative abundance in the CCP group compared with other groups (Figure 3.40). As expected, at the species level none of the prominent and red-complex periodontal pathogens were significantly present in any groups (Figure 3.42). Intriguingly, I found that 14 other species were significantly associated with the CCP group including six *Capnocytophaga* spp, three *Prevotella* spp, two *Fusobacterium* spp, two *Treponema* spp and *Leptotrichia* sp. oral taxon 879. It is unknown if these species are involved in the pathogenesis of RA, but most of these species are associated with the development of periodontitis. Therefore, currently I can only speculate that these species found associated with the CCP group may be implicated in the onset of RA via facilitating the development of periodontitis. My speculation is supported by the recent study where gram-negative anaerobes were

significantly more abundant in RA patients who were periodontally healthy, indicating a preclinical phase of periodontitis (Lopez-Oliva et al., 2018).

Among the predominant species in my study, *F. nucleatum*, a member of the orange complex and an important organism in plaque maturation, was found with the highest relative abundance in the CCP group. Previously *F. nucleatum* was detected in both the synovial fluid and dental plaque samples of RA patients diagnosed with periodontitis (Témoin et al., 2012) and was found in higher concentrations in anti-CCP positive patients with RA (Schmickler et al., 2017). The presence of *T. socranskii* was reported to be associated with periodontitis (Takeuchi et al., 2001), although conflicting results exist which may be due to differences between assays and the study populations (Riviere et al., 1997). In the present study, *Capnocytophaga* spp. accounted for nearly half of those significant species in the CCP group (6/14) (Figure 3.42). The genus *Capnocytophaga* encompasses a group of fastidious capnophilic and facultatively anaerobic Gram-negative bacilli. These bacteria are involved in causing endocarditis as well as arthritis, osteomyelitis or periodontitis (Piau et al., 2013). Similar to *P. gingivalis*, the LPS of *C. ochracea* can act as antagonists for human TLR4, which may help the bacterium escape from the innate immune system (Yoshimura et al., 2002). *Capnocytophaga* spp. are also involved in the formation of dental biofilms related to periodontal health and disease; in particular, a diffusible soluble component released by *C. ochracea* was identified as important in biofilm formation during co-culture with *F. nucleatum* (Jolivet-Gougeon et al., 2004). More interestingly, clinical isolates of *Capnocytophaga* spp. including *C. ochracea* were reported to exhibit the resistance to beta-lactam antibiotics through the activity of beta-lactamases, which may act as a major reason for unsuccessful periodontal treatment to control subgingival

pathogenic organisms (Jolivet-Gougeon et al., 2004). The specific *Leptotrichia* species found in the present study has not been completely characterized yet, but the *Leptotrichia* genus is known to be a member of the oral commensal microbiota and may act as an opportunistic pathogen associated with periodontal diseases and oral cavity abscesses (Eribe and Olsen, 2017). Taken together, these species associated with the anti-CCP positive at-risk individuals are worth further study as they may have an important role in the development of RA. Investigations of these species may provide the opportunity of finding methods for early risk assessment, diagnosis or prevention of RA.

4.2.3 Effect of RA and DMARD treatment on the subgingival microbiome

Anti-CCP positive RA patients were included in the present study who had been diagnosed with RA and treated with DMARD therapy for no more than three months (NORA) or at least six months (RA). Differences in subgingival microbiomes were also observed comparing NORA and RA groups and others using healthy site samples from individuals without periodontitis (Figure 3.41), which may be related to the effect of development of RA over time. Similarly, a prior study showed that RA may act as a condition shaping the subgingival microbiome, particularly promoting the growth of certain organisms (Lopez-Oliva et al., 2018). However, in contrast to my results, the microbial profiles were identified similar between NORA and chronic RA previously (Scher et al., 2012). It is should be noted that the NORA patients in their study were treatment-naïve which is different to my study group and the differences in the subgingival microbiomes could also be explained by the effect of DMARD treatment and the duration of the treatment (Romero-Sanchez et al., 2017).

It is highly possible that RA therapy, particularly those with proposed antibacterial properties, such as methotrexate or hydroxychloroquine (Greenstein et al., 2007, Rolain et al., 2007), will influence the subgingival microbiome. Moreover, the major tissue destruction associated with periodontitis results from the immune-inflammatory response to the bacterial challenge; thus it is possible that the immunomodulatory effects of the RA regime such as anti-tumour necrosis factor (TNF) therapy, can affect the development of periodontitis (Üstün et al., 2013), and may subsequently modify the subgingival microbiome. A recent study using shotgun sequencing identified alterations in the oral microbiome, which were partially restored by DMARD treatment (Zhang et al., 2015). However, the periodontal status of the participants was unknown, making the results somewhat ambiguous as observations might have been due to the effect of periodontal disease (Griffen et al., 2012). Additional studies with treatment of naive individuals receiving periodontal examination will be needed in the future to clearly determine the effect of RA disease and DMARD treatment on the subgingival microbiome.

4.2.4 Uniquely detected species including *A. actinomycetemcomitans* in the subgingival microbiome of individuals at-risk of RA

Dysbiosis of the subgingival microbiome related to the development of RA was also supported by identification of unique species present in the CCP group (section 3.10.7). Although each group had its own unique species, the number of the species was higher in the CCP group compared with other groups, either in healthy or diseased sites. Most interestingly, *A. actinomycetemcomitans* was only observed in the CCP cohort without periodontitis, with a prevalence of 15.4% (Appendix C10). A recent study demonstrated that *A. actinomycetemcomitans* can induce

hypercitrullination in host neutrophils through its leukotoxin-A and may act as a trigger of autoimmunity in RA (Konig et al., 2016). Different to the analysis of the unique species detected in each group, the core microbiota, referred to the species that were detected in at least 50% of samples above 0.2% relative abundance, was identified for each group (section 3.10.9). Although the size of a core species set is highly conditional, mostly depending on the detection threshold (Salonen et al., 2012), *A. actinomycetemcomitans* was absent in the list of core species as it was below the threshold of prevalence in this analysis (Table 3.11). Thus the unique of *A. actinomycetemcomitan* in the CCP group could be explained by the shortcomings of present study including variances between clinical samples and limited sequencing depth. However, it is conceivable that *A. actinomycetemcomitans* was scarce in the present study as it is more related to aggressive periodontitis, which was not observed in the current study population (Shaddox et al., 2012). Considering the implication of *A. actinomycetemcomitans* in autoimmunity in RA and its unique presence in the CCP group in this study, further analysis of *A. actinomycetemcomitans* and its leukotoxin-A would be interesting and might be helpful to understand the pathogenesis of RA.

4.2.5 Microbial diversity in the subgingival microbiome

Consistent with several previous studies (Abusleme et al., 2013b, Griffen et al., 2011), bacterial diversity (species richness and evenness) was higher in the periodontally diseased sites compared with the healthy sites, implying a shift from healthy to disease-associated subgingival microbiome (Figure 3.20). RA status had no significant influence on the bacterial diversity of periodontally diseased sites (Figure 3.21). Similar results were reported in another study using 454 pyrosequencing, where no significant differences in microbial diversity were

observed between chronic RA, untreated NORA and healthy controls (Scher et al., 2012). Interestingly, the diversity in HC, CCP and NORA group displayed a decreasing trend in both healthy and diseased sites, which might reflect an influence of the development of RA on the microbial diversity. However, I cannot exclude the potential effect of DMARD therapy on the microbial diversity of NORA patients although the treatment was less than 3 months. The findings that microbial diversity of RA patient did not follow the trend could also be explained by the potential effect of longer duration of DMARD treatment (> 6 months).

Moreover, the current study revealed that within the periodontally healthy and diseased sites bacterial diversity was significant lower in NORA patients compared with the CCP group (Figure 3.21). This result could be explained by the highest prevalence of periodontitis in the CCP group. This explanation was confirmed by analysing the periodontally healthy site samples only from non-periodontitis individuals, where no significant differences in microbial diversity were observed between groups (Figure 3.38). My findings were also supported by a recent study showing no significant difference in the microbial diversity between RA and HC who were periodontally healthy (Lopez-Oliva et al., 2018). However, I cannot exclude the possibility that early DMARD therapy (< 3 months) may decrease the microbial diversity of NORA patients and it is possible that periodontal disease associated microbiome may exhibit resistance to the treatment as the significantly lower bacterial diversity of the NORA group was only found in the healthy site samples but not in the diseased sites.

4.2.6 Co-occurrence network analysis

The present study also demonstrated that the CCP group harboured highly connected and more complex species inter-relationships by analysing the species co-occurrence in the diseased sites (Table 3.8). Among all the pair-wise correlations in the CCP group of diseased sites, 94.4% were positive, indicating that synergy of species in periodontitis may be related to the development of RA, rather than single species.

In both HC and CCP groups in the samples from healthy sites, a negative correlation was found between *A. naeslundii* and *P. gingivalis*. In the CCP group periodontally diseased site samples, *P. gingivalis* was negatively correlated with *Actinomyces oris* and three *Veillonella* species. Understanding the mechanism behind these correlations by future study may help us to find a method to adjust and control the composition of the subgingival microbiome and subsequently, prevent the development of related disease.

In summary, the present study demonstrates that dybiosis of the subgingival microbiome related with periodontitis precedes the onset of symptoms of RA and may drive the development of RA. Under the established periodontitis conditions, *P. gingivalis* and its PPAD may play an important role in the initiation of RA.

4.2.7 Potential functional capability

No significant difference was found when comparing the potential functional capabilities between groups in diseased site samples, implying that RA has no impact on the subgingival microbiome under periodontitis conditions. Twenty-nine functional units were significantly under-represented in the NORA group compared with the CCP group in the healthy sites (

Table 3.10), which could be the consequence of RA treatment. Additional caution must be taken when interpreting this output, because only the presence of the related sequences but not the actual expression could be determined by the method employed. Meta-transcriptomic analysis should be performed in future studies to provide information regarding any functional differences between groups, which would be more helpful for us to understand the mechanism(s) underlying the link between periodontitis and RA.

4.2.8 Strengths and limitations

To my knowledge, no prior study has assessed the profile of subgingival microbiomes from individuals at risk of development of RA. The periodontal status of every participant was examined thoroughly and the age, gender and smoking status of participants were well balanced between four groups in this study during the recruitment process. The small sample number of RA group represents a limitation of present study which could lead to more uncertainty when interpreting the results of comparison between RA and other groups. This also makes it difficult to compar with the data of other studies. Larger samples size will be needed to more clearly define the role of the subgingival microbiome in the development and progression of RA.

The present study is the first to utilize shotgun metagenomics to comprehensively analyse the oral microbial mechanism underlying the link between periodontitis and the development of RA. The 16S rRNA gene sequencing method was utilized in most previous studies to analyse the oral microbiome of early or chronic RA patients. This is cost-effective and efficient to detect alterations in bacterial populations (Lopez-Oliva et al., 2018, Scher et al., 2012). However, a major

limitation of the 16S amplicon method is that only a single region of the bacterial genome can be sequenced and it is difficult to distinguish the species when their 16S rRNA gene sequences display high similarities (Větrovský and Baldrian, 2013). Thus, classifications using the 16S rRNA method often lack accuracy at the species level. The shotgun metagenomics approach employed in the present study has been identified with several advantages over the 16S amplicon method, such as, more confident identification of bacterial species, increased detection of diversity and prediction of genes (Ranjan et al., 2016).

Data analysis after the sequencing is always a significant challenge for researchers because it is highly computationally demanding. In my study, sequencing reads were *de novo* assembled into contigs before downstream analysis, which is friendly to the hardware resources and has been proven to be more accurate for species identification (Ranjan et al., 2016).

There are many different pipelines and bioinformatics tools available to use but no agreement has been reached to set a standard data analysing method. An in-house pipeline was used in my study where almost every parameter can be adjusted during analysis and has been validated by previous studies (Li et al., 2017, Belstrøm et al., 2017). Particularly, DIAMOND, which was utilized in current study, has been identified as both fast and accurate (Ranjan et al., 2016). The identity cut-off was set at 95% in DIAMOND which is comparable to a similar study (Zhang et al., 2015), although the more stringent parameters decreases the annotation sensitivity (Ranjan et al., 2016).

Sequence data generated by shotgun metagenomics are affected by multiple sources of systematic variability, of which the major one is the differences in sequence

depth. Thus, normalization is essential for meaningful comparisons between different samples using shotgun metagenomics and different normalization methods may result in differences of interpretation of the end results (Pereira et al., 2018). Rather than rarefying the data to get similar total counts across all the samples, which may lose some precious fractions (McMurdie and Holmes, 2014), the straightforward total count method was used for the normalization. The argument for using this method exists because the sample is just a fraction of the original environment and the total number of reads obtained is not equal to the absolute number of microbes present in that environment (Weiss et al., 2017). Since the relative abundances are constrained by the simplex (sum to 1), traditional statistical methods assuming bacterial abundance to be independent variables are not applicable. Thus, the permutation test, with no requirement for data distribution (Collingridge, 2012), was used to identify the taxa that had a significant differential abundance between the study cohorts.

4.3 Detection of PAD or PAD-like enzyme in the bacterial species other than *P. gingivalis*

For a long time, PPAD was thought to be unique among prokaryotes, with *P. gingivalis* being the only bacterium known to produce and secrete it (McGraw et al., 1999, Gabarrini et al., 2015). This notion has recently been challenged by the evidence that PPAD homologues were found in other *Porphyromonas* species from non-human hosts, although the enzyme activity and substrate specificity have not been characterized yet (Gabarrini et al., 2018). In the present study, by aligning the DNA sequencing data to NCBI NR protein database with at least 70% identity, PAD from different bacterial species such as *C. concisus, P. gulae*, *P. melaninogenica*

and *P. veroralis* as described in NCBI database, were identified in the subgingival plaque samples (Table 3.12). There were also sequences annotated as hypothetical proteins related to *P. gingivalis* PAD-like proteins. Although errors in read mapping and incompleteness of the reference database may exist (Morgan et al., 2010) and the increased alignment sensitivity at 70% identity may lead to decreased precision (Ranjan et al., 2016), the present data suggest the presence of PAD or PAD-like enzymes in bacterial species other than *P. gingivalis*.

To further examine the presence of PAD or PAD-like enzymes, citrullination activity was measured in the cells of a range of *Prevotella* spp. using colorimetric assays with different substrates and with varying calcium concentrations. Albeit with a low level of activity compared with PPAD ($P < 0.05$, ANOVA test) , citrullination was observed in all *Prevotella* spp. in my study, including *P. nigrescens* OMZ227, *P. intermedia* OMZ326, *P. melaninogenica* NCTC 12963, *P. corporis* A818 and *P. intermedia* OMZ326. Among these species, only the genome of *P. intermedia* ATCC25611 (https://www.ncbi.nlm.nih.gov/protein/1132728393) and *P. melaninogenic* NCTC 12963 (https://www.ncbi.nlm.nih.gov/protein/WP_013265317.1) are available on NCBI and these contain sequences annotated as PAD. I did not, due to time constraints, determine the presence of PAD or PAD-like gene in those *Prevotella* spp. without the genome information used in the citrullination assay. To avoid any false positive results of the citrullination assay, control reactions containing only bacteria or substrate were included in my study. Although the details of the reaction and the exact substrate specificity are unknown, the present data also indicate that PAD or PAD-like enzymes in those *Prevotella* spp. are like PPAD in having a preference for C-terminal arginine residues (Figure 3.47), and the enzyme activity was not

dependent on the presence of calcium (Figure 3.45). The reaction conditions used in my study was optimal for PPAD, however, they might not be for the PAD or PAD-like enzymes in *Prevotella* spp. Different conditions such as pH and temperature need to be tested for these PAD or PAD-like enzymes in future studies.

For PPAD, there are mainly two forms as described previously (McGraw et al., 1999). The dominant form is cell associated, which is delivered by the outer membrane vesicles (OMV) and the other form is soluble secreted PPAD which accounts for a smaller fraction. However, it is unknown if the PAD or PAD-like enzyme of *Prevotella* spp. is more cell associated. Further studies are needed to determine the potential citrullination activity in the supernatant samples of *Prevotella* spp and to compare it with *P. gingivalis*.

Previous study has shown that endogenous citrullinated proteins were detected universally in *P. gingivalis* by western blot, but absent in other oral bacteria such as *P. intermedia* H13, *Prevotella oralis* ATCC 33269, *Capnocytophaga gingivalis* ATCC 33624, *Capnocytophaga ochracea* ATCC 27872, *etc*. (Gabarrini et al., 2015, Wegner et al., 2010). Apart from *P. gingivalis*, none of those species were included in my citrullination assay. Detection of the citrullinated proteins from these *Prevotella* spp. would be needed to confirm the presence of active PAD or PAD-like enzymes. PPAD is able to citrullinate the endogenous proteins of *P. gingivalis* (Stobernack et al., 2016). However, considering the potential differences between these PAD or PAD-like enzymes and PPAD, not only the endogenous citrullinated proteins, but also the human proteins such as fibrinogen and α-enolase should be examined.

Moreover, it has been reported that recombinant full-length PPAD (rPPAD) was auto-citrullinated *in vitro* and antibodies specific for citrullinated rPPAD were identified in RA patients (Quirke et al., 2014), although inconsistent results showed that auto-citrullination of PPAD might not occur i*n vivo* due to N-terminal cleavage of the protein (Konig et al., 2015). It would be interesting to know if the PAD or PAD-like enzymes from other species are able to be auto-citrullinated by themselves or citrullinated by PPAD, especially when considering the role of auto antibody production in the aetiology of RA.

Collectively, the PAD or PAD-like enzymes were detected in bacterial species such as *Prevotella* spp.by metagenomics studies, and representatives of *Prevotella* spp. were identified that exhibited some citrullination activity *in vitro* in a similar manner to PPAD. To my best knowledge, the present study is the first to highlight PAD or PAD-like enzymes in species other than *Porphyromonas* spp. Considering the predominant prevalence and abundance of these *Prevotella* spp. such as *P. intermedia* and *P. nigrescens* in the subgingival microbiome and their relationship to the periodontal disease (Zhang et al., 2017, Hong et al., 2015), the sum of their citrullination activity cannot be ignored and may contribute to the autoimmunity in of RA, even though the citrullination activity of these PAD or PAD-like enzymes is significantly lower than that of PPAD.

In addition to *Prevotella* spp., *P. gingivalis* was also found to harbour a gene encoding a type of PAD or PAD-like enzyme, in addition to PPAD, which provided a low level of citrullination activity even when the PPAD activity was depleted by mutation. More attention to its PAD-like enzyme might be needed when analysing the role of *P. gingivalis* in the link between periodontitis and RA.

Chapter 5

Conclusions

By comparing the Rgp and DPP activity between *P. gingivalis* W83 wild-type strain and its PPAD deficient mutant, my study suggested that PPAD may citrullinate those enzymes *in vitro* and impair their activity, although further studies are needed to confirm the citrullination of those enzymes by using western blot or mass spectrometry. It is unclear why *P. gingivalis* would reduce the activity of its important peptidase such as Rgp and DPP 11 by PPAD. However, with respect to the development of RA, these citrullinated proteins of *P. gingivalis* may lift the total citrullination level in the human host *in vivo* and contribute to the aetiology of RA.

By chemostat culture study, my results demonstrated that there must be an alternative strategy related to PPAD, which can be utilized by *P. gingivalis* to regulate the environmental pH, despite the fact that PPAD can generate ammonia through citrullination which may help neutralise acid. *In vitro*, PPAD may impair the acid-neutralizing activity or alkali-generating activity of *P. gingivalis* by affecting amino acid fermentation; this may be due to the citrullination of related enzymes, such as Rgp and DPP 11, by PPAD. Meanwhile, Rgp activity can be affected by the environmental pH, particularly, acidic culture conditions may lead to a lower Rgp activity.

By RNA sequencing analysis, my data suggested that both environmental pH and PPAD deficiency can affect gene expression of *P. gingivalis*. Although the mechanisms for the regulation of those genes are still not clear, up-regulated genes in the conditions without external pH control, such as *hmuY* and *hagA*, may help the bacteria adapt to the changes of the environment and facilitate bacterial growth. Expression of essential T9SS genes were up-regulated after removing pH control in *P. gingivalis* Δ*ppad* mutant, the strategy of which could be utilised by the bacteria to

promote the secretion of virulence factors such as Rgp and PPAD. Moreover, PPAD deficiency leading to the over-expression of the gene encoding thioredoxin family proteins may consequently influence the redox potential of the local environment which benefits the growth of the mutant.

In the clinical study, an increased prevalence of periodontitis was found in anti-CCP positive individuals who are at risk of RA development. Shotgun metagenomic analyses of subgingival dental plaque provided a detailed characterisation of the microbial communities in a large group of HC, CCP, NORA and RA patients. In conclusion, periodontitis and related dybiosis of the subgingival microbiome precede the onset of symptoms of RA and may drive the initiation of RA. Although the mechanisms still remain to be further elucidated by which periodontitis or related pathogens may trigger or facilitate the development of RA, *P. gingivalis* and its PPAD under the established periodontitis conditions may play an important role in the initiation of RA.

Furthermore, the analysis of shotgun sequencing data suggested the presence of PAD or PAD-like enzymes in bacterial species such as *C. concisus, P. gulae*, *P. melaninogenica, P. veroralis* and *P. gingivalis*. Representatives of *Prevotella* spp. were identified that exhibited some citrullination activity *in vitro* in a similar manner to PPAD. The sum of their citrullination activity cannot be ignored and may contribute to the autoimmunity in of RA, even though the activity of the PAD or PAD-like enzyme in a single species is significantly lower than that of PPAD.

Chapter 6

Future Study

In order to confirm the citrullination of Rgp and DPP11 by PPAD, mass spectrometry or western blot analysis should be performed to identify the citrullinated forms of these enzymes *in vitro*. Kgp, which is also important for the amino acid metabolism of *P. gingivalis* but not covered by the present study, should be analysed to see if it can be citrullinated by PPAD. Furthermore, qRT-PCR analysis should be carried out to determine the differences in the gene expression found by RNA sequencing.

To confirm the presence of PAD or PAD-like enzyme in the species other than *P. gingivalis*, the potential citrullinated proteins by these enzymes should be determined by using proteomic analysis. The reaction conditions of citrullination activity assays used in the present study were optimal for PPAD. Considering the potential differences between PPAD and PAD or PAD-like enzymes, different conditions such as pH and temperature need to be tested for the PAD or PAD-like enzymes. Both supernatant and different fractions of bacterial cells could be included to elucidate if those PAD or PAD-like enzyme is cell associated or secreted, or both. An appropriate bacterial strain without any citrullination potentials is also required as a negative control to confirm the activity of PAD or PAD-like enzymes in those species.

For the study on the clinical samples, further investigations with larger samples size and technical replication on the unique species present in the CCP group who are at risk of developing RA would be useful to identify other species that may contribute to the initiation of RA. Comprehensive analysis on the co-occurrence networks of subgingival microbiome of each cohort will provide more information about the species inter-relationships and help to identify any consortia of co-operating bacteria

related to the aetiology of RA. Moreover, any antagonistic relationships between bacterial species can be possibly utilized to develop approaches to control the microbial dysbiosis related with disease. Longitudinally, monitoring the status of anti-CCP positive individuals who are at risk of developing RA and investigation of the subgingival microbiome of those who develop to symptomatic RA would be helpful to understand the pathogenesis of RA. Trials of periodontal treatment on the anti-CCP positive individuals at-risk of RA could be included in the future study, which will provide better understanding of the link between periodontitis and RA.

List of References

References

AAS, J. A., PASTER, B. J., STOKES, L. N., OLSEN, I. & DEWHIRST, F. E. 2005. Defining the normal bacterial flora of the oral cavity. *Journal of Clinical Microbiology,* 43**,** 5721-5732.

ABDULLAH, S.-N., FARMER, E.-A., SPARGO, L., LOGAN, R. & GULLY, N. 2013. Porphyromonas gingivalis peptidylarginine deiminase substrate specificity. *Anaerobe,* 23**,** 102-108.

ABE, N., BABA, A., TAKII, R., NAKAYAMA, K., KAMAGUCHI, A., SHIBATA, Y., ABIKO, Y., OKAMOTO, K., KADOWAKI, T. & YAMAMOTO, K. 2004. Roles of Arg- and Lys-gingipains in coaggregation of Porphyromonas gingivalis: identification of its responsible molecules in translation products of rgpA, kgp, and hagA genes. *Biol Chem,* 385**,** 1041-7.

ABUSLEME, L., DUPUY, A. K., DUTZAN, N., SILVA, N., BURLESON, J. A., STRAUSBAUGH, L. D., GAMONAL, J. & DIAZ, P. I. 2013a. The subgingival microbiome in health and periodontitis and its relationship with community biomass and inflammation. *The ISME Journal,* 7**,** 1016-1025.

ABUSLEME, L., DUPUY, A. K., DUTZAN, N., SILVA, N., BURLESON, J. A., STRAUSBAUGH, L. D., GAMONAL, J. & DIAZ, P. I. 2013b. The subgingival microbiome in health and periodontitis and its relationship with community biomass and inflammation. *The Isme Journal,* 7**,** 1016.

ADUSE-OPOKU, J., DAVIES, N. N., GALLAGHER, A., HASHIM, A., EVANS, H. E., RANGARAJAN, M., SLANEY, J. M. & CURTIS, M. A. 2000. Generation of lys-gingipain protease activity in Porphyromonas gingivalis W50 is independent of Arg-gingipain protease activities. *Microbiology,* 146 (Pt 8)**,** 1933-40.

AL-QUTUB, M. N., BRAHAM, P. H., KARIMI-NASER, L. M., LIU, X., GENCO, C. A. & DARVEAU, R. P. 2006. Hemin-dependent modulation of the lipid A structure of Porphyromonas gingivalis lipopolysaccharide. *Infect Immun,* 74**,** 4474-85.

ALETAHA, D., NEOGI, T., SILMAN, A. J., FUNOVITS, J., FELSON, D. T., BINGHAM, C. O., 3RD, BIRNBAUM, N. S., BURMESTER, G. R., BYKERK, V. P., COHEN, M. D., COMBE, B., COSTENBADER, K. H., DOUGADOS, M., EMERY, P., FERRACCIOLI, G., HAZES, J. M., HOBBS, K., HUIZINGA, T. W., KAVANAUGH, A., KAY, J., KVIEN, T. K., LAING, T., MEASE, P., MENARD, H. A., MORELAND, L. W., NADEN, R. L., PINCUS, T., SMOLEN, J. S., STANISLAWSKA-BIERNAT, E., SYMMONS, D., TAK, P. P., UPCHURCH, K. S., VENCOVSKY, J., WOLFE, F. & HAWKER, G. 2010. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum,* 62**,** 2569-81.

ALPÍZAR-RODRÍGUEZ, D., PLUCHINO, N., CANNY, G., GABAY, C. & FINCKH, A. 2017. The role of female hormonal factors in the development of rheumatoid arthritis. *Rheumatology,* 56**,** 1254-1263.

AMANO, A. 2007. Disruption of epithelial barrier and impairment of cellular function by Porphyromonas gingivalis. *Front Biosci,* 12**,** 3965-74.

AMANO, A. 2010. Bacterial adhesins to host components in periodontitis. *Periodontol 2000,* 52**,** 12-37.

AMANO, A., NAKAGAWA, I., KATAOKA, K., MORISAKI, I. & HAMADA, S. 1999. Distribution of Porphyromonas gingivalis strains with fimA genotypes in periodontitis patients. *J Clin Microbiol,* 37**,** 1426-30.

ANDERS, S., PYL, P. T. & HUBER, W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics,* 31**,** 166-169.

ANZILOTTI, C., PRATESI, F., TOMMASI, C. & MIGLIORINI, P. 2010. Peptidylarginine deiminase 4 and citrullination in health and disease. *Autoimmunity Reviews,* 9**,** 158-160.

ARAUJO, V. M., MELO, I. M. & LIMA, V. 2015. Relationship between Periodontitis and Rheumatoid Arthritis: Review of the Literature. *Mediators Inflamm,* 2015**,** 259074.

ARMITAGE, G. C. & CULLINAN, M. P. 2010. Comparison of the clinical features of chronic and aggressive periodontitis. *Periodontol 2000,* 53**,** 12-27.

ARMSTRONG, C. L., MIRALDA, I., NEFF, A. C., TIAN, S., VASHISHTA, A., PEREZ, L., LE, J., LAMONT, R. J. & URIARTE, S. M. 2016. Filifactor alocis Promotes Neutrophil Degranulation and Chemotactic Activity. *Infection and Immunity,* 84**,** 3423-3433.

BANBULA, A., BUGNO, M., GOLDSTEIN, J., YEN, J., NELSON, D., TRAVIS, J. & POTEMPA, J. 2000. Emerging Family of Proline-Specific Peptidases of Porphyromonas gingivalis: Purification and Characterization of Serine Dipeptidyl Peptidase, a Structural and Functional Homologue of Mammalian Prolyl Dipeptidyl Peptidase IV. *Infection and Immunity,* 68**,** 1176-1182.

BELSTRØM, D., CONSTANCIAS, F., LIU, Y., YANG, L., DRAUTZ-MOSES, D. I., SCHUSTER, S. C., KOHLI, G. S., JAKOBSEN, T. H., HOLMSTRUP, P. & GIVSKOV, M. 2017. Metagenomic and metatranscriptomic analysis of saliva reveals disease-associated microbiota in patients with periodontitis and dental caries. *NPJ Biofilms and Microbiomes,* 3**,** 23.

BICKEL, M. & CIMASONI, G. 1985. The pH of human crevicular fluid measured by a new microanalytical technique. *Journal of Periodontal Research,* 20**,** 35-40.

BICKER, K. L. & THOMPSON, P. R. 2013. The protein arginine deiminases: Structure, function, inhibition, and disease. *Biopolymers,* 99**,** 155-63.

BIELECKA, E., SCAVENIUS, C., KANTYKA, T., JUSKO, M., MIZGALSKA, D., SZMIGIELSKI, B., POTEMPA, B., ENGHILD, J. J., PROSSNITZ, E. R., BLOM, A. M. & POTEMPA, J. 2014. Peptidyl Arginine Deiminase from Porphyromonas gingivalis Abolishes Anaphylatoxin C5a Activity. *The Journal of Biological Chemistry,* 289**,** 32481-32487.

BOS, W. H., WOLBINK, G. J., BOERS, M., TIJHUIS, G. J., DE VRIES, N., VAN DER HORST-BRUINSMA, I. E., TAK, P. P., VAN DE STADT, R. J., VAN DER LAKEN, C. J., DIJKMANS, B. A. & VAN SCHAARDENBURG, D. 2010. Arthritis development in patients with arthralgia is strongly associated with anti-citrullinated protein antibody status: a prospective cohort study. *Ann Rheum Dis,* 69**,** 490-4.

BOSTANCI, N. & BELIBASAKIS, G. N. 2012. Porphyromonas gingivalis: an invasive and evasive opportunistic oral pathogen. *FEMS Microbiol Lett,* 333**,** 1-9.

BRATTHALL, D., PETERSEN, P. E., STJERNSWARD, J. R. & BROWN, L. J. 2006. Oral and Craniofacial Diseases and Disorders. *In:* ND, JAMISON, D. T., BREMAN, J. G., MEASHAM, A. R., ALLEYNE, G., CLAESON, M., EVANS, D. B., JHA, P., MILLS, A. & MUSGROVE, P. (eds.) *Disease Control Priorities in Developing Countries.* Washington (DC): World Bank

The International Bank for Reconstruction and Development/The World Bank Group.

BRIGHT, R., PROUDMAN, S. M., ROSENSTEIN, E. D. & BARTOLD, P. M. 2015. Is there a link between carbamylation and citrullination in periodontal disease

and rheumatoid arthritis? *Med Hypotheses,* 84**,** 570-576.

BRINIG, M. M., LEPP, P. W., OUVERNEY, C. C., ARMITAGE, G. C. & RELMAN, D. A. 2003. Prevalence of bacteria of division TM7 in human subgingival plaque and their association with disease. *Applied and Environmental Microbiology,* 69**,** 1687-1694.

BRINKMANN, V., REICHARD, U., GOOSMANN, C., FAULER, B., UHLEMANN, Y., WEISS, D. S., WEINRAUCH, Y. & ZYCHLINSKY, A. 2004. Neutrophil Extracellular Traps Kill Bacteria. *Science,* 303**,** 1532.

BROGAN, J. M., LALLY, E. T., POULSEN, K., KILIAN, M. & DEMUTH, D. R. 1994. Regulation of Actinobacillus actinomycetemcomitans leukotoxin expression: analysis of the promoter regions of leukotoxic and minimally leukotoxic strains. *Infect Immun,* 62**,** 501-8.

BRUNNER, J., SCHERES, N., EL IDRISSI, N. B., DENG, D. M., LAINE, M. L., VAN WINKELHOFF, A. J. & CRIELAARD, W. 2010. The capsule of Porphyromonas gingivalis reduces the immune response of human gingival fibroblasts. *BMC Microbiology,* 10**,** 5-5.

BUCHFINK, B., XIE, C. & HUSON, D. H. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods,* 12**,** 59-60.

BURMISTRZ, M., RODRIGUEZ MARTINEZ, J. I., KROCHMAL, D., STANIEC, D. & PYRC, K. 2017. Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) RNAs in the Porphyromonas gingivalis CRISPR-Cas I-C System. *Journal of Bacteriology,* 199.

CHAFFEE, B. W. & WESTON, S. J. 2010. Association Between Chronic Periodontal Disease and Obesity: A Systematic Review and Meta-Analysis. *Journal of Periodontology,* 81**,** 1708-1724.

CHENG, Z., MEADE, J., MANKIA, K., EMERY, P. & DEVINE, D. A. 2017. Periodontal disease and periodontal bacteria as triggers for rheumatoid arthritis. *Best Practice & Research Clinical Rheumatology,* 31**,** 19-30.

COLLINGRIDGE, D. S. 2012. A Primer on Quantitized Data Analysis and Permutation Testing. *Journal of Mixed Methods Research,* 7**,** 81-97.

CONRADS, G., PELZ, K., HUGHES, B., SEYFARTH, I. & DEVINE, D. A. 1997. Optimized oligonucleotides for the differentiation of Prevotella intermedia and Prevotella nigrescens. *Oral Microbiol Immunol,* 12**,** 117-20.

CULLINAN, M. P. & SEYMOUR, G. J. 2013. Periodontal disease and systemic illness: will the evidence ever be enough? *Periodontol 2000,* 62**,** 271-86.

CURTIS, M. A., ADUSE-OPOKU, J. & RANGARAJAN, M. 2001. Cysteine Proteases of Porphyromonas Gingivalis. *Critical Reviews in Oral Biology & Medicine,* 12**,** 192-216.

CURTIS, M. A., PERCIVAL, R. S., DEVINE, D., DARVEAU, R. P., COATS, S. R., RANGARAJAN, M., TARELLI, E. & MARSH, P. D. 2011. Temperature-dependent modulation of Porphyromonas gingivalis lipid A structure and interaction with the innate host defenses. *Infect Immun,* 79**,** 1187-93.

D'EMPAIRE, G., T BAER, M. & GIBSON, F. 2006. *The K1 Serotype Capsular Polysaccharide of Porphyromonas gingivalis Elicits Chemokine Production from Murine Macrophages That Facilitates Cell Migration*.

DANILUK, T., TOKAJUK, G., CYLWIK-ROKICKA, D., ROZKIEWICZ, D., ZAREMBA, M. L. & STOKOWSKA, W. 2006. Aerobic and anaerobic bacteria in subgingival and supragingival plaques of adult patients with periodontal disease. *Adv Med Sci,* 51 Suppl 1**,** 81-5.

DARVEAU, R. P., PHAM, T. T., LEMLEY, K., REIFE, R. A., BAINBRIDGE, B. W., COATS, S. R., HOWALD, W. N., WAY, S. S. & HAJJAR, A. M. 2004. Porphyromonas gingivalis lipopolysaccharide contains multiple lipid A species that functionally interact with both toll-like receptors 2 and 4. *Infect Immun,* 72**,** 5041-51.

DE SMIT, M., WESTRA, J., VISSINK, A., DOORNBOS-VAN DER MEER, B., BROUWER, E. & VAN WINKELHOFF, A. J. 2012. Periodontitis in established rheumatoid arthritis patients: a cross-sectional clinical, microbiological and serological study. *Arthritis Res Ther,* 14**,** R222.

DE SMIT, M. J., WESTRA, J., BROUWER, E., JANSSEN, K. M. J., VISSINK, A. & VAN WINKELHOFF, A. J. 2015. Periodontitis and rheumatoid arthritis: what do we know? *J Periodontol,* 86**,** 1013-1019.

DEANE, K. D., DEMORUELLE, M. K., KELMENSON, L. B., KUHN, K. A., NORRIS, J. M. & HOLERS, V. M. 2017. Genetic and environmental risk factors for rheumatoid arthritis. *Best Pract Res Clin Rheumatol,* 31**,** 3-18.

DIAZ, P. I. & ROGERS, A. H. 2004. The effect of oxygen on the growth and physiology of Porphyromonas gingivalis. *Oral Microbiol Immunol,* 19**,** 88-94.

DIVARIS, K., MONDA, K. L., NORTH, K. E., OLSHAN, A. F., REYNOLDS, L. M., HSUEH, W. C., LANGE, E. M., MOSS, K., BARROS, S. P., WEYANT, R. J., LIU, Y., NEWMAN, A. B., BECK, J. D. & OFFENBACHER, S. 2013. Exploring the genetic basis of chronic periodontitis: a genome-wide association study. *Hum Mol Genet,* 22**,** 2312-24.

DIXON, D. R., BAINBRIDGE, B. W. & DARVEAU, R. P. 2004. Modulation of the innate immune response within the periodontium. *Periodontology 2000,* 35**,** 53-74.

DIXON, D. R. & DARVEAU, R. P. 2005. Lipopolysaccharide Heterogeneity: Innate Host Responses to Bacterial Modification of Lipid A Structure. *Journal of Dental Research,* 84**,** 584-595.

DO, T., DEVINE, D. & MARSH, P. D. 2013. Oral biofilms: molecular analysis, challenges, and future prospects in dental diagnostics. *Clinical, Cosmetic and Investigational Dentistry,* 5**,** 11-19.

EGGERT, F. M., DREWELL, L., BIGELOW, J. A., SPECK, J. E. & GOLDNER, M. 1991. The pH of gingival crevices and periodontal pockets in children, teenagers and adults. *Archives of Oral Biology,* 36**,** 233-238.

ELLEN, R. P. & GALIMANAS, V. B. 2005. Spirochetes at the forefront of periodontal infections. *Periodontology 2000,* 38**,** 13-32.

ERIBE, E. R. K. & OLSEN, I. 2017. Leptotrichia species in human infections II. *Journal of Oral Microbiology,* 9**,** 1368848.

ERIKSSON, K., NISE, L., KATS, A., LUTTROPP, E., CATRINA, A. I., ASKLING, J., JANSSON, L., ALFREDSSON, L., KLARESKOG, L., LUNDBERG, K. & YUCEL-LINDBERG, T. 2016. Prevalence of periodontitis in patients with established rheumatoid arthritis: a Swedish population based case-control study. *PLoS One,* 11**,** e0155956.

FINE, D. H., MARKOWITZ, K., FAIRLIE, K., TISCHIO-BERESKI, D., FERRENDIZ, J., FURGANG, D., PASTER, B. J. & DEWHIRST, F. E. 2013. A consortium of Aggregatibacter actinomycetemcomitans, Streptococcus parasanguinis, and Filifactor alocis is present in sites prior to bone loss in a longitudinal study of localized aggressive periodontitis. *J Clin Microbiol,* 51**,** 2850-61.

FISHER, B. A., CARTWRIGHT, A. J., QUIRKE, A. M., DE PABLO, P., ROMAGUERA, D., PANICO, S., MATTIELLO, A., GAVRILA, D.,

NAVARRO, C., SACERDOTE, C., VINEIS, P., TUMINO, R., LAPPIN, D. F., APAZIDOU, D., CULSHAW, S., POTEMPA, J., MICHAUD, D. S., RIBOLI, E. & VENABLES, P. J. 2015. Smoking, *Porphyromonas gingivalis* and the immune response to citrullinated autoantigens before the clinical onset of rheumatoid arthritis in a Southern European nested case–control study. *BMC Musculoskelet Disord,* 16**,** 331.

FOULQUIER, C., SEBBAG, M., CLAVEL, C., CHAPUY-REGAUD, S., AL BADINE, R., MECHIN, M. C., VINCENT, C., NACHAT, R., YAMADA, M., TAKAHARA, H., SIMON, M., GUERRIN, M. & SERRE, G. 2007. Peptidyl arginine deiminase type 2 (PAD-2) and PAD-4 but not PAD-1, PAD-3, and PAD-6 are expressed in rheumatoid arthritis synovium in close association with tissue inflammation. *Arthritis Rheum,* 56**,** 3541-53.

FUGGLE, N. R., SMITH, T. O., KAUL, A. & SOFAT, N. 2016. Hand to mouth: A systematic review and meta-analysis of the association between rheumatoid arthritis and periodontitis. *Front Immunol,* 7**,** 80.

GABARRINI, G., CHLEBOWICZ, M. A., VEGA QUIROZ, M. E., VELOO, A. C. M., ROSSEN, J. W. A., HARMSEN, H. J. M., LAINE, M. L., VAN DIJL, J. M. & VAN WINKELHOFF, A. J. 2018. Conserved Citrullinating Exoenzymes in Porphyromonas Species. *J Dent Res,* 97**,** 556-562.

GABARRINI, G., DE SMIT, M., WESTRA, J., BROUWER, E., VISSINK, A., ZHOU, K., A. ROSSEN, J. W., STOBERNACK, T., VAN DIJL, J. M. & JAN VAN WINKELHOFF, A. 2015. The peptidylarginine deiminase gene is a conserved feature of Porphyromonas gingivalis. *Scientific Reports,* 5**,** 13936.

GAWRON, K., BERETA, G., NOWAKOWSKA, Z., LAZARZ-BARTYZEL, K., LAZARZ, M., SZMIGIELSKI, B., MIZGALSKA, D., BUDA, A., KOZIEL, J., ORUBA, Z., CHOMYSZYN-GAJEWSKA, M. & POTEMPA, J. 2014. Peptidylarginine deiminase from Porphyromonas gingivalis contributes to infection of gingival fibroblasts and induction of prostaglandin E2 -signaling pathway. *Mol Oral Microbiol,* 29**,** 321-32.

GENCO, R. J. & VAN DYKE, T. E. 2010. Reducing the risk of CVD in patients with periodontitis. *Nature Reviews Cardiology,* 7**,** 479.

GERLAG, D. M., RAZA, K., VAN BAARSEN, L. G. M., BROUWER, E., BUCKLEY, C. D., BURMESTER, G. R., GABAY, C., CATRINA, A. I., COPE, A. P., CORNELIS, F., DAHLQVIST, S. R., EMERY, P., EYRE, S., FINCKH, A., GAY, S., HAZES, J. M., VAN DER HELM-VAN MIL, A., HUIZINGA, T. W. J., KLARESKOG, L., KVIEN, T. K., LEWIS, C., MACHOLD, K. P., RÖNNELID, J., SCHAARDENBURG, D. V., SCHETT, G., SMOLEN, J. S., THOMAS, S., WORTHINGTON, J. & TAK, P. P. 2012. EULAR recommendations for terminology and research in individuals at risk of rheumatoid arthritis: report from the Study Group for Risk Factors for Rheumatoid Arthritis. *Annals of the Rheumatic Diseases,* 71**,** 638.

GOEMAERE, S., ACKERMAN, C., GOETHALS, K., DE KEYSER, F., VAN DER STRAETEN, C., VERBRUGGEN, G., MIELANTS, H. & VEYS, E. M. 1990. Onset of symptoms of rheumatoid arthritis in relation to age, sex and menopausal transition. *J Rheumatol,* 17**,** 1620-2.

GREENSTEIN, R. J., SU, L., HAROUTUNIAN, V., SHAHIDI, A. & BROWN, S. T. 2007. On the action of methotrexate and 6-mercaptopurine on M. avium subspecies paratuberculosis. *PLoS One,* 2**,** e161.

GREGERSEN, P. K., SILVER, J. & WINCHESTER, R. J. 1987. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis Rheum,* 30**,** 1205-13.

GRIFFEN, A. L., BEALL, C. J., CAMPBELL, J. H., FIRESTONE, N. D., KUMAR, P. S., YANG, Z. K., PODAR, M. & LEYS, E. J. 2011. Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. *The Isme Journal,* 6**,** 1176.

GRIFFEN, A. L., BEALL, C. J., CAMPBELL, J. H., FIRESTONE, N. D., KUMAR, P. S., YANG, Z. K., PODAR, M. & LEYS, E. J. 2012. Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. *Isme j,* 6**,** 1176-85.

GUO, Y., NGUYEN, K. A. & POTEMPA, J. 2010. Dichotomy of gingipains action as virulence factors: from cleaving substrates with the precision of a surgeon's knife to a meat chopper-like brutal degradation of proteins. *Periodontol 2000,* 54**,** 15-44.

GYORGY, B., TOTH, E., TARCSA, E., FALUS, A. & BUZAS, E. I. 2006. Citrullination: a posttranslational modification in health and disease. *Int J Biochem Cell Biol,* 38**,** 1662-77.

HAJISHENGALLIS, G., ABE, T., MAEKAWA, T., HAJISHENGALLIS, E. & LAMBRIS, J. D. 2013. Role of complement in host-microbe homeostasis of the periodontium. *Seminars in immunology,* 25**,** 65-72.

HAJISHENGALLIS, G., DARVEAU, R. P. & CURTIS, M. A. 2012. The keystone-pathogen hypothesis. *Nat Rev Microbiol,* 10**,** 717-25.

HAJISHENGALLIS, G. & LAMONT, R. J. 2012. Beyond the red complex and into more complexity: the polymicrobial synergy and dysbiosis (PSD) model of periodontal disease etiology. *Mol Oral Microbiol,* 27**,** 409-19.

HAJISHENGALLIS, G., LIANG, S., PAYNE, M. A., HASHIM, A., JOTWANI, R., ESKAN, M. A., MCINTOSH, M. L., ALSAM, A., KIRKWOOD, K. L., LAMBRIS, J. D., DARVEAU, R. P. & CURTIS, M. A. 2011. Low-abundance biofilm species orchestrates inflammatory periodontal disease through the commensal microbiota and complement. *Cell Host Microbe,* 10**,** 497-506.

HAJISHENGALLIS, G., MAEKAWA, T., ABE, T., HAJISHENGALLIS, E. & LAMBRIS, J. D. Complement Involvement in Periodontitis: Molecular Mechanisms and Rational Therapeutic Approaches. *In:* LAMBRIS, J. D., EKDAHL, K. N., RICKLIN, D. & NILSSON, B., eds. Immune Responses to Biosurfaces, 2015// 2015 Cham. Springer International Publishing, 57-74.

HAMADA, S., AMANO, A., KIMURA, S., NAKAGAWA, I., KAWABATA, S. & MORISAKI, I. 1998. The importance of fimbriae in the virulence and ecology of some oral bacteria. *Oral Microbiology and Immunology,* 13**,** 129-138.

HAN, N., WHITLOCK, J. & PROGULSKE-FOX, A. 1996. The hemagglutinin gene A (hagA) of Porphyromonas gingivalis 381 contains four large, contiguous, direct repeats. *Infection and Immunity,* 64**,** 4000-4007.

HANIOKA, T., MATSUSE, R., SHIGEMOTO, Y., OJIMA, M. & SHIZUKUISHI, S. 2005. Relationship between periodontal disease status and combination of biochemical assays of gingival crevicular fluid. *J Periodontal Res,* 40**,** 331-8.

HASHIMOTO, M., YAMAZAKI, T., HAMAGUCHI, M., MORIMOTO, T., YAMORI, M., ASAI, K., ISOBE, Y., FURU, M., ITO, H., FUJII, T., TERAO, C., MORI, M., MATSUO, T., YOSHITOMI, H., YAMAMOTO, K., YAMAMOTO, W., BESSHO, K. & MIMORI, T. 2015. Periodontitis and *Porphyromonas gingivalis* in preclinical stage of arthritis patients. *PLoS One,* 10**,** e0122121.

HEDSTROM, A. K., STAWIARZ, L., KLARESKOG, L. & ALFREDSSON, L. 2018. Smoking and susceptibility to rheumatoid arthritis in a Swedish population-based case-control study. *Eur J Epidemiol*.

HEIDARI, B. 2011. Rheumatoid Arthritis: Early diagnosis and treatment outcomes. *Caspian J Intern Med,* 2**,** 161-70.

HERBERT, B. A., NOVINCE, C. M. & KIRKWOOD, K. L. 2016. Aggregatibacter actinomycetemcomitans, a potent immunoregulator of the periodontal host defense system and alveolar bone homeostasis. *Molecular oral microbiology,* 31**,** 207-227.

HITCHON, C. A., CHANDAD, F., FERUCCI, E. D., WILLEMZE, A., IOAN-FACSINAY, A., VAN DER WOUDE, D., MARKLAND, J., ROBINSON, D., ELIAS, B., NEWKIRK, M., TOES, R. M., HUIZINGA, T. W. & EL-GABALAWY, H. S. 2010. Antibodies to porphyromonas gingivalis are associated with anticitrullinated protein antibodies in patients with rheumatoid arthritis and their relatives. *J Rheumatol,* 37**,** 1105-12.

HOGLUND ABERG, C., KWAMIN, F., CLAESSON, R., DAHLEN, G., JOHANSSON, A. & HAUBEK, D. 2014. Progression of attachment loss is strongly associated with presence of the JP2 genotype of Aggregatibacter actinomycetemcomitans: a prospective cohort study of a young adolescent population. *J Clin Periodontol,* 41**,** 232-41.

HONG, B.-Y., FURTADO ARAUJO, M. V., STRAUSBAUGH, L. D., TERZI, E., IOANNIDOU, E. & DIAZ, P. I. 2015. Microbiome Profiles in Periodontitis in Relation to Host and Disease Characteristics. *PLoS ONE,* 10**,** e0127077.

HOSKISSON, P. A. & HOBBS, G. 2005. Continuous culture--making a comeback? *Microbiology,* 151**,** 3153-9.

HOW, K. Y., PENG SONG, K. & GAN CHAN, K. 2016. *Porphyromonas gingivalis: An Overview of Periodontopathic Pathogen Below the Gum Line*.

HUNT, L. & EMERY, P. 2014. Defining populations at risk of rheumatoid arthritis: the first steps to prevention. *Nat Rev Rheumatol,* 10**,** 521-30.

IMAMURA, T. 2003. The role of gingipains in the pathogenesis of periodontal disease. *J Periodontol,* 74**,** 111-8.

IMAMURA, T., POTEMPA, J., PIKE, R. N., MOORE, J. N., BARTON, M. H. & TRAVIS, J. 1995. Effect of free and vesicle-bound cysteine proteinases of Porphyromonas gingivalis on plasma clot formation: implications for bleeding tendency at periodontitis sites. *Infect Immun,* 63**,** 4877-82.

JOHANSSON, L., SHERINA, N., KHARLAMOVA, N., POTEMPA, B., LARSSON, B., ISRAELSSON, L., POTEMPA, J., RANTAPAA-DAHLQVIST, S. & LUNDBERG, K. 2016. Concentration of antibodies against *Porphyromonas gingivalis* is increased before the onset of symptoms of rheumatoid arthritis. *Arthritis Res Ther,* 18**,** 201.

JOLIVET-GOUGEON, A., TAMANAI-SHACOORI, Z., DESBORDES, L., BURGGRAEVE, N., CORMIER, M. & BONNAURE-MALLET, M. 2004. Genetic Analysis of an Ambler Class A Extended-Spectrum Beta-Lactamase from Capnocytophaga ochracea. *Journal of Clinical Microbiology,* 42**,** 888-890.

JU, F., XIA, Y., GUO, F., WANG, Z. & ZHANG, T. 2014. Taxonomic relatedness shapes bacterial assembly in activated sludge of globally distributed wastewater treatment plants. *Environ Microbiol,* 16**,** 2421-32.

KADOWAKI, T., YUKITAKE, H., NAITO, M., SATO, K., KIKUCHI, Y., KONDO, Y., SHOJI, M. & NAKAYAMA, K. 2016. A two-component system regulates gene expression of the type IX secretion component proteins via an ECF sigma factor. *Sci Rep,* 6**,** 23288.

KALLBERG, H., DING, B., PADYUKOV, L., BENGTSSON, C., RONNELID, J., KLARESKOG, L. & ALFREDSSON, L. 2011. Smoking is a major

preventable risk factor for rheumatoid arthritis: estimations of risks after various exposures to cigarette smoke. *Ann Rheum Dis,* 70**,** 508-11.

KAMAGUCHI, A., OHYAMA, T., SAKAI, E., NAKAMURA, R., WATANABE, T., BABA, H. & NAKAYAMA, K. 2003. Adhesins encoded by the gingipain genes of Porphyromonas gingivalis are responsible for co-aggregation with Prevotella intermedia. *Microbiology,* 149**,** 1257-64.

KASSEBAUM, N. J., BERNABÉ, E., DAHIYA, M., BHANDARI, B., MURRAY, C. J. L. & MARCENES, W. 2014. Global Burden of Severe Periodontitis in 1990-2010: A Systematic Review and Meta-regression. *Journal of Dental Research,* 93**,** 1045-1053.

KATO, H., TAGUCHI, Y., TOMINAGA, K., UMEDA, M. & TANAKA, A. 2014. Porphyromonas gingivalis LPS inhibits osteoblastic differentiation and promotes pro-inflammatory cytokine production in human periodontal ligament stem cells. *Arch Oral Biol,* 59**,** 167-75.

KAUR, S., BRIGHT, R., PROUDMAN, S. M. & BARTOLD, P. M. 2014. Does periodontal treatment influence clinical and biochemical measures for rheumatoid arthritis? A systematic review and meta-analysis. *Semin Arthritis Rheum,* 44**,** 113-22.

KENNEY, E. B. & ASH, M. M., JR. 1969. Oxidation reduction potential of developing plaque, periodontal pockets and gingival sulci. *J Periodontol,* 40**,** 630-3.

KILIAN, M., FRANDSEN, E. V. G., HAUBEK, D. & POULSEN, K. 2006. The etiology of periodontal disease revisited by population genetic analysis. *Periodontology 2000,* 42**,** 158-179.

KIRKWOOD, K. L., CIRELLI, J. A., ROGERS, J. E. & GIANNOBILE, W. V. 2007. Novel host response therapeutic approaches to treat periodontal diseases. *Periodontology 2000,* 43**,** 294-315.

KNIPP, M. & VASAK, M. 2000. A colorimetric 96-well microtiter plate assay for the determination of enzymatically formed citrulline. *Anal Biochem,* 286**,** 257-64.

KOLENBRANDER, P. E., PALMER JR, R. J., PERIASAMY, S. & JAKUBOVICS, N. S. 2010. Oral multispecies biofilm development and the key role of cell–cell distance. *Nature Reviews Microbiology,* 8**,** 471.

KOLENBRANDER, P. E., PALMER, R. J., RICKARD, A. H., JAKUBOVICS, N. S., CHALMERS, N. I. & DIAZ, P. I. 2006. Bacterial interactions and successions during plaque development. *Periodontology 2000,* 42**,** 47-79.

KONIG, M. F., ABUSLEME, L., REINHOLDT, J., PALMER, R. J., TELES, R. P., SAMPSON, K., ROSEN, A., NIGROVIC, P. A., SOKOLOVE, J., GILES, J. T., MOUTSOPOULOS, N. M. & ANDRADE, F. 2016. *Aggregatibacter actinomycetemcomitans*-induced hypercitrullination links periodontal infection to autoimmunity in rheumatoid arthritis. *Sci Transl Med,* 8**,** 369ra176.

KONIG, M. F. & ANDRADE, F. 2016. A Critical Reappraisal of Neutrophil Extracellular Traps and NETosis Mimics Based on Differential Requirements for Protein Citrullination. *Front Immunol,* 7**,** 461.

KONIG, M. F., PARACHA, A. S., MONI, M., BINGHAM, C. O., 3RD & ANDRADE, F. 2015. Defining the role of *Porphyromonas gingivalis* peptidylarginine deiminase (PPAD) in rheumatoid arthritis through the study of PPAD biology. *Ann Rheum Dis,* 74**,** 2054-61.

KORO, C., BIELECKA, E., DAHL-KNUDSEN, A., ENGHILD JAN, J., SCAVENIUS, C., BRUN JOHAN, G., BINDER, V., HELLVARD, A., BERGUM, B., JONSSON, R., POTEMPA, J., BLOM ANNA, M. & MYDEL, P. 2014. Carbamylation of immunoglobulin abrogates activation of the

classical complement pathway. *European Journal of Immunology,* 44**,** 3403-3412.

KUMAR, P. S., GRIFFEN, A. L., BARTON, J. A., PASTER, B. J., MOESCHBERGER, M. L. & LEYS, E. J. 2003. New bacterial species associated with chronic periodontitis. *Journal of Dental Research,* 82**,** 338-344.

KUMAR, P. S., LEYS, E. J., BRYK, J. M., MARTINEZ, F. J., MOESCHBERGER, M. L. & GRIFFEN, A. L. 2006. Changes in periodontal health status are associated with bacterial community shifts as assessed by quantitative 16S cloning and sequencing. *J Clin Microbiol,* 44**,** 3665-73.

LAINE, M. L., APPELMELK, B. J. & VAN WINKELHOFF, A. J. 1996. Novel polysaccharide capsular serotypes in Porphyromonas gingivalis. *J Periodontal Res,* 31**,** 278-84.

LAINE, M. L. & VAN WINKELHOFF, A. J. 1998. Virulence of six capsular serotypes of Porphyromonas gingivalis in a mouse model. *Oral Microbiol Immunol,* 13**,** 322-5.

LALLA, E. & PAPAPANOU, P. N. 2011. Diabetes mellitus and periodontitis: a tale of two common interrelated diseases. *Nature Reviews Endocrinology,* 7**,** 738.

LASICA, A. M., KSIAZEK, M., MADEJ, M. & POTEMPA, J. 2017. The Type IX Secretion System (T9SS): Highlights and Recent Insights into Its Structure and Function. *Frontiers in Cellular and Infection Microbiology,* 7**,** 215.

LAUGISCH, O., WONG, A., SROKA, A., KANTYKA, T., KOZIEL, J., NEUHAUS, K., SCULEAN, A., VENABLES, P. J., POTEMPA, J., MÖLLER, B. & EICK, S. 2016. Citrullination in the periodontium—a possible link between periodontitis and rheumatoid arthritis. *Clin Oral Investig,* 20**,** 675-683.

LEKE, N., GRENIER, D., GOLDNER, M. & MAYRAND, D. 1999. Effects of hydrogen peroxide on growth and selected properties of Porphyromonas gingivalis. *FEMS Microbiol Lett,* 174**,** 347-53.

LEPP, P. W., BRINIG, M. M., OUVERNEY, C. C., PALM, K., ARMITAGE, G. C. & RELMAN, D. A. 2004. Methanogenic Archaea and human periodontal disease. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 6176-6181.

LEWIS, J. P., IYER, D. & ANAYA-BERGMAN, C. 2009. Adaptation of Porphyromonas gingivalis to microaerophilic conditions involves increased consumption of formate and reduced utilization of lactate. *Microbiology,* 155**,** 3758-3774.

LI, D., LUO, R., LIU, C. M., LEUNG, C. M., TING, H. F., SADAKANE, K., YAMASHITA, H. & LAM, T. W. 2016. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods,* 102**,** 3-11.

LI, H. & DURBIN, R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics,* 26**,** 589-95.

LI, J., ZHAO, F., WANG, Y., CHEN, J., TAO, J., TIAN, G., WU, S., LIU, W., CUI, Q., GENG, B., ZHANG, W., WELDON, R., AUGUSTE, K., YANG, L., LIU, X., CHEN, L., YANG, X., ZHU, B. & CAI, J. 2017. Gut microbiota dysbiosis contributes to the development of hypertension. *Microbiome,* 5**,** 14.

LI, N. & COLLYER, C. A. 2011. Gingipains from Porphyromonas gingivalis – Complex domain structures confer diverse functions. *European Journal of Microbiology & Immunology,* 1**,** 41-58.

LIVAK, K. J. & SCHMITTGEN, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.

Methods, 25**,** 402-8.

LOPEZ-OLIVA, I., PAROPKARI, A. D., SARASWAT, S., SERBAN, S., YONEL, Z., SHARMA, P., DE PABLO, P., RAZA, K., FILER, A., CHAPPLE, I., DIETRICH, T., GRANT, M. M. & KUMAR, P. S. 2018. Dysbiotic Subgingival Microbial Communities in Periodontally Healthy Patients With Rheumatoid Arthritis. *Arthritis Rheumatol,* 70**,** 1008-1013.

LOVE, M. I., HUBER, W. & ANDERS, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol,* 15**,** 550.

LUNDBERG, K., WEGNER, N., YUCEL-LINDBERG, T. & VENABLES, P. J. 2010. Periodontitis in RA—the citrullinated enolase connection. *Nature Reviews Rheumatology,* 6**,** 727.

MANKIA, K. & EMERY, P. 2016. A new window of opportunity in rheumatoid arthritis: targeting at-risk individuals. *Curr Opin Rheumatol,* 28**,** 260-6.

MARESZ, K. J., HELLVARD, A., SROKA, A., ADAMOWICZ, K., BIELECKA, E., KOZIEL, J., GAWRON, K., MIZGALSKA, D., MARCINSKA, K. A., BENEDYK, M., PYRC, K., QUIRKE, A. M., JONSSON, R., ALZABIN, S., VENABLES, P. J., NGUYEN, K. A., MYDEL, P. & POTEMPA, J. 2013. *Porphyromonas gingivalis* facilitates the development and progression of destructive arthritis through its unique bacterial peptidylarginine deiminase (PAD). *PLoS Pathog,* 9**,** e1003627.

MAROTTE, H., FARGE, P., GAUDIN, P., ALEXANDRE, C., MOUGIN, B. & MIOSSEC, P. 2006. The association between periodontal disease and joint destruction in rheumatoid arthritis extends the link between the HLA-DR shared epitope and severity of bone destruction. *Annals of the Rheumatic Diseases,* 65**,** 905.

MARSH, P. D. 1994. Microbial ecology of dental plaque and its significance in health and disease. *Adv Dent Res,* 8**,** 263-71.

MARSH, P. D. 2003. Are dental diseases examples of ecological catastrophes? *Microbiology-Sgm,* 149**,** 279-294.

MARSH, P. D. 2005. Dental plaque: biological significance of a biofilm and community life-style. *Journal of Clinical Periodontology,* 32**,** 7-15.

MARSH, P. D. 2006. Dental plaque as a biofilm and a microbial community – implications for health and disease. *BMC Oral Health,* 6**,** S14-S14.

MARSH, P. D. 2016. *Marsh and Martin's Oral Microbiology*, Elsevier.

MARSH, P. D. & DEVINE, D. A. 2011. How is the development of dental biofilms influenced by the host? *Journal of Clinical Periodontology,* 38**,** 28-35.

MARSH, P. D., MCDERMID, A. S., MCKEE, A. S. & BASKERVILLE, A. 1994. The effect of growth rate and haemin on the virulence and proteolytic activity of Porphyromonas gingivalis W50. *Microbiology,* 140 (Pt 4)**,** 861-5.

MARSH, P. D., MOTER, A. & DEVINE, D. A. 2011. Dental plaque biofilms: communities, conflict and control. *Periodontol 2000,* 55**,** 16-35.

MARTIN, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011,* 17.

MATTEY, D. L., DAWES, P. T., GONZALEZ-GAY, M. A., GARCIA-PORRUA, C., THOMSON, W., HAJEER, A. H. & OLLIER, W. E. 2001. HLA-DRB1 alleles encoding an aspartic acid at position 70 protect against development of rheumatoid arthritis. *J Rheumatol,* 28**,** 232-9.

MCGRAW, W. T., POTEMPA, J., FARLEY, D. & TRAVIS, J. 1999. Purification, characterization, and sequence analysis of a potential virulence factor from Porphyromonas gingivalis, peptidylarginine deiminase. *Infect Immun,* 67**,** 3248-56.

MCKEE, A. S., MCDERMID, A. S., WAIT, R., BASKERVILLE, A. & MARSH, P. D. 1988. Isolation of colonial variants of Bacteroides gingivalis W50 with a reduced virulence. *J Med Microbiol,* 27**,** 59-64.

MCMURDIE, P. J. & HOLMES, S. 2014. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol,* 10**,** e1003531.

MIKULS TED, R., THIELE GEOFFREY, M., DEANE KEVIN, D., PAYNE JEFFREY, B., O'DELL JAMES, R., YU, F., SAYLES, H., WEISMAN MICHAEL, H., GREGERSEN PETER, K., BUCKNER JANE, H., KEATING RICHARD, M., DERBER LEZLIE, A., ROBINSON WILLIAM, H., HOLERS, V. M. & NORRIS JILL, M. 2012. Porphyromonas gingivalis and disease-related autoantibodies in individuals at increased risk of rheumatoid arthritis. *Arthritis & Rheumatism,* 64**,** 3522-3530.

MIKULS, T. R., PAYNE, J. B., DEANE, K. D. & THIELE, G. M. 2016. Autoimmunity of the lung and oral mucosa in a multisystem inflammatory disease: The spark that lights the fire in rheumatoid arthritis? *J Allergy Clin Immunol* 137**,** 28-34.

MIKULS, T. R., THIELE, G. M., DEANE, K. D., PAYNE, J. B., O'DELL, J. R., YU, F., SAYLES, H., WEISMAN, M. H., GREGERSEN, P. K., BUCKNER, J. H., KEATING, R. M., DERBER, L. A., ROBINSON, W. H., HOLERS, V. M. & NORRIS, J. M. 2012. *Porphyromonas gingivalis* and disease-related autoantibodies in individuals at increased risk of rheumatoid arthritis. *Arthritis Rheum,* 64**,** 3522-30.

MONTGOMERY, A. B., KOPEC, J., SHRESTHA, L., THEZENAS, M. L., BURGESS-BROWN, N. A., FISCHER, R., YUE, W. W. & VENABLES, P. J. 2016. Crystal structure of Porphyromonas gingivalis peptidylarginine deiminase: implications for autoimmunity in rheumatoid arthritis. *Ann Rheum Dis,* 75**,** 1255-61.

MORGAN, J. L., DARLING, A. E. & EISEN, J. A. 2010. Metagenomic Sequencing of an In Vitro-Simulated Microbial Community. *PLOS ONE,* 5**,** e10209.

NAKAGAWA, I., AMANO, A., KIMURA, R. K., NAKAMURA, T., KAWABATA, S. & HAMADA, S. 2000. Distribution and Molecular Characterization ofPorphyromonas gingivalis Carrying a New Type offimA Gene. *Journal of Clinical Microbiology,* 38**,** 1909-1914.

NAKAGAWA, I., AMANO, A., OHARA-NEMOTO, Y., ENDOH, N., MORISAKI, I., KIMURA, S., KAWABATA, S. & HAMADA, S. 2002. Identification of a new variant of fimA gene of Porphyromonas gingivalis and its distribution in adults and disabled populations with periodontitis. *Journal of Periodontal Research,* 37**,** 425-432.

NAKAO, R., TAKASHIBA, S., KOSONO, S., YOSHIDA, M., WATANABE, H., OHNISHI, M. & SENPUKU, H. 2014. Effect of Porphyromonas gingivalis outer membrane vesicles on gingipain-mediated detachment of cultured oral epithelial cells and immune responses. *Microbes and Infection,* 16**,** 6-16.

NAKAYAMA, K., RATNAYAKE, D. B., TSUKUBA, T., KADOWAKI, T., YAMAMOTO, K. & FUJIMURA, S. 1998. Haemoglobin receptor protein is intragenically encoded by the cysteine proteinase-encoding genes and the haemagglutinin-encoding gene of Porphyromonas gingivalis. *Mol Microbiol,* 27**,** 51-61.

NAKAYAMA, K., YOSHIMURA, F., KADOWAKI, T. & YAMAMOTO, K. 1996. Involvement of arginine-specific cysteine proteinase (Arg-gingipain) in fimbriation of Porphyromonas gingivalis. *J Bacteriol,* 178**,** 2818-24.

NAM, J. L., HUNT, L., HENSOR, E. M. A. & EMERY, P. 2016. Enriching case

selection for imminent RA: the use of anti-CCP antibodies in individuals with new non-specific musculoskeletal symptoms – a cohort study. *Annals of the Rheumatic Diseases,* 75**,** 1452.

NATIVEL, B., COURET, D., GIRAUD, P., MEILHAC, O., D'HELLENCOURT, C. L., VIRANAÏCKEN, W. & DA SILVA, C. R. 2017. Porphyromonas gingivalis lipopolysaccharides act exclusively through TLR4 with a resilience between mouse and human. *Scientific Reports,* 7**,** 15789.

NEMOTO, T. K. & OHARA-NEMOTO, Y. 2016. Exopeptidases and gingipains in Porphyromonas gingivalis as prerequisites for its amino acid metabolism. *Japanese Dental Science Review,* 52**,** 22-29.

NIELEN, M. M., VAN SCHAARDENBURG, D., REESINK, H. W., VAN DE STADT, R. J., VAN DER HORST-BRUINSMA, I. E., DE KONING, M. H., HABIBUW, M. R., VANDENBROUCKE, J. P. & DIJKMANS, B. A. 2004. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. *Arthritis Rheum,* 50**,** 380-6.

NISHIMATA, H., OHARA-NEMOTO, Y., BABA, T. T., HOSHINO, T., FUJIWARA, T., SHIMOYAMA, Y., KIMURA, S. & NEMOTO, T. K. 2014. Identification of Dipeptidyl-Peptidase (DPP)5 and DPP7 in Porphyromonas endodontalis, Distinct from Those in Porphyromonas gingivalis. *PLOS ONE,* 9**,** e114221.

ODA, H., SAIKI, K., TONOSAKI, M., YAJIMA, A. & KONISHI, K. 2009. Participation of the secreted dipeptidyl and tripeptidyl aminopeptidases in asaccharolytic growth of Porphyromonas gingivalis. *J Periodontal Res,* 44**,** 362-7.

OHARA-NEMOTO, Y., ROUF, S. M., NAITO, M., YANASE, A., TETSUO, F., ONO, T., KOBAYAKAWA, T., SHIMOYAMA, Y., KIMURA, S., NAKAYAMA, K., SAIKI, K., KONISHI, K. & NEMOTO, T. K. 2014. Identification and characterization of prokaryotic dipeptidyl-peptidase 5 from Porphyromonas gingivalis. *J Biol Chem,* 289**,** 5436-48.

OHARA-NEMOTO, Y., SHIMOYAMA, Y., KIMURA, S., KON, A., HARAGA, H., ONO, T. & NEMOTO, T. K. 2011. Asp- and Glu-specific Novel Dipeptidyl Peptidase 11 of Porphyromonas gingivalis Ensures Utilization of Proteinaceous Energy Sources. *Journal of Biological Chemistry,* 286**,** 38115-38127.

OLCZAK, T., SIMPSON, W., LIU, X. & GENCO, C. A. 2005. Iron and heme utilization in Porphyromonas gingivalis. *FEMS Microbiol Rev,* 29**,** 119-44.

OLCZAK, T., SOSICKA, P. & OLCZAK, M. 2015. HmuY is an important virulence factor for Porphyromonas gingivalis growth in the heme-limited host environment and infection of macrophages. *Biochem Biophys Res Commun,* 467**,** 748-53.

OLSEN, I. & SINGHRAO, S. K. 2018. Importance of heterogeneity in Porhyromonas gingivalis lipopolysaccharide lipid A in tissue specific inflammatory signalling. *Journal of Oral Microbiology,* 10**,** 1440128.

OSIRO, D., FILHO, R. B., ASSIS, O. B. G., JORGE, L. A. D. C. & COLNAGO, L. A. 2012. Measuring bacterial cells size with AFM. *Brazilian journal of microbiology : [publication of the Brazilian Society for Microbiology],* 43**,** 341-347.

PADYUKOV, L., SILVA, C., STOLT, P., ALFREDSSON, L. & KLARESKOG, L. 2004. A gene-environment interaction between smoking and shared epitope genes in HLA-DR provides a high risk of seropositive rheumatoid arthritis. *Arthritis Rheum,* 50**,** 3085-92.

PARAMAESVARAN, M., NGUYEN, K. A., CALDON, E., MCDONALD, J. A.,

NAJDI, S., GONZAGA, G., LANGLEY, D. B., DECARLO, A., CROSSLEY, M. J., HUNTER, N. & COLLYER, C. A. 2003. Porphyrin-mediated cell surface heme capture from hemoglobin by Porphyromonas gingivalis. *J Bacteriol,* 185**,** 2528-37.

PARK, O. J., YI, H., JEON, J. H., KANG, S. S., KOO, K. T., KUM, K. Y., CHUN, J., YUN, C. H. & HAN, S. H. 2015. Pyrosequencing Analysis of Subgingival Microbiota in Distinct Periodontal Conditions. *J Dent Res,* 94**,** 921-7.

PASTER, B. J., OLSEN, I., AAS, J. A. & DEWHIRST, F. E. 2006. The breadth of bacterial diversity in the human periodontal pocket and other oral sites. *Periodontology 2000,* 42**,** 80-87.

PATHIRANA, R. D., O'BRIEN-SIMPSON, N. M. & REYNOLDS, E. C. 2010. Host immune responses to Porphyromonas gingivalis antigens. *Periodontol 2000,* 52**,** 218-37.

PATHIRANA, R. D., O'BRIEN-SIMPSON, N. M., VEITH, P. D., RILEY, P. F. & REYNOLDS, E. C. 2006. Characterization of proteinase-adhesin complexes of Porphyromonas gingivalis. *Microbiology,* 152**,** 2381-94.

PAYET, J., GOULVESTRE, C., BIALE, L., AVOUAC, J., WIPFF, J., JOB-DESLANDRE, C., BATTEUX, F., DOUGADOS, M., KAHAN, A. & ALLANORE, Y. 2014. Anticyclic citrullinated peptide antibodies in rheumatoid and nonrheumatoid rheumatic disorders: experience with 1162 patients. *J Rheumatol,* 41**,** 2395-402.

PEREIRA, M. B., WALLROTH, M., JONSSON, V. & KRISTIANSSON, E. 2018. Comparison of normalization methods for the analysis of metagenomic gene abundance data. *BMC Genomics,* 19**,** 274.

PÉREZ-CHAPARRO, P. J., GONÇALVES, C., FIGUEIREDO, L. C., FAVERI, M., LOBÃO, E., TAMASHIRO, N., DUARTE, P. & FERES, M. 2014. Newly Identified Pathogens Associated with Periodontitis: A Systematic Review. *Journal of Dental Research,* 93**,** 846-858.

PETERSEN, P. E. & OGAWA, H. 2012. The global burden of periodontal disease: towards integration with chronic disease prevention and control. *Periodontol 2000,* 60**,** 15-39.

PIAU, C., ARVIEUX, C., BONNAURE-MALLET, M. & JOLIVET-GOUGEON, A. 2013. Capnocytophaga spp. involvement in bone infections: a review. *International Journal of Antimicrobial Agents,* 41**,** 509-515.

PIHLSTROM, B. L., MICHALOWICZ, B. S. & JOHNSON, N. W. 2005. Periodontal diseases. *Lancet,* 366**,** 1809-20.

PISANO, E., CABRAS, T., MONTALDO, C., PIRAS, V., INZITARI, R., OLMI, C., CASTAGNOLA, M. & MESSANA, I. 2005. Peptides of human gingival crevicular fluid determined by HPLC-ESI-MS. *European Journal of Oral Sciences,* 113**,** 462-468.

POTEMPA, J., PIKE, R. & TRAVIS, J. 1995. The multiple forms of trypsin-like activity present in various strains of Porphyromonas gingivalis are due to the presence of either Arg-gingipain or Lys-gingipain. *Infection and Immunity,* 63**,** 1176-1182.

POTEMPA, J., PIKE, R. & TRAVIS, J. 1997. Titration and mapping of the active site of cysteine proteinases from Porphyromonas gingivalis (gingipains) using peptidyl chloromethanes. *Biol Chem,* 378**,** 223-30.

POTEMPA, J. & TRAVIS, J. 1996. Porphyromonas gingivalis proteinases in periodontitis, a review. *Acta Biochim Pol,* 43**,** 455-65.

PYRC, K., MILEWSKA, A., KANTYKA, T., SROKA, A., MARESZ, K., KOZIEL, J., NGUYEN, K. A., ENGHILD, J. J., KNUDSEN, A. D. & POTEMPA, J.

2013. Inactivation of epidermal growth factor by Porphyromonas gingivalis as a potential mechanism for periodontal tissue damage. *Infect Immun,* 81**,** 55-64.

QUIRKE, A. M., LUGLI, E. B., WEGNER, N., HAMILTON, B. C., CHARLES, P., CHOWDHURY, M., YTTERBERG, A. J., ZUBAREV, R. A., POTEMPA, J., CULSHAW, S., GUO, Y., FISHER, B. A., THIELE, G., MIKULS, T. R. & VENABLES, P. J. 2014. Heightened immune response to autocitrullinated Porphyromonas gingivalis peptidylarginine deiminase: a potential mechanism for breaching immunologic tolerance in rheumatoid arthritis. *Ann Rheum Dis,* 73**,** 263-9.

RANGARAJAN, M., ADUSE-OPOKU, J., SLANEY, J. M., YOUNG, K. A. & CURTIS, M. A. 1997. The prpR1 and prR2 arginine-specific protease genes of Porphyromonas gingivalis W50 produce five biochemically distinct enzymes. *Mol Microbiol,* 23**,** 955-65.

RANJAN, R., RANI, A., METWALLY, A., MCGEE, H. S. & PERKINS, D. L. 2016. Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochem Biophys Res Commun,* 469**,** 967-77.

RANTAPÄÄ-DAHLQVIST, S., DE JONG, B. A. W., BERGLIN, E., HALLMANS, G., WADELL, G., STENLUND, H., SUNDIN, U. & VAN VENROOIJ, W. J. 2003. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. *Arthritis & Rheumatism,* 48**,** 2741-2749.

RAYCHAUDHURI, S., SANDOR, C., STAHL, E. A., FREUDENBERG, J., LEE, H.-S., JIA, X., ALFREDSSON, L., PADYUKOV, L., KLARESKOG, L., WORTHINGTON, J., SIMINOVITCH, K. A., BAE, S.-C., PLENGE, R. M., GREGERSEN, P. K. & DE BAKKER, P. I. W. 2012. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nature Genetics,* 44**,** 291.

REA, D., VAN ELZEN, R., DE WINTER, H., VAN GOETHEM, S., LANDUYT, B., LUYTEN, W., SCHOOFS, L., VAN DER VEKEN, P., AUGUSTYNS, K., DE MEESTER, I., FULOP, V. & LAMBEIR, A. M. 2017. Crystal structure of Porphyromonas gingivalis dipeptidyl peptidase 4 and structure-activity relationships based on inhibitor profiling. *Eur J Med Chem,* 139**,** 482-491.

RIVIERE, G. R., DEROUEN, T. A., KAY, S. L., AVERA, S. P., STOUFFER, V. K. & HAWKINS, N. R. 1997. Association of oral spirochetes from sites of periodontal health with development of periodontitis. *J Periodontol,* 68**,** 1210-4.

ROLAIN, J. M., COLSON, P. & RAOULT, D. 2007. Recycling of chloroquine and its hydroxyl analogue to face bacterial, fungal and viral infections in the 21st century. *Int J Antimicrob Agents,* 30**,** 297-308.

ROMERO-SANCHEZ, C., RODRIGUEZ, C., SANTOS-MORENO, P., MESA, A. M., LAFAURIE, G. I., GIRALDO, Q. S., DE-AVILA, J., CASTILLO, D. M., DURAN, M., CHALEM, P. C., BELLO GUALTERO, J. M. & VALLE-ONATE, R. 2017. Is the Treatment with Biological or Non-biological DMARDS a Modifier of Periodontal Condition in Patients with Rheumatoid Arthritis? *Curr Rheumatol Rev,* 13**,** 139-151.

ROPER, J. M., RAUX, E., BRINDLEY, A. A., SCHUBERT, H. L., GHARBIA, S. E., SHAH, H. N. & WARREN, M. J. 2000. The enigma of cobalamin (Vitamin B12) biosynthesis in Porphyromonas gingivalis. Identification and characterization of a functional corrin pathway. *J Biol Chem,* 275**,** 40316-23.

ROUF, S. M., OHARA-NEMOTO, Y., HOSHINO, T., FUJIWARA, T., ONO, T. &

NEMOTO, T. K. 2013. Discrimination based on Gly and Arg/Ser at position 673 between dipeptidyl-peptidase (DPP) 7 and DPP11, widely distributed DPPs in pathogenic and environmental gram-negative bacteria. *Biochimie,* 95**,** 824-32.

SAKAMOTO, Y., SUZUKI, Y., IIZUKA, I., TATEOKA, C., ROPPONGI, S., FUJIMOTO, M., INAKA, K., TANAKA, H., YAMADA, M., OHTA, K., GOUDA, H., NONAKA, T., OGASAWARA, W. & TANAKA, N. 2015. Structural and mutational analyses of dipeptidyl peptidase 11 from Porphyromonas gingivalis reveal the molecular basis for strict substrate specificity. *Scientific Reports,* 5**,** 11151.

SALONEN, A., SALOJÄRVI, J., LAHTI, L. & DE VOS, W. M. 2012. The adult intestinal core microbiota is determined by analysis depth and health status. *Clinical Microbiology and Infection,* 18**,** 16-20.

SATO, K., YUKITAKE, H., NARITA, Y., SHOJI, M., NAITO, M. & NAKAYAMA, K. 2013. Identification of Porphyromonas gingivalis proteins secreted by the Por secretion system. *FEMS Microbiol Lett,* 338**,** 68-76.

SCHATZLE, M., FADDY, M. J., CULLINAN, M. P., SEYMOUR, G. J., LANG, N. P., BURGIN, W., ANERUD, A., BOYSEN, H. & LOE, H. 2009. The clinical course of chronic periodontitis: V. Predictive factors in periodontal disease. *J Clin Periodontol,* 36**,** 365-71.

SCHELLEKENS, G. A., VISSER, H., DE JONG, B. A., VAN DEN HOOGEN, F. H., HAZES, J. M., BREEDVELD, F. C. & VAN VENROOIJ, W. J. 2000. The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. *Arthritis Rheum,* 43**,** 155-63.

SCHER, J. U., UBEDA, C., EQUINDA, M., KHANIN, R., BUISCHI, Y., VIALE, A., LIPUMA, L., ATTUR, M., PILLINGER, M. H., WEISSMANN, G., LITTMAN, D. R., PAMER, E. G., BRETZ, W. A. & ABRAMSON, S. B. 2012. Periodontal disease and the oral microbiota in new-onset rheumatoid arthritis. *Arthritis Rheum,* 64**,** 3083-94.

SCHIFFERLE, R. E., SHOSTAD, S. A., BAYERS-THERING, M. T., DYER, D. W. & NEIDERS, M. E. 1996. Effect of protoporphyrin IX limitation on porphyromonas gingivalis. *J Endod,* 22**,** 352-5.

SCHMICKLER, J., RUPPRECHT, A., PATSCHAN, S., PATSCHAN, D., MÜLLER GERHARD, A., HAAK, R., MAUSBERG RAINER, F., SCHMALZ, G., KOTTMANN, T. & ZIEBOLZ, D. 2017. Cross-Sectional Evaluation of Periodontal Status and Microbiologic and Rheumatoid Parameters in a Large Cohort of Patients With Rheumatoid Arthritis. *Journal of Periodontology,* 88**,** 368-379.

SCHWENZER, A., QUIRKE, A. M., MARZEDA, A. M., WONG, A., MONTGOMERY, A. B., SAYLES, H. R., EICK, S., GAWRON, K., CHOMYSZYN-GAJEWSKA, M., LAZARZ-BARTYZEL, K., DAVIS, S., POTEMPA, J., KESSLER, B. M., FISCHER, R., VENABLES, P. J., PAYNE, J. B., MIKULS, T. R. & MIDWOOD, K. S. 2017. Association of Distinct Fine Specificities of Anti-Citrullinated Peptide Antibodies With Elevated Immune Responses to Prevotella intermedia in a Subgroup of Patients With Rheumatoid Arthritis and Periodontitis. *Arthritis Rheumatol,* 69**,** 2303-2313.

SCOTT, C. F., WHITAKER, E. J., HAMMOND, B. F. & COLMAN, R. W. 1993. Purification and characterization of a potent 70-kDa thiol lysyl-proteinase (Lys-gingivain) from Porphyromonas gingivalis that cleaves kininogens and fibrinogen. *J Biol Chem,* 268**,** 7935-42.

SHADDOX, L. M., HUANG, H., LIN, T., HOU, W., HARRISON, P. L., AUKHIL,

I., WALKER, C. B., KLEPAC-CERAJ, V. & PASTER, B. J. 2012. Microbiological characterization in children with aggressive periodontitis. *J Dent Res,* 91**,** 927-33.

SHANNON, P., MARKIEL, A., OZIER, O., BALIGA, N. S., WANG, J. T., RAMAGE, D., AMIN, N., SCHWIKOWSKI, B. & IDEKER, T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res,* 13**,** 2498-504.

SHI, J., KNEVEL, R., SUWANNALAI, P., VAN DER LINDEN, M. P., JANSSEN, G. M., VAN VEELEN, P. A., LEVARHT, N. E., VAN DER HELM-VAN MIL, A. H., CERAMI, A., HUIZINGA, T. W., TOES, R. E. & TROUW, L. A. 2011. Autoantibodies recognizing carbamylated proteins are present in sera of patients with rheumatoid arthritis and predict joint damage. *Proc Natl Acad Sci U S A,* 108**,** 17372-7.

SHI, J., VAN DE STADT LOTTE, A., LEVARHT, E. W. N., HUIZINGA TOM, W. J., TOES RENÉ, E. M., TROUW LEENDERT, A. & VAN SCHAARDENBURG, D. 2012. Brief Report: Anti–Carbamylated Protein Antibodies Are Present in Arthralgia Patients and Predict the Development of Rheumatoid Arthritis. *Arthritis & Rheumatism,* 65**,** 911-915.

SHI, J., WILLEMZE, A., JANSSEN, G. M., VAN VEELEN, P. A., DRIJFHOUT, J. W., CERAMI, A., HUIZINGA, T. W., TROUW, L. A. & TOES, R. E. 2013. Recognition of citrullinated and carbamylated proteins by human antibodies: specificity, cross-reactivity and the 'AMC-Senshu' method. *Ann Rheum Dis,* 72**,** 148-50.

SHI, Y., RATNAYAKE, D. B., OKAMOTO, K., ABE, N., YAMAMOTO, K. & NAKAYAMA, K. 1999. Genetic Analyses of Proteolysis, Hemoglobin Binding, and Hemagglutination of Porphyromonas gingivalis : CONSTRUCTION OF MUTANTS WITH A COMBINATION OF rgpA, rgpB, kgp, AND hagA. *Journal of Biological Chemistry,* 274**,** 17955-17960.

SILMAN, A. J. & PEARSON, J. E. 2002. Epidemiology and genetics of rheumatoid arthritis. *Arthritis Research,* 4**,** S265-S272.

SINGH, A., WYANT, T., ANAYA-BERGMAN, C., ADUSE-OPOKU, J., BRUNNER, J., LAINE, M. L., CURTIS, M. A. & LEWIS, J. P. 2011. The Capsule of Porphyromonas gingivalis Leads to a Reduction in the Host Inflammatory Response, Evasion of Phagocytosis, and Increase in Virulence. *Infection and Immunity,* 79**,** 4533-4542.

SLOTS, J. 2005. Herpesviruses in periodontal diseases. *Periodontology 2000,* 38**,** 33-62.

SOCRANSKY, S. S. & HAFFAJEE, A. D. 2005. Periodontal microbial ecology. *Periodontol 2000,* 38**,** 135-87.

SOCRANSKY, S. S., HAFFAJEE, A. D., CUGINI, M. A., SMITH, C. & KENT, R. L., JR. 1998. Microbial complexes in subgingival plaque. *J Clin Periodontol,* 25**,** 134-44.

SOKKA, T., TOLOZA, S., CUTOLO, M., KAUTIAINEN, H., MAKINEN, H., GOGUS, F., SKAKIC, V., BADSHA, H., PEETS, T., BARANAUSKAITE, A., GEHER, P., UJFALUSSY, I., SKOPOULI, F. N., MAVROMMATI, M., ALTEN, R., POHL, C., SIBILIA, J., STANCATI, A., SALAFFI, F., ROMANOWSKI, W., ZAROWNY-WIERZBINSKA, D., HENROHN, D., BRESNIHAN, B., MINNOCK, P., KNUDSEN, L. S., JACOBS, J. W., CALVO-ALEN, J., LAZOVSKIS, J., PINHEIRO GDA, R., KARATEEV, D., ANDERSONE, D., REXHEPI, S., YAZICI, Y. & PINCUS, T. 2009. Women, men, and rheumatoid arthritis: analyses of disease activity, disease

characteristics, and treatments in the QUEST-RA study. *Arthritis Res Ther,* 11**,** R7.

SOLBIATI, J. & FRIAS-LOPEZ, J. 2018. Metatranscriptome of the Oral Microbiome in Health and Disease. *Journal of Dental Research,* 97**,** 492-500.

STABHOLZ, A., SOSKOLNE, W. A. & SHAPIRA, L. 2010. Genetic and environmental risk factors for chronic periodontitis and aggressive periodontitis. *Periodontol 2000,* 53**,** 138-53.

STOBERNACK, T., GLASNER, C., JUNKER, S., GABARRINI, G., DE SMIT, M., DE JONG, A., OTTO, A., BECHER, D., VAN WINKELHOFF, A. J. & VAN DIJL, J. M. 2016. Extracellular Proteome and Citrullinome of the Oral Pathogen Porphyromonas gingivalis. *Journal of Proteome Research,* 15**,** 4532-4543.

STOLT, P., BENGTSSON, C., NORDMARK, B., LINDBLAD, S., LUNDBERG, I., KLARESKOG, L. & ALFREDSSON, L. 2003. Quantification of the influence of cigarette smoking on rheumatoid arthritis: results from a population based case-control study, using incident cases. *Ann Rheum Dis,* 62**,** 835-41.

SUPEK, F., BOŠNJAK, M., ŠKUNCA, N. & ŠMUC, T. 2011. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLOS ONE,* 6**,** e21800.

SVENDSEN, A. J., JUNKER, P., HOUEN, G., KYVIK, K. O., NIELSEN, C., SKYTTHE, A. & HOLST, R. 2017. Incidence of Chronic Persistent Rheumatoid Arthritis and the Impact of Smoking: A Historical Twin Cohort Study. *Arthritis Care Res (Hoboken),* 69**,** 616-624.

SZAFRAŃSKI, S. P., DENG, Z.-L., TOMASCH, J., JAREK, M., BHUJU, S., MEISINGER, C., KÜHNISCH, J., SZTAJER, H. & WAGNER-DÖBLER, I. 2015. Functional biomarkers for chronic periodontitis and insights into the roles of Prevotella nigrescens and Fusobacterium nucleatum; a metatranscriptome analysis. *Npj Biofilms And Microbiomes,* 1**,** 15017.

TAKAHASHI, N. 2003. Acid-neutralizing activity during amino acid fermentation by Porphyromonas gingivalis, Prevotella intermedia and Fusobacterium nucleatum. *Oral Microbiol Immunol,* 18**,** 109-13.

TAKAHASHI, N. & SATO, T. 2001. Preferential Utilization of Dipeptides by Porphyromonas gingivalis. *Journal of Dental Research,* 80**,** 1425-1429.

TAKAHASHI, N., SATO, T. & YAMADA, T. 2000. Metabolic pathways for cytotoxic end product formation from glutamate- and aspartate-containing peptides by Porphyromonas gingivalis. *J Bacteriol,* 182**,** 4704-10.

TAKAHASHI, N. & SCHACHTELE, C. F. 1990. Effect of pH on the Growth and Proteolytic Activity of Porphyromonas gingivalis and Bacteroides intermedius. *Journal of Dental Research,* 69**,** 1266-1269.

TAKEUCHI, Y., UMEDA, M., SAKAMOTO, M., BENNO, Y., HUANG, Y. & ISHIKAWA, I. 2001. Treponema socranskii, Treponema denticola, and Porphyromonas gingivalis are associated with severity of periodontal tissue destruction. *J Periodontol,* 72**,** 1354-63.

TEDESCHI, S. K., BERMAS, B. & COSTENBADER, K. H. 2013. Sexual disparities in the incidence and course of SLE and RA. *Clinical Immunology,* 149**,** 211-218.

TÉMOIN, S., CHAKAKI, A., ASKARI, A., EL-HALABY, A., FITZGERALD, S., MARCUS, R. E., HAN, Y. W. & BISSADA, N. F. 2012. Identification of oral bacterial DNA in synovial fluid of arthritis patients with native and failed prosthetic joints. *Journal of clinical rheumatology : practical reports on rheumatic & musculoskeletal diseases,* 18**,** 117-121.

TRIANTAFILOU, M., GAMPER, F. G., LEPPER, P. M., MOURATIS, M. A.,

SCHUMANN, C., HAROKOPAKIS, E., SCHIFFERLE, R. E., HAJISHENGALLIS, G. & TRIANTAFILOU, K. 2007. Lipopolysaccharides from atherosclerosis-associated bacteria antagonize TLR4, induce formation of TLR2/1/CD36 complexes in lipid rafts and trigger TLR2-induced inflammatory responses in human vascular endothelial cells. *Cell Microbiol,* 9**,** 2030-9.

TROUW, L. A. & MAHLER, M. 2012. Closing the serological gap: promising novel biomarkers for the early diagnosis of rheumatoid arthritis. *Autoimmun Rev,* 12**,** 318-22.

ÜSTÜN, K., ERCIYAS, K., KISACIK, B., SEZER, U., PEHLIVAN, Y., ÖZTUZCU, S., GÜNDOĞAR, H. & ONAT, A. M. 2013. Host Modulation in Rheumatoid Arthritis Patients with TNF Blockers Significantly Decreases Biochemical Parameters in Periodontitis. *Inflammation,* 36**,** 1171-1177.

VAN DER WOUDE, D., LIE, B. A., LUNDSTROM, E., BALSA, A., FEITSMA, A. L., HOUWING-DUISTERMAAT, J. J., VERDUIJN, W., NORDANG, G. B., ALFREDSSON, L., KLARESKOG, L., PASCUAL-SALCEDO, D., GONZALEZ-GAY, M. A., LOPEZ-NEVOT, M. A., VALERO, F., ROEP, B. O., HUIZINGA, T. W., KVIEN, T. K., MARTIN, J., PADYUKOV, L., DE VRIES, R. R. & TOES, R. E. 2010. Protection against anti-citrullinated protein antibody-positive rheumatoid arthritis is predominantly associated with HLA-DRB1*1301: a meta-analysis of HLA-DRB1 associations with anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis in four European populations. *Arthritis Rheum,* 62**,** 1236-45.

VAN WINKELHOFF, A. J., LAINE, M. L., TIMMERMAN, M. F., VAN DER WEIJDEN, G. A., ABBAS, F., WINKEL, E. G., ARIEF, E. M. & VAN DER VELDEN, U. 1999. Prevalence and serotyping of Porphyromonas gingivalis in an Indonesian population. *J Clin Periodontol,* 26**,** 301-5.

VERNAL, R., DIAZ-GUERRA, E., SILVA, A., SANZ, M. & GARCIA-SANZ, J. A. 2014. Distinct human T-lymphocyte responses triggered by Porphyromonas gingivalis capsular serotypes. *J Clin Periodontol,* 41**,** 19-30.

VESSEY, M. P., VILLARD-MACKINTOSH, L. & YEATES, D. 1987. Oral contraceptives, cigarette smoking and other factors in relation to arthritis. *Contraception,* 35**,** 457-64.

VĚTROVSKÝ, T. & BALDRIAN, P. 2013. The Variability of the 16S rRNA Gene in Bacterial Genomes and Its Consequences for Bacterial Community Analyses. *PLOS ONE,* 8**,** e57923.

VOSSENAAR, E. R., ZENDMAN, A. J., VAN VENROOIJ, W. J. & PRUIJN, G. J. 2003. PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease. *Bioessays,* 25**,** 1106-18.

WADE, W. 2002. Unculturable bacteria - the uncharacterized organisms that cause oral infections. *Journal of the Royal Society of Medicine,* 95**,** 81-83.

WANG, Z., NICHOLLS, S. J., RODRIGUEZ, E. R., KUMMU, O., HORKKO, S., BARNARD, J., REYNOLDS, W. F., TOPOL, E. J., DIDONATO, J. A. & HAZEN, S. L. 2007. Protein carbamylation links inflammation, smoking, uremia and atherogenesis. *Nat Med,* 13**,** 1176-84.

WEGNER, N., WAIT, R., SROKA, A., EICK, S., NGUYEN, K.-A., LUNDBERG, K., KINLOCH, A., CULSHAW, S., POTEMPA, J. & VENABLES, P. J. 2010. Peptidylarginine deiminase from Porphyromonas gingivalis citrullinates human fibrinogen and α-enolase: Implications for autoimmunity in rheumatoid arthritis. *Arthritis and rheumatism,* 62**,** 2662-2672.

WEISS, S., XU, Z. Z., PEDDADA, S., AMIR, A., BITTINGER, K., GONZALEZ, A., LOZUPONE, C., ZANEVELD, J. R., VÁZQUEZ-BAEZA, Y., BIRMINGHAM, A., HYDE, E. R. & KNIGHT, R. 2017. Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome,* 5**,** 27.

WHITE, P. C., CHICCA, I. J., COOPER, P. R., MILWARD, M. R. & CHAPPLE, I. L. 2016. Neutrophil Extracellular Traps in Periodontitis: A Web of Intrigue. *J Dent Res,* 95**,** 26-34.

WHITE, P. C., CHICCA, I. J., COOPER, P. R., MILWARD, M. R. & CHAPPLE, I. L. C. 2015. Neutrophil Extracellular Traps in Periodontitis: A Web of Intrigue. *Journal of Dental Research,* 95**,** 26-34.

WOJTOWICZ, H., GUEVARA, T., TALLANT, C., OLCZAK, M., SROKA, A., POTEMPA, J., SOLA, M., OLCZAK, T. & GOMIS-RUTH, F. X. 2009. Unique structure and stability of HmuY, a novel heme-binding protein of Porphyromonas gingivalis. *PLoS Pathog,* 5**,** e1000419.

WOOD, S. R., KIRKHAM, J., MARSH, P. D., SHORE, R. C., NATTRESS, B. & ROBINSON, C. 2000. Architecture of Intact Natural Human Plaque Biofilms Studied by Confocal Laser Scanning Microscopy. *Journal of Dental Research,* 79**,** 21-27.

YOSHIMURA, A., KANEKO, T., KATO, Y., GOLENBOCK, D. T. & HARA, Y. 2002. Lipopolysaccharides from Periodontopathic Bacteria Porphyromonas gingivalis and Capnocytophaga ochracea Are Antagonists for Human Toll-Like Receptor 4. *Infection and Immunity,* 70**,** 218-225.

YOST, S., DURAN-PINEDO, A. E., TELES, R., KRISHNAN, K. & FRIAS-LOPEZ, J. 2015. Functional signatures of oral dysbiosis during periodontitis progression revealed by microbial metatranscriptome analysis. *Genome Med,* 7**,** 27.

ZENOBIA, C. & HAJISHENGALLIS, G. 2015. Porphyromonas gingivalis virulence factors involved in subversion of leukocytes and microbial dysbiosis. *Virulence,* 6**,** 236-243.

ZHANG, X., ZHANG, D., JIA, H., FENG, Q., WANG, D., LIANG, D., WU, X., LI, J., TANG, L., LI, Y., LAN, Z., CHEN, B., LI, Y., ZHONG, H., XIE, H., JIE, Z., CHEN, W., TANG, S., XU, X., WANG, X., CAI, X., LIU, S., XIA, Y., LI, J., QIAO, X., AL-AAMA, J. Y., CHEN, H., WANG, L., WU, Q.-J., ZHANG, F., ZHENG, W., LI, Y., ZHANG, M., LUO, G., XUE, W., XIAO, L., LI, J., CHEN, W., XU, X., YIN, Y., YANG, H., WANG, J., KRISTIANSEN, K., LIU, L., LI, T., HUANG, Q., LI, Y. & WANG, J. 2015. The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. *Nature Medicine,* 21**,** 895.

ZHANG, Y., ZHEN, M., ZHAN, Y., SONG, Y., ZHANG, Q. & WANG, J. 2017. Population-Genomic Insights into Variation in Prevotella intermedia and Prevotella nigrescens Isolates and Its Association with Periodontal Disease. *Frontiers in Cellular and Infection Microbiology,* 7**,** 409.

List of Abbreviations

ATCC	American type culture collection
ACPA	Anti-citrullinated protein antibodies
CAL	Clinical attachment loss
CCP	Cyclic citrullinated peptide
CFU	Colony forming units
DMARDs	Disease-modifying anti-rheumatic drugs
DPP	Dipeptidyl peptidase
EULAR	European League Against Rheumatism
GCF	Gingival crevicular fluid
HC	Healthy control
Kgp	lysine gingipain
LPS	Lipopolysaccharide
NGS	Next generation sequencing
NORA	New-onset rheumatoid arthritis
OUT	Operational taxonomic units
PBS	Phosphate buffered saline
PCoA	Principal coordinates analysis
qRT-PCR	Quantitative real-time polymerase chain reaction
PPAD	*P. gingivalis* peptidylarginine deiminase
RA	Rheumatoid arthritis
RF	Rheumatoid factor
Rgp	Arginine gingipain

Appendix A Recipes for Buffers Used

Sodium boric acid (SB) buffer

50x Stock:

20 g NaOH

pH 8.0 with 120 g H_3BO_3 (boric acid powder)

bring to 1 L with deionized H_2O

1x Working Buffer:

20 ml of 50x stock

980 ml dH2O

Appendix B Scripts Used for Data Analysis

B 1 Cutadapt

```
cutadapt -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC -A
AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCATT -
o $1 -p $2 $3 $4
#note: $1 is the output:  *_R1_trimmed.fastq.gz, $2 is *_R2_trimmed.fastq.gz,$3 is the
input,*R1.fastq.gz, $4 is *R2.fastq.gz
```

B 2 Sickle

```
#note: $1 is the input, *_R1_trimmed.fastq.gz,$2 is *_R2_trimmed.fastq.gz,$3 is the output,
*_R1_sickle.fastq, $4 is *_R2_sickle.fastq,$5 is *_single_sickel.fastq
```

B 3 Burrow-Wheeler Aligner

```
bwa index $1
```

B 4 Samtools import

```
samtools import pginw83.fna.fai $1 $2
#note: $1 is the .sam file and $2 the .bam file
```

B 5 Samtools sort

```
samtools sort $1>$2
#note: $1 should be the bam file to sort, $2 is the sorted bam file
```

B 6 Htseq-count

```
htseq-count -s no -f bam -t CDS --additional-attr=gene_name $1 $2 > $3
#note: $1 is the sorted bam file, $2 is the gtf file (in this case pginw83.gtf) and $3 is the output as
*_readcount.txt
```

B 7 Megahit

```
megahit -1 $1 -2 $2 -o $3
#note: $1 is read1, $2 is read2 and $3 is the path to the directory where the results will be saved to.
```

B 8 Diamond

```
diamond blastx -p 47 -d /nobackup/dennttd/diamond/nr_prot_fasta/nr.dmnd  -q $1 -k 1 --id 95 --
more-sensitive --daa $2
#note: $1 is the fastq or fasta file (either trimmed reads or the generated contigs from megahit), $2 is
the *.daa ouput file
```

B 9 TopGO (R script)

```
#load gene to GO annotation txt file, (geneID GO1,GO2)
geneID2GO <- readMappings(file = "~/Desktop/Readcount of
Wildtype/results(4conditions/topGO/go2gene.map.txt")
# build genelist including background gene list and significant genes, background genes are the name
of the allgenes
geneList <- factor(as.integer(allgenes%in%significantgenes))
names(geneList) <- go2gene$Gene
#create Godata object
GOdata <- new("topGOdata", ontology = "BP", allGenes = geneList,annot = annFUN.gene2GO,
gene2GO = geneID_Go)
GOdata
#run the erichment test by classic fisher test
resultFisher <- runTest(GOdata, algorithm = "classic", statistic = "fisher")
resultFisher
#generate results table
allRes <- GenTable(GOdata, classicFisher = resultFisher, ranksOf = "classicFisher",topNodes =
30,numChar=99)
```

```
allRes
#Visualising the GO structure
par(cex = 0.1)+showSigOfNodes(GOdata, score(resultFis), firstSigNodes = 5, useInfo = 'all')
```

B 10 Network analysis (R script)

```
#load phyloseq object of species raw count  and transfor to relative abundance
library(phloseq)
> load("/Users/Zijian/Desktop/Hsite_species.Rdata")
rltdt_Hsite<-transform_sample_counts(Hsite_species,function(x) x/sum(x) )
#using microbiome package to filter species (relative abundance>0,prevalence>20%)
library(microbiome)
>rltdt_Hsite_filtered<-core(rltdt_Hsite, detection = 0,prevalence = 20/100)
# generate the table as the input for network analysis (rowsnames are species, columns names are samples)
write.csv(as.data.frame(otu_table(rltdt_Hsite_filtered)),file = "~/Desktop/input.csv")
# network analysis using the 1.Pairwise_correlations.R and 2.Network_analysis.R (https://github.com/RichieJu520/Co-occurrence_Network_Analysis)
# gml file generated from the script will be used for graph ploting in cytospace
# abundance table and taxanomic table of species will also be generated and incorporated into cytospace for ploting (mapping with color and size of node).
```

Appendix C Supplementary Results

C 1 Differentially expressed genes in *P. gingivalis* W83 Δ*ppad* mutant compared with the wild-type when grown under identical conditions in a chemostat (one-fold or more, adjusted $P < 0.01$, DESeq2).

Culture stages	Total	Genes
1st & 2nd steady-states, early transition stage, late transition stage	25	PG_0646
		PG_0495
		PG_1179
		PG_0275
		PG_1178
		PG_1868
		PG_0174
		htrA
		PG_1124
		PG_0645
		PG_0707
		PG_0173
		PG_1180
		folD
		PG_1858
		PG_1543
		PG_0555
		PG_0195
		hmuR
		PG_0432
		PG_1424
		hmuY
		PG_0686
		PG_0421

		PG_1553
1^{st} & 2^{nd} steady-states, early transition stage	6	PG_1181
		PG_1641
		PG_1556
		PG_1421
		PG_1019
		PG_1554
2^{nd} steady-state, early transition stage, late transition stage	10	PG_0725
		PG_0602
		pgaA
		PG_1491
		PG_0047
		PG_0906
		PG_0739
		porT
		PG_0735
		PG_0759
1^{st} steady-state, early transition stage	15	PG_1555
		PG_1715
		PG_1729
		PG_1837
		PG_1426
		sodB
		rplS
		PG_1374
		PG_1547
		PG_2101
		PG_1236
		PG_2102
		trx

		rpmE
		ftn
1st & 2nd steady-states	11	queA
		PG_0283
		PG_1869
		PG_1514
		PG_1982
		tpr
		PG_1983
		PG_0554
		PG_1485
		PG_0433
		folK
early transition stage, late transition stage	8	nhaA
		PG_0745
		rpsR
		PG_0598
		PG_0009
		topA
		PG_2168
		ribE
2nd steady-state, early transition stage	4	PG_0737
		PG_0708
		PG_0624
		PG_1020
2nd steady-state, late transition stage	37	cutC
		efp-1
		ftsA
		lpxK
		PG_0161

PG_0242
PG_0291
PG_0308
PG_0430
PG_0434
PG_0491
PG_0565
PG_0585
PG_0612
PG_0613
PG_0616
PG_0702
PG_0717
PG_0720
PG_0721
PG_0724
PG_0740
PG_0750
PG_0757
PG_0928
PG_0929
PG_1591
PG_1795
PG_1823
PG_1974
PG_1988
PG_2008
PG_2106
ppnK
prtC

		pruA
		rplT
1st steady-state	62	gmk
		cas1
		PG_2095
		PG_1893
		PG_1984
		PG_0508
		PG_0513
		topB-2
		PG_0451
		cas4
		PG_0510
		PG_2100
		PG_2131
		PG_1663
		PG_2134
		PG_1679
		PG_0060
		PG_0411
		PG_0117
		PG_0871
		PG_1015
		PG_0509
		PG_0458
		PG_0218
		PG_0682
		PG_0216
		PG_1738
		PG_1478

	yngK-1
	PG_1986
	PG_1359
	PG_1486
	PG_1490
	cas2-1
	PG_0285
	PG_1480
	PG_0874
	PG_1479
	panE
	PG_1987
	PG_2103
	PG_1594
	PG_0215
	PG_1746
	PG_0018
	PG_0546
	cas2-2
	PG_0511
	PG_0293
	PG_1817
	PG_1857
	PG_1662
	PG_0972
	PG_0917
	PG_1874
	PG_1481
	PG_1969
	PG_0274

		PG_1745
		PG_0214
		PG_0680
		PG_1326
early transition stage	35	PG_1021
		PG_0710
		PG_1492
		PG_0865
		PG_1482
		PG_0749
		PG_0839
		PG_1008
		PG_0617
		PG_0569
		PG_0786
		rpsF
		PG_1841
		PG_1010
		PG_0371
		manC
		PG_0444
		PG_0799
		PG_1975
		PG_0752
		PG_2225
		rpmG
		PG_0323
		PG_2058
		PG_0326
		rplM

		PG_0685
		PG_2120
		PG_0619
		PG_1842
		PG_1196
		typA
		fur
		PG_0039
		PG_1667
late transition stage	21	etfA-2
		PG_1380
		PG_1270
		yaaA
		PG_0179
		PG_0633
		PG_0718
		PG_0756
		mutA
		PG_1164
		feoB-2
		PG_2064
		PG_0734
		PG_0020
		PG_1069
		rpoD
		PG_0148
		PG_0744
		PG_1507
		frdB
		gcvH

2nd steady-state	62	PG_0068
		PG_1709
		PG_1273
		dinF
		PG_0562
		PG_2017
		PG_1684
		cmk
		rnhB
		PG_1067
		PG_0738
		ftsz
		PG_0382
		PG_1666
		PG_0424
		PG_0987
		PG_1085
		map
		murG
		groEL
		PG_0627
		PG_0055
		thiC
		PG_1989
		PG_0614
		PG_0492
		PG_2139
		PG_0726
		PG_1022
		PG_0438

PG_0611
sppA
PG_1285
PG_0747
PG_1271
PG_0712
fkpA
PG_0723
PG_0985
prtQ
PG_0575
ispH
thiL
PG_1108
cas3
PG_0864
rpmH
PG_0558
PG_0258
PG_0986
recF
PG_0447
clpB
PG_0579
PG_0182
PG_0610
rpoA
PG_2209
PG_0490
PG_0704

PG_0647
PG_0932

C 2 Differentially expressed genes of *P. gingivalis* W83 wild-type after removing pH control (one-fold or more, adjusted $P < 0.01$, DESeq2).

Culture stage (compared with the 1st steady-state)	Total	Genes
early transition stage, late transition stage, 2nd steady-state	39	PG_0283
		PG_1176
		PG_1489
		PG_1109
		PG_1475
		PG_1020
		PG_1482
		PG_1494
		PG_1486
		topB-2
		lacZ-1
		PG_1571
		PG_1490
		PG_1769
		PG_1485
		PG_1476
		PG_1554
		PG_0285
		PG_0930
		PG_2131
		PG_1480
		PG_2063
		PG_0493
		PG_1479
		PG_1975
		PG_2134
		PG_1679

		PG_1481
		PG_1594
		PG_0337
		PG_0281
		PG_0717
		PG_0718
		PG_1022
		PG_2133
		PG_0282
		PG_2114
		PG_1478
		PG_1178
early transition stage, late transition stage	6	PG_1180
		hslR
		PG_1181
		PG_0799
		PG_1113
		PG_0680
early transition stage, 2nd steady-state	3	PG_1619
		ribD
		PG_1977
late transition stage, 2nd steady-state	15	PG_1534
		PG_0064
		PG_0049
		PG_0749
		PG_1398
		PG_0847
		PG_2064
		PG_0065
		PG_0419

		PG_1296
		PG_0871
		PG_0090
		PG_0215
		PG_0214
		PG_0216
early transition stage	4	PG_0495
		PG_2226
		PG_1555
		PG_0158
late transition stage	8	pth
		PG_1474
		PG_0928
		PG_1462
		PG_1995
		PG_0470
		PG_0404
		PG_1787
2^{nd} steady-state	25	PG_1572
		PG_0180
		PG_1270
		recX
		PG_0333
		PG_1837
		PG_0006
		PG_0879
		PG_1271
		PG_0732
		hmuY
		PG_1175

	PG_1102
	PG_1098
	PG_0720
	PG_0874
	megL
	PG_0985
	PG_1666
	nrd
	PG_0612
	PG_0063
	PG_0429
	PG_2116
	PG_2040

C 3 Differentially expressed genes of *P. gingivalis* W83 Δ*ppad* mutant after removing pH control (one-fold or more, adjusted $P < 0.01$, DESeq2).

Culture stage (compared with the 1st steady-state)	Total	Genes
early transition stage, late transition stage, 2nd steady-state	4	PG_2017
		fabG
		PG_0090
		PG_0612
early transition stage, 2nd steady-state	1	PG_0865
late transition stage, 2nd steady-state	51	PG_0611
		PG_1819
		PG_0646
		PG_1374
		PG_0410
		PG_1497
		PG_0585
		epsD
		PG_1988
		PG_1858
		PG_0565
		PG_2018
		efp-1
		PG_1987
		PG_0613
		PG_0757
		rnhA
		PG_2071
		PG_1908
		rplY
		PG_1664
		PG_0232

		PG_1982
		cas3
		PG_1985
		PG_1493
		PG_0411
		PG_1729
		PG_0174
		groEL
		PG_1837
		PG_0293
		ribF
		fabF
		PG_1212
		PG_0645
		PG_0558
		rplT
		hmuY
		PG_1999
		PG_2048
		ftsA
		murC
		PG_0614
		htpG
		PG_1211
		PG_2019
		PG_0173
		hagE
		gcvH
		acpP
early transition stage	6	PG_0009

		PG_1008
		PG_0698
		PG_2120
		PG_1527
		PG_0371
late transition stage	9	PG_1164
		PG_0133
		PG_1380
		mraY
		tagD
		PG_1507
		mrr
		dnaJ
		PG_1742
2^{nd} steady-stage	95	yngK-1
		PG_0461
		PG_1492
		sppA
		PG_1986
		PG_0068
		PG_2127
		PG_1893
		PG_1984
		nrfA
		rplS
		PG_0725
		PG_0495
		cas4
		PG_0562
		PG_1514

	PG_0491
	PG_0906
	cas2-1
	PG_2100
	cydA
	PG_0345
	PG_1684
	PG_0737
	oxyR
	PG_1543
	PG_1127
	rnhB
	PG_2161
	PG_0971
	PG_1067
	PG_1663
	ftsz
	PG_0344
	PG_1235
	rpsO
	PG_0575
	PG_1014
	nrd
	PG_2020
	lpxK
	PG_0582
	PG_1221
	PG_0987
	porT
	(PG_0751)

	PG_0226
	PG_0450
	murG
	rpmH
	cas2-2
	PG_0183
	PG_1983
	PG_1891
	PG_0291
	PG_1817
	thiC
	PG_2149
	PG_0109
	PG_1989
	PG_0986
	PG_0441
	PG_0447
	clpB
	PG_1659
	PG_0579
	PG_0492
	PG_1660
	PG_2139
	pckA
	dut
	PG_0027
	PG_1001
	PG_1767
	PG_0678
	PG_0610

	PG_0724
	purN
	PG_0832
	PG_0329
	PG_0110
	PG_1665
	serS
	PG_0438
	rpmI
	rpmB
	PG_1358
	PG_0704
	infC
	PG_1747
	PG_2216
	PG_1226
	PG_0647
	uvrC
	PG_1892
	PG_1326

C 4 Bacterial genera associated with different groups

Group	* Genera	# Corrected *P*
Healthy site		
HC	*Peptostreptococcus*	0.0277
CCP	*Capnocytophaga*	0.0199
	Treponema	0.0238
NORA	*Methylobacterium*	0.0159
	Pseudopropionibacterium	0.0316
RA	*Bacillus*	0.0159
	Vibrio	0.0394
	Bradyrhizobium	0.0471
Diseased site		
CCP	*Porphyromonas*	0.0355
	Treponema	0.0471
NORA	*Neisseria*	0.0040
	Cardiobacterium	0.0040
	Haemophilus	0.0119
	Corynebacterium	0.0119
	Klebsiella	0.0471
RA	*Stomatobaculum*	0.0040

*: Genera with significantly higher relative abundance between groups were investigated by using the permutation test (one-sided *signassoc* function, R).
#: Sidak's correction was applied for multiple testing

C 5 Uniquely detected species in HC, CCP, NORA and RA groups from periodontally healthy sites.

Number of unique species	Species	Number of infected samples	%Prevalence	Total counts
H_HC (32 samples)				
22	*[Eubacterium] sulci*	1	3.13%	26
	Acidovorax delafieldii	1	3.13%	827
	Capnocytophaga haemolytica	1	3.13%	1609
	Granulicatella elegans	1	3.13%	106
	Haemophilus sp. HMSC068C11	1	3.13%	72
	Haemophilus sp. HMSC61B11	1	3.13%	42
	Lactobacillus kimchicus	1	3.13%	1428
	Mitsuokella sp. oral taxon 131	1	3.13%	1852
	Mogibacterium pumilum	3	9.38%	799
	Neisseria cinerea	2	6.25%	387
	Neisseria polysaccharea	1	3.13%	39
	Paracoccus sp. 228	1	3.13%	198
	Paracoccus sp. PAMC 22219	1	3.13%	67
	Paracoccus sp. S4493	1	3.13%	38
	Piscicoccus intestinalis	1	3.13%	1219
	Propionibacterium sp. 409-HC1	1	3.13%	15
	Propionibacterium sp. 434-HC2	1	3.13%	20
	Rhodopseudomonas palustris	2	6.25%	50
	Selenomonas ruminantium	1	3.13%	18
	Streptococcus sp. CCH8-G7	1	3.13%	123
	Streptococcus sp. HMSC072D05	1	3.13%	18
	Streptococcus sp. NLAE-zl-C503	1	3.13%	58
H_CCP (48samples)				

91	*Achromobacter sp. 2789STDY5608615*	1	2.08%	21
	Actinomyces graevenitzii	2	4.17%	441
	Actinomyces sp. HMSC035G02	3	6.25%	229
	Actinomyces sp. HPA0247	2	4.17%	152
	Actinomyces sp. ICM54	4	8.33%	396
	Afipia broomeae	1	2.08%	6
	Alishewanella jeotgali	1	2.08%	23
	Bifidobacterium adolescentis	1	2.08%	7
	Bifidobacterium moukalabense	3	6.25%	111
	Bilophila wadsworthia	1	2.08%	176
	Bradyrhizobium liaoningense	1	2.08%	9
	Bradyrhizobium sp.	1	2.08%	8
	Campylobacter curvus	2	4.17%	158
	Campylobacter sp. 10_1_50	1	2.08%	128
	Corynebacterium aurimucosum	1	2.08%	197
	Corynebacterium pseudogenitalium	1	2.08%	248
	Corynebacterium sp. OG2	1	2.08%	132
	Corynebacterium tuberculostearicum	1	2.08%	108
	Eikenella sp. NML99-0057	1	2.08%	81
	Enhydrobacter aerosaccus	1	2.08%	118
	Enterococcus saccharolyticus	1	2.08%	18
	Gardnerella vaginalis	1	2.08%	14
	Kocuria sp. UCD-OTCP	1	2.08%	73
	Lactobacillus acidophilus	1	2.08%	22
	Lactobacillus amylovorus	1	2.08%	45
	Lactobacillus casei	1	2.08%	48
	Lactobacillus crispatus	1	2.08%	80

Lactobacillus fermentum	2	4.17%	287
Lactobacillus gasseri	1	2.08%	1760
Lactobacillus kitasatonis	1	2.08%	35
Lactobacillus paracasei	1	2.08%	68
Lactobacillus rhamnosus	1	2.08%	897
Lactobacillus salivarius	2	4.17%	1946
Lactobacillus ultunensis	1	2.08%	27
Methylobacterium sp. MB200	1	2.08%	10
Microcella alkaliphila	1	2.08%	32
Micrococcales bacterium 73-15	1	2.08%	67
Micrococcus sp. HMSC31B01	1	2.08%	12
Moraxella osloensis	1	2.08%	93
Mycobacterium llatzerense	1	2.08%	28
Neisseria sp. HMSC03D10	2	4.17%	20
Neisseria sp. HMSC056A03	4	8.33%	181
Neisseria sp. HMSC065D04	3	6.25%	216
Neisseria sp. HMSC068C04	1	2.08%	33
Neisseria sp. HMSC069H12	3	6.25%	200
Neisseria sp. HMSC06F02	2	4.17%	48
Neisseria sp. HMSC072F04	1	2.08%	17
Neisseria sp. HMSC074B07	3	6.25%	93
Neisseria sp. HMSC15C08	2	4.17%	21
Neisseria sp. HMSC70E02	2	4.17%	103
Neisseria subflava	3	6.25%	115
Paracoccus tibetensis	1	2.08%	22
Paracoccus yeei	1	2.08%	127
Paraprevotella clara CAG:116	1	2.08%	8
Pectobacterium carotovorum	1	2.08%	5
Peptostreptococcaceae bacterium oral taxon 113	2	4.17%	44

Phyllobacterium sp. OV277	1	2.08%	20
Prevotella amnii	2	4.17%	29
Prevotella bryantii	1	2.08%	8
Prevotella jejuni	2	4.17%	136
Pseudonocardia sp. P2	1	2.08%	53
Pyramidobacter piscolens	1	2.08%	17
Ralstonia pickettii	1	2.08%	159
Rhodobacter sphaeroides	2	4.17%	283
Roseomonas aerilata	1	2.08%	115
Sanguibacter keddieii	3	6.25%	239
Staphylococcus capitis	1	2.08%	656
Staphylococcus sp. E463	1	2.08%	21
Streptococcus peroris	1	2.08%	439
Streptococcus sp. BS35b	1	2.08%	38
Streptococcus sp. CCH8-C6	2	4.17%	41
Streptococcus sp. CM6	1	2.08%	58
Streptococcus sp. F0441	1	2.08%	29
Streptococcus sp. GMD4S	1	2.08%	6
Streptococcus sp. HMSC062B01	1	2.08%	81
Streptococcus sp. HMSC066F01	1	2.08%	61
Streptococcus sp. HMSC073F11	1	2.08%	69
Streptococcus sp. HMSC076C09	1	2.08%	23
Streptococcus sp. HMSC077F03	1	2.08%	65
Streptococcus sp. HMSC34B10	1	2.08%	424
Streptococcus sp. HPH0090	1	2.08%	31
Streptococcus sp. SR1	1	2.08%	50
Streptococcus sp. VT 162	3	6.25%	65
Tepidiphilus thermophilus	1	2.08%	127
Thioflexothrix psekupsii	1	2.08%	6
uncultured bacterium	1	2.08%	22

	Veillonella sp. ACP1	1	2.08%	78
	Veillonella sp. HPA0037	1	2.08%	48
	Veillonella sp. ICM51a	1	2.08%	59
	Xanthomonas campestris	1	2.08%	509
	Xylanimonas cellulosilytica	1	2.08%	59
H_NORA (26 samples)				
74	*[Eubacterium] minutum*	1	3.85%	78
	Acinetobacter guillouiae	1	3.85%	47
	Acinetobacter pittii	1	3.85%	18
	Acinetobacter sp. 479375	1	3.85%	33
	Acinetobacter sp. Root1280	1	3.85%	21
	Acinetobacter ursingii	1	3.85%	1525
	Actinobacillus ureae	1	3.85%	2
	Agrobacterium tumefaciens	1	3.85%	80
	alpha proteobacterium L41A	1	3.85%	82
	Bacillus subtilis	1	3.85%	3
	Bacteroides stercoris	1	3.85%	2
	Bradyrhizobium lupini	1	3.85%	101
	Brevundimonas diminuta	1	3.85%	37
	Brevundimonas nasdae	1	3.85%	46
	Brevundimonas sp. BAL3	1	3.85%	21
	Brevundimonas sp. DS20	1	3.85%	52
	Brevundimonas sp. KM4	1	3.85%	26
	Brevundimonas sp. Leaf280	1	3.85%	17
	Brevundimonas sp. SH203	2	7.69%	527
	Candidatus Saccharibacteria bacterium 32-49-12	1	3.85%	2
	Caulobacterales bacterium RIFOXYB1_FULL_67_16	1	3.85%	18
	Chryseobacterium balustinum	1	3.85%	76

Chryseobacterium indoltheticum	1	3.85%	37
Chryseobacterium scophthalmum	1	3.85%	32
Delftia lacustris	1	3.85%	17
Enterobacter cloacae	1	3.85%	24
Enterococcus gallinarum	1	3.85%	152
Flavobacteria bacterium BBFL7	1	3.85%	2
Halothiobacillus sp. 28-55-5	1	3.85%	6
Helicobacter pylori	1	3.85%	4
Klebsiella michiganensis	1	3.85%	448
Klebsiella oxytoca	1	3.85%	149
Klebsiella sp. AS10	1	3.85%	28
Klebsiella sp. OBRC7	1	3.85%	41
Kocuria rhizophila	1	3.85%	38
Lachnospiraceae bacterium A4	1	3.85%	98
Lactobacillus murinus	1	3.85%	910
Lactobacillus reuteri	1	3.85%	99
Lactobacillus taiwanensis	1	3.85%	311
Methylobacterium salsuginis	1	3.85%	5
Methylobacterium sp. AMS5	1	3.85%	6
Methylobacterium sp. C1	1	3.85%	7
Methylobacterium sp. Leaf361	1	3.85%	5
Microbacterium barkeri	1	3.85%	761
Microbacterium maritypicum	1	3.85%	21
Microbacterium oxydans	1	3.85%	54
Microbacterium paraoxydans	1	3.85%	17
Microbacterium sp. Ag1	1	3.85%	20
Microbacterium sp. UCD-TDU	1	3.85%	36
Microbacterium sp. URHA0036	1	3.85%	18
Moorella glycerini	1	3.85%	7
Neisseria sp. HMSC15G01	1	3.85%	46

	Paenibacillus sp. St-s	1	3.85%	9
	Paraprevotella clara	1	3.85%	2
	Porphyromonadaceae bacterium COT-184 OH4590	1	3.85%	4
	Prevotella multisaccharivorax	1	3.85%	21
	Prevotella sp. HMSC077E09	1	3.85%	12
	Prevotella sp. KHD1	1	3.85%	3
	Prevotella timonensis	1	3.85%	9
	Raoultella ornithinolytica	1	3.85%	17
	Rhizobium sp. IRBG74	1	3.85%	60
	Rhizobium sp. UR51a	1	3.85%	20
	Rothia sp. HMSC068F09	1	3.85%	34
	Rothia sp. HMSC072B03	1	3.85%	50
	Shuttleworthia satelles	1	3.85%	9
	Sphingobacterium mizutaii	1	3.85%	24
	Sphingobacterium sp. IITKGP-BTPF85	1	3.85%	651
	Sphingobacterium sp. ML3W	1	3.85%	20
	Sphingobacterium sp. PM2-P1-29	1	3.85%	46
	Sphingobacterium spiritivorum	1	3.85%	2651
	Sphingomonas sp. 67-36	1	3.85%	6
	Streptococcus sp. oral taxon 056	1	3.85%	244
	Williamsia muralis	1	3.85%	162
	Williamsia sp. D3	1	3.85%	151
H_RA (10 samples)				
27	*Aerococcus sp. HMSC23C02*	1	10.00%	140
	Aeromicrobium sp. Leaf245	1	10.00%	203
	Aeromicrobium sp. Leaf272	1	10.00%	50
	Atopobium sp. ICM42b	1	10.00%	45
	Bacillus azotoformans	1	10.00%	1464

Beggiatoa sp. PS	1	10.00%	139
Candidatus Saccharibacteria bacterium RAAC3_TM7_1	1	10.00%	2
Cloacibacterium normanense	1	10.00%	120
Janibacter indicus	1	10.00%	175
Klebsiella aerogenes	1	10.00%	174
Lactobacillus brevis	1	10.00%	2
Leuconostoc mesenteroides	1	10.00%	130
Pedobacter panaciterrae	1	10.00%	2
Porphyromonas sp. CAG:1061	1	10.00%	1
Prevotella sp. AGR2160	1	10.00%	1
Prevotella sp. CAG:1058	1	10.00%	1
Prevotella sp. oral taxon 299	1	10.00%	100
Prevotellaceae bacterium HUN156	1	10.00%	1
Proteiniphilum acetatigenes	1	10.00%	121
Pseudomonas mendocina	1	10.00%	103
Pseudomonas syringae	1	10.00%	118
Solibacillus silvestris	1	10.00%	185
Streptococcus dysgalactiae	1	10.00%	170
Streptococcus sp. GMD6S	1	10.00%	118
Veillonella rodentium	1	10.00%	29
Veillonella sp. AS16	1	10.00%	117
Vibrio cholerae	1	10.00%	1

C 6 Uniquely detected species in HC, CCP, NORA and RA groups from periodontally diseased sites.

Number of unique species	Species	Number of infected samples	%Prevalence	Total counts
D_HC (18 samples in total)				
14	*Actinomyces polynesiensis*	1	5.56%	1491
	Actinomyces provencensis	1	5.56%	217
	Actinomyces sp. ICM47	1	5.56%	56
	Actinomyces sp. ICM54	1	5.56%	79
	Actinomyces sp. ICM58	1	5.56%	54
	Anaerobacillus alkalilacustris	1	5.56%	564
	Bifidobacterium moukalabense	1	5.56%	26
	Capnocytophaga haemolytica	1	5.56%	90
	Neisseria sp. HMSC059F02	1	5.56%	77
	Neisseria sp. HMSC070E12	1	5.56%	95
	Peptostreptococcaceae bacterium AS15	1	5.56%	126
	Prevotella bergensis	1	5.56%	22
	Prevotella multisaccharivorax	1	5.56%	27
	Propionibacterium sp. oral taxon192	1	5.56%	130
D_CCP (39 samples in total)				
166	*Acetobacter orientalis*	1	2.56%	20
	Achromobacter xylosoxidans	1	2.56%	155
	Acidipropionibacterium jensenii	1	2.56%	1
	Acidovorax delafieldii	1	2.56%	3259
	Acidovorax sp. CCH12-A4	2	5.13%	1282
	Acidovorax sp. KKS102	1	2.56%	1816

Acidovorax sp. NO-1	1	2.56%	959
Acidovorax temperans	1	2.56%	1381
Acinetobacter johnsonii	1	2.56%	478
Afipia broomeae	1	2.56%	1863
Anaerococcus prevotii	1	2.56%	1
Aquabacterium sp. NJ1	1	2.56%	2
Atopobium sp. BS2	1	2.56%	65
Bacillus azotoformans	1	2.56%	605
Bacillus cereus	1	2.56%	1858
Bacteroides pyogenes	1	2.56%	4
Bacteroides stercoris	1	2.56%	15
Bacteroides uniformis	1	2.56%	1
Blastococcus sp. DSM 44268	1	2.56%	1
Brachybacterium paraconglomeratum	1	2.56%	2
Brachybacterium sp. HMSC06H03	1	2.56%	2
Brachybacterium sp. sponge	1	2.56%	2
Brachybacterium sp. SW0106-09	1	2.56%	8
Bradyrhizobiaceae bacterium SG-6C	1	2.56%	545
Bradyrhizobium liaoningense	1	2.56%	2
Bradyrhizobium manausense	1	2.56%	1
Bradyrhizobium sp.	1	2.56%	1
Bradyrhizobium sp. CCBAU 43298	1	2.56%	1
Bradyrhizobium sp. DFCI-1	1	2.56%	2
Bradyrhizobium sp. WSM3983	1	2.56%	1
Bradyrhizobium valentinum	1	2.56%	1
Brevibacterium casei	4	10.26%	3585

Brevibacterium sp. VCM10	1	2.56%	1
Brevundimonas sp. DS20	1	2.56%	1
Brevundimonas sp. KM4	1	2.56%	2
Brevundimonas sp. Leaf280	1	2.56%	1
Brevundimonas sp. SH203	1	2.56%	4
Brevundimonas vesicularis	1	2.56%	1
Campylobacter curvus	1	2.56%	755
Campylobacter sp. FOBRC14	1	2.56%	281
candidate division TM7 genomosp. GTL1	1	2.56%	1
Candidatus Kryptobacter tengchongensis	1	2.56%	106
Candidatus Saccharibacteria bacterium 32-49-10	1	2.56%	526
Candidatus Saccharibacteria bacterium 32-49-12	2	5.13%	469
Candidatus Saccharibacteria bacterium 32-50-10	2	5.13%	1904
Candidatus Saccharibacteria bacterium 32-50-13	1	2.56%	1
Candidatus Saccharibacteria bacterium GW2011_GWC2_48_9	2	5.13%	762
Candidatus Saccharibacteria bacterium RAAC3_TM7_1	2	5.13%	1286
Caulobacter sp. CCH5-E12	1	2.56%	796
Caulobacter vibrioides	1	2.56%	931
Cloacibacterium normanense	1	2.56%	1
Clostridiales bacterium 1_7_47FAA	1	2.56%	1
Clostridium sp. FS41	1	2.56%	492

Clostridium sulfidigenes	1	2.56%	1
Corynebacterium ammoniagenes	1	2.56%	1
Corynebacterium falsenii	1	2.56%	2
Corynebacterium pseudogenitalium	1	2.56%	685
Corynebacterium sp. HMSC08D02	1	2.56%	5
Corynebacterium sp. HMSC28B08	2	5.13%	6
Corynebacterium urealyticum	1	2.56%	1
Corynebacterium variabile	1	2.56%	1
Cupriavidus gilardii	1	2.56%	31
Cupriavidus sp. HPC(L)	1	2.56%	11
Cutibacterium granulosum	1	2.56%	193
Desulfovibrio fairfieldensis	1	2.56%	38
Desulfovibrio sp. 3_1_syn3	1	2.56%	70
Desulfovibrio sp. 6_1_46AFAA	1	2.56%	32
Eikenella sp. NML96-A-049	1	2.56%	33
Enhydrobacter aerosaccus	1	2.56%	2845
Enterobacter cloacae	1	2.56%	1
Gemella sp. oral taxon 928	1	2.56%	1
Granulicatella sp. HMSC30F09	1	2.56%	105
Granulicatella sp. HMSC31F03	1	2.56%	89
Herbaspirillum huttiense	2	5.13%	709
Herbaspirillum seropedicae	1	2.56%	936
Herbaspirillum sp. B39	1	2.56%	962
Janibacter indicus	1	2.56%	1
Kocuria palustris	1	2.56%	2
Kocuria rhizophila	3	7.69%	724
Kocuria sp. HMSC066H03	1	2.56%	12

Kocuria sp. ICS0012	1	2.56%	3
Lactobacillus casei	1	2.56%	124
Lactobacillus crispatus	1	2.56%	221
Lactobacillus gasseri	1	2.56%	334
Lactobacillus oris	1	2.56%	749
Lactobacillus paracasei	1	2.56%	208
Lactobacillus rhamnosus	1	2.56%	864
Lactobacillus sakei	1	2.56%	526
Lactobacillus salivarius	1	2.56%	396
Lactobacillus vaginalis	1	2.56%	320
Leuconostoc mesenteroides	1	2.56%	734
Leuconostoc sp. DORA_2	1	2.56%	1
Methylobacterium sp. 77	1	2.56%	963
Methylobacterium sp. ap11	1	2.56%	1
Methylobacterium sp. Leaf123	2	5.13%	4
Methylobacterium sp. Leaf456	1	2.56%	924
Methylobacterium sp. Leaf90	1	2.56%	1
Methylobacterium sp. MB200	3	7.69%	2855
Methylobacterium sp. UNCCL110	1	2.56%	1
Methylobacterium sp. yr596	1	2.56%	8
Methylobacterium variabile	1	2.56%	2
Micrococcus luteus	4	10.26%	594
Micrococcus lylae	1	2.56%	1
Moraxella osloensis	1	2.56%	1390
Mycobacterium chelonae	1	2.56%	793
Mycobacterium gordonae	1	2.56%	1525
Mycobacterium mucogenicum	1	2.56%	593
Mycobacterium sp. Root135	1	2.56%	1
Mycobacterium szulgai	1	2.56%	3
Neisseria sp. HMSC055H02	1	2.56%	213

Neisseria sp. HMSC074B07	1	2.56%	36
Paracoccus sp. 228	2	5.13%	711
Paracoccus tibetensis	1	2.56%	53
Peptoniphilus sp. oral taxon 386	1	2.56%	3
Porphyromonadaceae bacterium COT-184 OH4590	1	2.56%	2
Porphyromonas cangingivalis	2	5.13%	3
Prevotella amnii	6	15.38%	595
Prevotella aurantiaca	7	17.95%	2413
Prevotella copri	1	2.56%	1
Prevotella jejuni	2	5.13%	278
Propionibacterium sp. CC003-HC2	1	2.56%	2
Propionibacterium sp. KPL1844	1	2.56%	144
Pseudoalteromonas luteoviolacea	2	5.13%	8
Pseudomonas fluorescens	1	2.56%	2
Pseudomonas sp. CB1	1	2.56%	16
Pseudomonas sp. FSL W5-0203	1	2.56%	1
Pseudomonas syringae	2	5.13%	5
Pseudonocardia sp. 73-21	1	2.56%	2
Psychrobacter cibarius	1	2.56%	2
Rhodobacter sphaeroides	1	2.56%	181
Rhodopseudomonas pseudopalustris	1	2.56%	1
Rudaea cellulosilytica	1	2.56%	1
Schwartzia succinivorans	1	2.56%	17
Shewanella decolorationis	1	2.56%	1182
Sphingobium lucknowense	1	2.56%	1
Sphingobium sp. C100	1	2.56%	2
Sphingomonas hankookensis	1	2.56%	2

Sphingomonas sp. 67-36	2	5.13%	858
Sphingomonas sp. CCH9-F2	2	5.13%	5175
Sphingomonas sp. Leaf10	1	2.56%	1
Sphingomonas sp. Leaf23	1	2.56%	2
Sphingomonas sp. Leaf257	1	2.56%	1
Sphingomonas sp. OV641	1	2.56%	1
Staphylococcus capitis	1	2.56%	1
Staphylococcus epidermidis	6	15.38%	1324
Staphylococcus saccharolyticus	1	2.56%	1
Stenotrophomonas sp. HMSC10F06	1	2.56%	864
Streptococcus massiliensis	1	2.56%	182
Streptococcus peroris	1	2.56%	374
Streptococcus sp. 263_SSPC	1	2.56%	389
Streptococcus sp. AS14	1	2.56%	55
Streptococcus sp. CCH8-G7	1	2.56%	28
Streptococcus sp. CM6	1	2.56%	53
Streptococcus sp. HMSC056D07	1	2.56%	1
Streptococcus sp. HMSC071D03	1	2.56%	1
Streptococcus sp. HMSC34B10	1	2.56%	356
Streptococcus sp. oral taxon 056	1	2.56%	768
Streptococcus sp. SR1	1	2.56%	63
Streptococcus suis	1	2.56%	2
Tepidiphilus margaritifer	1	2.56%	1360
Tepidiphilus thermophilus	1	2.56%	3165
Tissierella praeacuta	3	7.69%	542
Treponema porcinum	1	2.56%	25
uncultured bacterium	1	2.56%	13
Veillonella sp. DORA_B_18_19_23	2	5.13%	72

	Veillonella sp. oral taxon 158	2	5.13%	49
D_NORA (16 samples in total)				
22	*Alishewanella agri*	1	6.25%	2
	Alishewanella jeotgali	1	6.25%	2
	Bacillus pseudomycoides	1	6.25%	2
	Bacteroides vulgatus	1	6.25%	2
	Bilophila sp. 4_1_30	1	6.25%	150
	Bilophila wadsworthia	1	6.25%	135
	Methylobacterium gossipiicola	1	6.25%	2
	Methylobacterium platani	1	6.25%	2
	Methylobacterium sp. B34	1	6.25%	4
	Methylobacterium sp. C1	1	6.25%	2
	Methylobacterium sp. Leaf125	1	6.25%	3
	Methylobacterium sp. Leaf361	1	6.25%	3
	Neisseria sp. HMSC077D05	1	6.25%	160
	Porphyromonas sp. COT-052 OH4946	1	6.25%	60
	Prevotella shahii	1	6.25%	3
	Prevotella timonensis	1	6.25%	3
	Pseudomonas psychrotolerans	1	6.25%	47
	Streptococcus sp. HMSC066F01	1	6.25%	49
	Tessaracoccus flavescens	1	6.25%	2
	Tessaracoccus flavus	1	6.25%	2
	Treponema phagedenis	1	6.25%	2
	Verrucomicrobia bacterium IMCC26134	2	12.50%	45
D_RA (7 samples in total)				
29	*Acidaminococcus sp. CAG:917*	1	14.29%	2
	Actinomyces graevenitzii	1	14.29%	454
	Actinomyces radicidentis	1	14.29%	2

Actinomyces sp. HPA0247	1	14.29%	229
Actinomyces sp. pika_113	1	14.29%	2
Bacteroides sp. CAG:709	1	14.29%	2
Cellulomonas sp. HZM	1	14.29%	1
Corynebacterium provencense	1	14.29%	1
Cryptobacterium curtum	1	14.29%	44
Cyanothece sp. CCY0110	1	14.29%	1
Dermacoccus sp. Ellin185	1	14.29%	2
Gallionellales bacterium RBG_16_56_9	1	14.29%	2
Gramella forsetii	1	14.29%	1
Isoptericola variabilis	1	14.29%	93
Leptotrichia goodfellowii	1	14.29%	2
Moorella glycerini	1	14.29%	1
Moritella viscosa	1	14.29%	6
Olsenella sp. DNF00959	1	14.29%	1
Paenibacillus sp. FSL H8-0259	1	14.29%	1
Paenibacillus sp. St-s	1	14.29%	1
Propionibacterium freudenreichii	1	14.29%	1
Rhodococcus wratislaviensis	1	14.29%	2
Simonsiella muelleri	1	14.29%	2
Streptococcus sp. C150	1	14.29%	1
Streptococcus sp. HSISM1	1	14.29%	73
Veillonella sp. ACP1	1	14.29%	105
Veillonella sp. HPA0037	1	14.29%	68
Veillonella sp. ICM51a	1	14.29%	73
Xanthomonas citri	1	14.29%	2

C 7 List of core species (> 50% prevalence, > 0.2% relative abundance) of each group in periodontally healthy sites

Groups	Numbers of species	Species
HC, CCP, NORA, RA	17	*Prevotella nigrescens*
		Actinomyces massiliensis
		Actinomyces oris
		Actinomyces viscosus
		Actinomyces naeslundii
		Fusobacterium nucleatum
		Actinomyces dentalis
		Actinomyces israelii
		Rothia dentocariosa
		Actinomyces gerencseriae
		Actinomyces sp. oral taxon 414
		Actinomyces sp. oral taxon 175
		Streptococcus pneumoniae
		Streptococcus sanguinis
		Tannerella forsythia
		Corynebacterium matruchotii
		Actinomyces sp. oral taxon 171
HC, CCP, NORA	2	*Streptococcus oralis*
		Pseudopropionibacterium propionicum
HC, CCP, RA	6	*Olsenella sp. oral taxon 807*
		Actinomyces sp. oral taxon 170
		Treponema medium
		Bacteroidetes oral taxon 274
		Veillonella parvula
		Campylobacter gracilis
HC, NORA, RA	1	*Actinobaculum sp. oral taxon 183*

HC, CCP	3	*Capnocytophaga ochracea*
		Treponema socranskii
		Prevotella oris
HC, NORA	1	*Actinomyces sp. HMSC08A09*
HC, RA	3	*Actinomyces johnsonii*
		Tannerella sp. oral taxon HOT-286
		Actinomyces sp. oral taxon 448
CCP, RA	2	*Treponema denticola*
		Veillonella dispar
HC	3	*Prevotella denticola*
		Streptococcus mitis
		Actinomyces sp. oral taxon 849
CCP	5	*Prevotella sp. oral taxon 317*
		Campylobacter showae
		Campylobacter rectus
		Selenomonas sp. CM52
		Treponema vincentii
NORA	1	*Porphyromonas gingivalis*
RA	4	*Bradyrhizobium sp. BTAi1*
		Candidatus Bacteroides
		periocalifornicus
		Alloprevotella tannerae
		Porphyromonas endodontalis

C 8 List of core species (> 50% prevalence, > 0.2% relative abundance) of each group in periodontally diseased sites

Groups	number of species	Species
HC, CCP, NORA, RA	25	*Prevotella nigrescens*
		Candidatus Bacteroides periocalifornicus
		Olsenella sp. oral taxon 807
		Actinomyces massiliensis
		Actinomyces sp. oral taxon 448
		Actinomyces oris
		Actinomyces viscosus
		Actinomyces naeslundii
		Fusobacterium nucleatum
		Treponema medium
		Actinomyces dentalis
		Bacteroidetes oral taxon 274
		Actinomyces israelii
		Rothia dentocariosa
		Prevotella pleuritidis
		Treponema denticola
		Actinomyces sp. oral taxon 414
		Actinomyces gerencseriae
		Actinomyces sp. oral taxon 175
		Treponema socranskii
		Corynebacterium matruchotii
		Alloprevotella tannerae
		Tannerella forsythia
		Prevotella oris
		Campylobacter gracilis
HC, CCP, NORA	1	*Campylobacter showae*

HC, CCP, RA	13	*Prevotella denticola*
		Treponema maltophilum
		Treponema lecithinolyticum
		Fretibacterium fastidiosum
		Campylobacter rectus
		Capnocytophaga ochracea
		Selenomonas sputigena
		Selenomonas sp. CM52
		Parvimonas micra
		[Eubacterium] brachy
		Porphyromonas endodontalis
		Treponema vincentii
		Treponema sp. OMZ 838
HC, NORA, RA	2	*Prevotella sp. oral taxon 317*
		Actinobaculum sp. oral taxon 183
HC, CCP	1	*Tannerella sp. oral taxon HOT-286*
HC, NORA	2	*Actinomyces sp. oral taxon 171*
		Pseudopropionibacterium propionicum
HC, RA	1	*Dialister invisus CAG:218*
CCP, RA	2	*Filifactor alocis*
		Veillonella parvula
NORA, RA	1	*Streptococcus oralis*
HC	3	*Prevotella sp. HMSC073D09*
		Mogibacterium sp. CM50
		Prevotella conceptionensis
CCP	2	*Bradyrhizobium sp. BTAi1*
		Prevotella intermedia
NORA	6	*Actinomyces sp. HMSC08A09*
		Cardiobacterium hominis
		Streptococcus mitis

		Streptococcus pneumoniae
		Porphyromonas gingivalis
		Streptococcus sanguinis
RA	4	*Selenomonas noxia*
		Kingella oralis
		Oribacterium sp. oral taxon 078
		Leptotrichia wadei

C 9 Bacterial genera associated with different groups in periodontally healthy site samples from individuals without periodontitis

Group	* Genera	# Corrected *P*
CCP	*Mogibacterium*	0.0277
NORA	*Cardiobacterium*	0.0277
	Methylobacterium	0.0080
	Sphingomonas	0.0394
RA	*Achromobacter*	0.0471
	Bacillus	0.0316
	Bradyrhizobium	0.0394
	Delftia	0.0471
	Escherichia	0.0394
	Staphylococcus	0.0199
	Stenotrophomonas	0.0471

*: Genera with significantly higher relative abundance between groups were investigated by using the permutation test (one-sided *signassoc* function, R).
#: Sidak's correction was applied for multiple testing

C 10 Uniquely detected species in HC, CCP, NORA and RA groups in periodontally healthy sites samples from individuals without periodontitis.

Number of unique species	Species	Number of infected samples	%Prevalence	Total counts
HC (19 samples)				
23	*[Eubacterium] sulci*	1	5.26	26
	Actinomyces cardiffensis	1	5.26	674
	Granulicatella elegans	1	5.26	106
	Haemophilus haemolyticus	2	10.53	436
	Haemophilus sp. HMSC068C11	1	5.26	72
	Haemophilus sp. HMSC61B11	1	5.26	42
	Micrococcus luteus	1	5.26	12
	Mogibacterium pumilum	2	10.53	456
	Neisseria cinerea	2	10.53	387
	Neisseria polysaccharea	1	5.26	39
	Nitrincola lacisaponensis	1	5.26	78
	Paracoccus sp. 228	1	5.26	198
	Paracoccus sp. PAMC 22219	1	5.26	67
	Paracoccus sp. S4493	1	5.26	38
	Propionibacterium sp. 409-HC1	1	5.26	15
	Propionibacterium sp. 434-HC2	1	5.26	20
	Rhodopseudomonas palustris	2	10.53	50
	Selenomonas ruminantium	1	5.26	18
	Slackia exigua	3	15.79	650
	Streptococcus sp. CCH8-G7	1	5.26	123
	Streptococcus sp. NLAE-zl-C503	1	5.26	58
	Streptococcus sp. oral taxon 058	2	10.53	134
	Verrucomicrobia bacterium IMCC26134	2	10.53	126

CCP (13 samples)				
60	*Actinomyces graevenitzii*	1	7.69	8
	Actinomyces sp. ICM54	1	7.69	108
	Afipia broomeae	1	7.69	6
	Aggregatibacter actinomycetemcomitans	2	15.38	1632
	Bifidobacterium adolescentis	1	7.69	7
	Bifidobacterium longum	1	7.69	8
	Bifidobacterium moukalabense	1	7.69	7
	Bifidobacterium sp. MSTE12	2	15.38	125
	Bradyrhizobium liaoningense	1	7.69	9
	Bradyrhizobium sp.	1	7.69	8
	Eikenella sp. NML080894	1	7.69	81
	Eikenella sp. NML99-0057	1	7.69	81
	Lactobacillus casei	1	7.69	48
	Lactobacillus fermentum	1	7.69	161
	Lactobacillus gasseri	1	7.69	1760
	Lactobacillus johnsonii	1	7.69	59
	Lactobacillus paracasei	1	7.69	68
	Lactobacillus rhamnosus	1	7.69	897
	Lactobacillus salivarius	1	7.69	1923
	Lactobacillus vaginalis	1	7.69	321
	Neisseria sp. HMSC03D10	1	7.69	13
	Neisseria sp. HMSC055H02	1	7.69	22
	Neisseria sp. HMSC056A03	1	7.69	40
	Neisseria sp. HMSC064E01	1	7.69	34
	Neisseria sp. HMSC065D04	1	7.69	83
	Neisseria sp. HMSC068C04	1	7.69	33
	Neisseria sp. HMSC069H12	1	7.69	27
	Neisseria sp. HMSC06F02	1	7.69	31

Neisseria sp. HMSC072F04	1	7.69	17
Neisseria sp. HMSC074B07	1	7.69	43
Neisseria sp. HMSC077D05	2	15.38	182
Neisseria sp. HMSC15C08	1	7.69	15
Neisseria sp. HMSC70E02	1	7.69	72
Neisseria subflava	1	7.69	24
Paraprevotella clara CAG:116	1	7.69	8
Peptostreptococcaceae bacterium oral taxon 113	1	7.69	6
Porphyromonas gulae	2	15.38	40
Prevotella bryantii	1	7.69	8
Prevotella jejuni	1	7.69	43
Prevotella shahii	1	7.69	21
Prevotella sp. C561	1	7.69	98
Prevotella sp. KH2C16	1	7.69	13
Prevotella sp. oral taxon 306	2	15.38	297
Pseudoramibacter alactolyticus	1	7.69	53
Pyramidobacter piscolens	1	7.69	17
Rhodobacter sphaeroides	1	7.69	64
Selenomonas sp. F0473	1	7.69	1110
Selenomonas sp. oral taxon 136	1	7.69	58
Selenomonas sp. oral taxon 478	1	7.69	99
Streptococcus downei	1	7.69	28
Streptococcus sobrinus	1	7.69	435
Streptococcus sp. 263_SSPC	1	7.69	56
Thioflexothrix psekupsii	1	7.69	6
uncultured bacterium	1	7.69	22
Veillonella sp. ACP1	1	7.69	78
Veillonella sp. DORA_B_18_19_23	1	7.69	44

	Veillonella sp. HPA0037	1	7.69	48
	Veillonella sp. ICM51a	1	7.69	59
	Vibrio parahaemolyticus	1	7.69	6
	Xanthomonas citri	1	7.69	7
NORA (12 samples)				
33	*Alkalibacterium thalassium*	1	8.33	158
	Fusobacterium sp. CM21	1	8.33	11
	Halothiobacillus sp. 28-55-5	1	8.33	6
	Methylobacterium aquaticum	1	8.33	5
	Methylobacterium extorquens	2	16.67	40
	Methylobacterium phyllosphaerae	1	8.33	10
	Methylobacterium pseudosasicola	2	16.67	37
	Methylobacterium salsuginis	1	8.33	5
	Methylobacterium sp. 285MFTsu5.1	2	16.67	24
	Methylobacterium sp. AMS5	1	8.33	6
	Methylobacterium sp. ARG-1	2	16.67	45
	Methylobacterium sp. B1	1	8.33	16
	Methylobacterium sp. C1	1	8.33	7
	Methylobacterium sp. Leaf361	1	8.33	5
	Methylobacterium sp. Leaf456	1	8.33	6
	Methylobacterium sp. UNC378MF	2	16.67	44
	Methylobacterium sp. UNCCL125	1	8.33	20
	Novosphingobium sp. Chol11	2	16.67	18
	Olsenella profusa	1	8.33	330

	Peptostreptococcaceae bacterium AS15	2	16.67	204
	Prevotella aurantiaca	1	8.33	14
	Prevotella bivia	1	8.33	12
	Prevotella multisaccharivorax	1	8.33	21
	Prevotella sp. HMSC077E09	1	8.33	12
	Prevotella timonensis	1	8.33	9
	Shuttleworthia satelles	1	8.33	9
	Sphingomonas sp. 67-36	1	8.33	6
	Sphingomonas sp. CCH9-F2	1	8.33	8
	Streptococcus infantis	2	16.67	54
	Streptococcus sp. oral taxon 056	1	8.33	244
	Streptomyces purpurogeneiscleroticus	2	16.67	42
	Williamsia muralis	1	8.33	162
	Williamsia sp. D3	1	8.33	151
RA (8 samples)				
30	*Aerococcus sp. HMSC23C02*	1	12.50	140
	Aeromicrobium sp. Leaf245	1	12.50	203
	Aeromicrobium sp. Leaf272	1	12.50	50
	Bacillus azotoformans	1	12.50	1464
	Beggiatoa sp. PS	1	12.50	139
	Candidatus Saccharibacteria bacterium RAAC3_TM7_1	1	12.50	2
	Cloacibacterium normanense	1	12.50	120
	Clostridioides difficile	2	25.00	371
	Delftia acidovorans	2	25.00	113
	Enterococcus casseliflavus	1	12.50	221
	Janibacter indicus	1	12.50	175
	Klebsiella aerogenes	1	12.50	174

Lactobacillus brevis	1	12.50	2
Leuconostoc mesenteroides	1	12.50	130
Pedobacter panaciterrae	1	12.50	2
Porphyromonas sp. CAG:1061	1	12.50	1
Porphyromonas uenonis	1	12.50	1
Prevotella sp. AGR2160	1	12.50	1
Prevotella sp. CAG:1058	1	12.50	1
Prevotella sp. oral taxon 299	1	12.50	100
Prevotellaceae bacterium HUN156	1	12.50	1
Proteiniphilum acetatigenes	1	12.50	121
Pseudomonas mendocina	1	12.50	103
Pseudomonas syringae	1	12.50	118
Sanguibacteroides justesenii	1	12.50	227
Solibacillus silvestris	1	12.50	185
Staphylococcus epidermidis	1	12.50	10
Streptococcus dysgalactiae	1	12.50	170
Streptococcus sp. GMD6S	1	12.50	118
Vibrio cholerae	1	12.50	1

Appendix D Publications and Presentations

Publication:

CHENG, Z., MEADE, J., MANKIA, K., EMERY, P. & DEVINE, D. A. 2017. Periodontal disease and periodontal bacteria as triggers for rheumatoid arthritis. *Best Practice & Research Clinical Rheumatology,* 31**,** 19-30.

Conference abstract:

CHENG, Z., DO, T., MANKIA, K., MEADE, J. L., HUNT, L., NAM, J., TUGNAIT, A., SPEIRS, A., CLEREHUGH, V., EMERY, P. & DEVINE, D. 2017. The subgingival microbiomes in periodontitis and health of individuals with rheumatoid arthritis and at risk of developing rheumatoid arthritis. *Journal of Oral Microbiology,* 9**,** 1325216.

Oral presentation:

Cheng, Z., Do, T., Meade, J., Hunt, L., Mankia, K., Nam, J., Tugnait, A., Speirs, A., Clerehugh, D.V., Emery, P., Devine, D**.** Periodontal Disease and the Subgingival Microbiome in Preclinical Rheumatoid Arthritis. Oral Microbiology & Immunology Group of the British Society for Oral & Dental Research (OMIG), Wales, UK: 2016.

Cheng, Z., Do, T., Mankia, K., Meade, J.L., Hunt, L., Nam, J., Tugnait, A., Speirs, A., Clerehugh, V., Emery, P., Devine, D. Preliminary Analyses of the Microbiome in Rheumatoid Arthritis-associated Periodontitis. School of dentistry research day, University of Leeds. Leeds, UK: 2017. (First prize)

Poster presentation:

Cheng, Z., Do, T., Mankia, K., Meade, J.L., Hunt, L., Nam, J., Tugnait, A., Speirs, A., Clerehugh, V., Emery, P., Devine, D. Preliminary Analyses of the Microbiome in Rheumatoid Arthritis-associated Periodontitis. International Association for Dental Research (IADR), San Francisco, USA: 2017.