METAGENOMIC ANALYSIS OF PERIODONTAL BACTERIA ASSOCIATED WITH GENERALIZED AGGRESSIVE PERIODONTITIS

THESIS

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in the Graduate School of The Ohio State University

By

Khaled Altabtbaei

Graduate Program in Dentistry

The Ohio State University

2016

Master's Examination Committee:

Dr. Purnima Kumar, Advisor

Dr. Pooja Maney

Dr. John Walters

Copyright by

Khaled Altabtbaei

2016

ABSTRACT

Background: Generalized aggressive periodontitis (GAP) is a microbial disease with distinct clinical profile to the more clinically prevalent chronic periodontitis. Until recently, the capability to study the microbiome of the disease has not been possible.

Methods: 17 patients with GAP were enrolled. Patients were excluded if they had uncontrolled systemic disease, pregnancy, or had received antibiotic therapy in the last 3 months. Paper point samples were taken from the deep and shallow sulci of the affected patients. Genomic DNA was isolated, sheared, size selected and sequenced using Illumina. Sequences were filtered, and functionally annotated using the MG-RAST pipeline. The intra- and inter-group functional profiles were compared between the deep and shallow pockets, as well as with healthy individuals and with patients with chronic periodontitis (CP), using Wald test and FDR (DESeq2).

Results: Statistical analysis revealed no significant differences in the functional diversity between the shallow and deep pockets in generalized aggressive periodontitis, thus establishing that globalized functional capacity in GAP. Deep sites in GAP have higher functional capability than healthy sites, especially in regards to chemotaxis and motility.

Finally, the functional capability of generalized aggressive periodontitis is distinct from chronic periodontitis.

Conclusion: Distinct functional differences are found between health and disease in generalized aggressive periodontitis. These pockets, despite looking clinically healthy, have the same functional capacity as the deeper pockets in those patients. Moreover, these functional capability are distinct from the disease state in chronic periodontitis.

Dedicated to
My family; close and far

ACKNOWLEDGEMENTS

I wish to thank Dr. Purnima Kumar for giving me the opportunity to work on this project. Her presence, guidance and availability has made it possible for me to complete this thesis.

I also want to thank my committee members, Dr. John D. Walters and Dr. Pooja Maney. Their advice, and contribution was crucial to the development of this work.

I also wish to mention that this project could not have worked without the precious help of Dr. Shareef Dabdoub, Dr. Matt Mason, and Dr. Sukirth Ganesan. They have made it possible for me to conduct the laboratory manipulation, and formulate the data into a cogent thesis.

I want to thank as well all the members of the Ohio State Periodontology Department. Patient recruitment would have been nearly impossible without their vigilance and assistance. I would also like to thank Dr. Laura Porra for her help with recruiting patients in LSU.

Finally, I would like to thank my family and friends for their constant support.

VITA

March 9, 1985 …Born: Kuwait city, Kuwait

2009 … Obtained B. D. M.: University of Kuwait

FIELD OF STUDY

Major Field: Dentistry

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

Periodontitis

Periodontitis is a microbially-driven inflammatory disease that is characterized by destruction of the supporting tissues of the teeth (periodontium).[1] According to the 2009-2012 NHANES data, 47% of adults in the US suffer from periodontitis, with 8.9% suffering from severe forms of the disease.[2] The disease is diagnosed based on clinical findings of inflammation, pocketing and attachment loss, and hence, it is possible that multiple diseases with the same clinical presentation, i.e., inflammation and destruction of the periodontium, may have multifactorial etiologies.

In 1999, the American Academy of Periodontology, based on the available knowledge, grouped periodontitis into two categories based on clinical, familial and immunological presentation: chronic, and aggressive.[3] Despite the fact that each category may be composed of a heterogeneous group of diseases, the distinction between chronic and aggressive periodontitis is important clinically, since they respond differently to therapy, especially, non-surgical therapy.[4,5] Further, the long-term prognosis of the affected teeth, and dental implants placed in these patients differs between the two disease phenotypes.[6]

Aggressive Periodontitis

Aggressive periodontitis is a disease that is classified based on evidence of familial aggregation and a clinical presentation of rapid attachment loss in systemically healthy young adults. Patients with this disease typically demonstrate significant destruction of supporting structures that is not consistent with amount of microbial deposits. It is sub-classified into localized aggressive periodontitis (LAP) and generalized aggressive periodontitis (GAP). LAP affects the first molar/incisor regions with interproximal attachment loss on at least 2 permanent teeth, one of which is a first molar, with involvement of no more than 2 teeth other than the first molars and incisors. GAP patients have at least 3 permanent teeth affected by interproximal attachment loss other than the first molars and incisors.[7]

Localized aggressive periodontitis usually occurs at the time of puberty, while generalized aggressive periodontitis generally occurs in people under the age of 30. At the time of reclassification, evidence suggested that there is a robust serum antibody response to the infecting agents, in contrast to individuals with generalized aggressive periodontitis, who have poor serum response. Both conditions were thought to be associated with elevated proportions of *Aggregatibacter* (previously *Actinobacillus*) *actinomycetemcomitans*, and in some populations, *Porphyromonas gingivalis*. Phagocyte abnormalities and hyperresponsive macrophages were also associated with the two disease entities.[7]

Epidemiology of Aggressive Periodontitis

In the United States, as part of a comprehensive nationwide survey of school-age children, approximately 14,013 adolescents (Grade 8-12) were examined. It was estimated that aggressive periodontitis occurs in 0.4% in adolescents 13-15 years old, and 0.8% in adolescents 16-19 years old, establishing an increasing prevalence with age. Based on that survey, Löe and Brown estimated that there are 70,000 and 17,000 adolescents suffering from LAP and GAP, respectively. Of those that are 14-17 years old age, 0.53% had localized aggressive periodontitis, and 0.13% had generalized aggressive periodontitis.[8] In a follow-up study, 91 subjects were re-examined 6 years later, Brown found that the areas with pre-existing attachment loss continued to have attachment loss, and some unaffected teeth started developing attachment loss. 35% of those with localized aggressive periodontitis had progressed to become generalized, while 62% had stayed the same. Of those that had generalized aggressive periodontitis, 82% continued to have the same diagnosis.[9]

Globally, similar studies have been conducted, revealing similar statistics. In 14-17 year old Dutch population, the prevalence of aggressive periodontitis was estimated to be 0.1-0.2%. A Swiss study in 19-20 years old found the prevalence to be 0.13%[10] In the UK, it was estimated to be 0.1-0.2%.[11] Studies in Chile and Brazil found the prevalence to be 0.32% and 5.5%, respectively. In Asia, studies in school-aged children in Turkey and Iran found the prevalence to be 0.6% and 0.13%, respectively. Finally, in a Sudanese children population, the prevalence was found to be 3.4%.[10]

Demographics of Aggressive Periodontitis

It was estimated that the odds ratio for detecting localized aggressive periodontitis in a 17 years old US adolescent compared to a 14 years old to be 3.8. There is also evidence that the disease is more prevalent in males than in females, 0.78% compared to 0.52%, respectively.[10] In those that are affected by GAP, the male to female ratio was 4.3:1.[8] Similar result from a 3-year longitudinal Brazilian study corroborate the increase prevalence of aggressive periodontitis at 13 and 16 years old.[10]

There is also a predilection for the disease to occur in young people of African origin. In the national survey in the United States, the prevalence was highest in Black Americans (2.6%), followed by Hispanics (0.5%), while it was the lowest in Caucasians (0.06%).[10] This distinct difference was found in other studies worldwide. In the UK, 1% of those of Afro-Caribbean origins were affected compared to Asians (0.3%) and Caucasians (0.1%).[11] A Sudanese study has found that there is a higher prevalence of aggressive periodontitis in children of African origins (6%) than those of Afro-Arab origins (2.3%).[10]

Aggressive and Chronic periodontitis - Clinical Difference

Both conditions, in essence, encompass a variety of diseases that cause an immune-inflammatory response against a dysbiotic microbiome.[12] Several factors can affect this inflammation, such as genetic factors, smoking, and systemic diseases that modify these immunologic disorders.[3,13,14]

The main clinical feature that distinguishes the two variations of the disease is the rate of attachment loss. Chronic periodontitis is characterized by the slow to moderate

progression of periodontal destruction, while aggressive periodontitis is characterized by the rapid destruction of the periodontium.[15] This difference in the rate of progression is presumably due to a myriad of factors that modify the inflammatory response, for example: the genetic makeup and the inflammatory response of the patient, the microbial antigens within in the microbiome, and environmental factors.[16] A secondary clinical feature that has been used to distinguish between the two diseases is the amount of local factors such as plaque and calculus compared to the attachment loss. Aggressive periodontitis generally has few local factors compared to the attachment loss, while patients with chronic periodontitis demonstrate an extensive amount of plaque and calculus. However, the amount of local factors is not always distinct between the two entities, as some patients with aggressive periodontitis, especially those with generalized aggressive periodontitis, present with extensive local factors.[12] This can lead to confusion as to whether the patient exhibits chronic or generalized aggressive forms of the disease.

Periodontal health: A Polymicrobial Condition

In periodontally healthy pockets, a biofilm exists that has a small number of microorganisms with a limited range of microorganisms. Each one of these microorganisms come with its own set of genetic composition and phenotypic expression. These microorganisms, through quorum sensing, have their phenotypic expression either enhanced or suppressed depending on its neighboring bacteria. Some species also have the ability to shelter the microbiome from foreign insults by producing the antimicrobial resistant agents such as b-lactamase, which reduces the fluctuations that could occur in the biofilm.[17]

This allows the bacteria to work together in consortium to confer properties and functions that no single cell is able to fully perform on its own, much like a multicellular organism. A well-documented example is the co-dependence of the Veillonellae species on the lactate that is produced by the Streprococcus species as food source. This biofilm exists in harmony with the host, and possibly incur beneficial effects on the host such as create a barrier for non-indigenous pathogens to colonize the biofilm, inhibit the overgrowth of opportunistic microorganisms, mature the host immune system to recognize pathogenic microorganisms versus those associated with health (Reviewed - Kumar 2005[18]). This microbiome is stable in its composition, and in some people, remains stable over time, even in cases of changes in the oral environment such as ingestion of food.[19,20]

Bacterial Identification Methods

Much of the knowledge regarding the bacterial etiology of periodontitis has been derived using culture-based and targeted molecular methods. Culturing creates a bias in the identification of etiologic agents, since it promotes the growth of certain kinds of microorganisms in preference to others. While this type of identification methodology is important in monoinfections, it is less important in polymicrobial diseases. The metabolic pathways expressed in planktonic microorganisms are different than when they are part of a biofilm. For example, the genomic expression of *Porphyromonas gingivalis* (P.g.) in its planktonic state is 18% different than that when it is in biofilm. Moreover, some fastidious microorganisms, such as spirochetes, require other microorganisms before colonization and cannot be grown solitarily.[21] It is also currently not possible to culture

every species, even in combination with other known species. A method that circumvents some of these issues is DNA-DNA hybridization. This method uses the properties of hybridization of the DNA double helix to identify the microorganism, where one strand from a specific organism is used as a probe that would hybridize with the complementary strand of the same species if it were present in the biofilm. This method allows for relatively fast identification a wide variety of microorganisms without the need for culturing. This method obviously requires prior knowledge of the target species in a community and requires stringent hybridization conditions to keep heterogeneous hybridization from other species low. Based on that, it becomes obvious that it still does not address the variety of microorganisms that are still not known, and its usage still requires knowledge on what is in the microbiome before testing it.

As technology progressed, it became possible to identify the species in the microbiome through open-ended methods. A highly conserved region of prokaryotic and archaeal DNA that codes for rDNA, hence termed as 16S rDNA, has been used to identify the constituents of the microbiome. By comparing this region in unknown bacteria to databases of bacteria with known 16s rDNA regions, it is possible to identify the bacteria and archaea in an open ended fashion. Once the organism is identified, its functional capability can be inferred. Currently, this method is a relatively cheap and fast method to identify the bacteria in the microbiome with moderate resolution. Since this method uses the highly conserved portion of the prokaryotic gene and as such identifies the genera/species that the bacteria is, it does not take into account mutations or horizontal gene transfers that may have occurred. This is especially true in some

genera/species where the species that have recently reclassified, notably those belonging to the families *Actinomyces*, and *Enterobacteriaceae*.[22,23]

Advances in the sequencing technology has made it possible to sequence the entire genome of the bacterial community, which is known as the metagenome. Since the entire genome of the microbiome is analyzed, the identification of the bacterial constituents can be done based on similarity of long sequences of DNA as opposed to only the 16s rDNA sequences. This provides higher identification accuracy at the bacterial species and subspecies level.[24] Moreover, the DNA sequences can be used to identify the genes and their actual pathogenic potential.

Microbiology of Periodontitis at the Community Level

The theory that there are specific microbial organisms that cause periodontitis was proposed in 1976 by Loesche et al (reviewed by Rosier et al., Frontiers in Cellular and Infection Microbiology). Socransky et al grouped bacteria based on their presence in healthy individuals or individuals with chronic periodontitis, that is, bacteria that are health compatible or associated with disease. His group has observed that some clusters of bacteria appear to be mostly associated with either health or disease. A distinct complex made of *Porphyromonas gingivalis, Tannerella forsythia*, and *Treponema denticola,* which they called the red complex, was mostly found in the patients with chronic periodontitis. Moreover, some clusters of bacteria were found to be rarely associated with other clusters of bacteria. For example, the green and yellow complexes were less commonly associated with the members of the more pathogenic complexes. This lead to the speculation that the presence of certain groups of bacteria could lead to

making the biofilm less hospitable to other bacteria. Furthermore, another pathogenic complex was found to be closely associated with the red complex, albeit less so with disease, which was termed as the orange complex. It was speculated that the presence of the orange complex is needed for the colonization of the red complex, and as such show a temporal relationship in colonization as well as the co-dependence of different species.[25] Open-ended microbial identification methods have also corroborated this finding.[26,27] Further research has led to a relatively small group of microorganisms that are associated with disease: *P. gingivalis, P. endodontalis T. forsythia, A. actinomycetemcomitans, Prevotella intermedia, P. melaninogenica, P. denticola, P. nigrescens, P. corporis, P. disiens, Fusobacterium nucleatum, Parvimonas micra, Eikenella corrodens , Capnocytophaga gingivalis, Treponema denticola, T. socranskii, T. maltophilum, Treponema sp. Smibert-3, T. lecithinolyticum, Treponema putidum sp. nov, Eubacterium nodatum, E. saphenum, Campylobacter rectus, Filifactor alocis, Cryptobacterium curtum, Mogibacterium timidum, Peptostreptococcus magnus, Slackia exigua, Enterococcus faecalis, Escherichia coli and Bartonella sp.*[28]

Microbiota associated with GAP

Our knowledge of aggressive periodontitis is limited due to its comparatively rare prevalence, which leads to difficulties in garnering a large cohort of patients for study.[29] Moreover, due to changes in disease identification criteria over the last six decades, studies conducted done prior to the 1999 reclassification did not take into consideration the difference between localized and generalized aggressive periodontitis, and therefore translation of those findings to be applicable to current disease classification is

problematic.[17,30] These early studies grouped both aggressive disease entities together, or simply did not identify whether the sampled patients have either form.[31,32] Finally, direct comparison of different results is difficult due to evidence that different isolation methods, probing/sampling sequence, number of sampled teeth, and paper point size could change the microbial profile of the sample.[33,34] That being said, previous efforts have shed light on the microbiota present in GAP.

A pilot investigation by Laksmana examined the 16s rDNA of 2 subjects with GAP and identified 208 species/phylotypes, with 129 species/phylotypes that were shared by both patients. 28-42 species represented 90% of all subgingival bacteria in each sample. *F. nucleatum, Prevotella spp., Porphyromonas gingivalis, A. actinomycetemcomitans* were identified in both samples.[35]

Faveri looked at the 16s rDNA sequences of the microbiota in 10 patients with GAP. 57% of the microbiome were found to be from species have not been previously cultivated. From these samples, 110 species-level taxa were detected, with 70 of them being prevalent in the majority of individuals. The most prevalent genera were *Selemonas* and *Streptococcus*. Other commonly identified genera in these subjects included *Eubacterium* and *Peptostreptococcus*, Members of the phyla *Spirocheta* and *Actinobacteria* were also identified. Interestingly, none of the red complex pathogens, *F. nucleatum*, or *A. actinomycetemcomitans* were found in any of the samples.[36]

CP and GAP: Microbial difference

In 2002, a meta-analysis by Mombelli looked at the microbial difference between aggressive and chronic periodontitis in studies that specifically compared chronic and

aggressive periodontitis (both forms) based on the prevalence and frequency of detection of the five species- *Porphyromonas gingivalis, Prevotella intermedia, Aggregatibacter actinomycetemcomitans, Tanerella forsythia*, and *Campylobacter rectus*. The studies included in this meta-analysis used a variety of methods to detect the microorganisms, such as culturing, DNA-probes, and indirect immunofluorescence. It was found that the presence or absence of these microorganisms did not discriminate between the two disease entities. Since then, a few studies have explicitly compared the microbiomes of the two disease entities. Since there is evidence that the microbiome is specific to the race of the individual[37], the following studies will be presented depending on the geographic region.

Europe

In Germany, Riep (GAP=44, CP=46;16s rDNA) found that A.a., P.g., P.i., T.f., Treponemas denticola-like, treponema lecithinolyticum, Campylobacter rectus, Capnocytophaga ochracea, Fusobacterium spp., and F. nucleatum. Significantly more Treponema lecithinolyticum found in GAP vs CP. The rest were not significantly different. [38] In the UK (CP=183, GAP=84; PCR), Nibali found no association with the two clinical pictures and the detection of A.a. And/or P.g.[39]

South America

In Brazil, Rescala et al compared the abundances of 40 different species using DNA-DNA hybrization, and found no difference between generalized aggressive and chronic periodontitis (GAP=17,CP=20).[40] The same conclusion was reached by Lourenço

et al after analyzing 16S rDNA sequencing of 24 GAP and 35 CP patients. Comparisons of all species have found that detection of Aa, Cardiobacterium hominis, Peptostreptococcaceae sp., P. alactolyticus, and absence of Fretibacterium spp., Fusobacterium naviforme/Fusobacterium nucleatum ss vincentii and Granulicatella adiacens/Granulicatella elegans were associated with GAP compared to CP.[41]

In a Chilean study (GAP=6, CP=17) that used microscopic identification of 8 different species (*A.actinomyecetemcomitans, P.gingivalis, P. intermedia/nigrescens, E. corrodens, F. nucleatum, Capnocytophaga sp., C. rectus, P. micra*), Gajardo found that C. rectus was isolated significantly more frequently and in larger numbers in aggressive periodontitis compared to chronic periodontitis.[42] Casarin (GAP= 40, CP=28) found more *A. actinomyecetemcomitans* and *P. gingivalis*. in GAP compared to CP.[43]

Asia

In Chinese populations, Li et al (GAP=10,CP=10) sequenced the 16S rDNA of bacteria in GAP patients and compared them to a first-degree relatives that have CP. GAP patients had lower taxonomic diversity than patients with CP. The ten most abundant shared species between the two conditions were *Porphyromonas gingivalis, Porphyromonas endodontalis, Tannerella forsythia, Treponema medium, Leptotrichia hofstadii, Treponema denticola, Prevotella intermedia*, and *Prevotella loescheii*, representing only 44.12% and 19.35% of the total constituents of the microbiome in CP and GAP respectively. The unique highly abundant species in GAP were *Treponema denticola, Prevotella intermedia, Prevotella loescheii, Selenomonas infelix, and Prevotella tannerae*, each having an abundance of less than 1%. On the other hand, the

unique highly abundant species in CP were *Leptotrichia wadei* (2.82%), *Fusobacterium canifelinum* (2.48%), *Corynebacterium matruchotii* (1.16%), and *Prevotella scopos* (0.7%). They also identified a higher abundance of P. gingivalis in GAP (35.88%) compared to CP (11.26%). A. actinomyecetemcomitans was not present in any of the GAP samples. The constituents of the microbiome of those with aggressive periodontitis were compared to both others with the same condition, and to their relatives. Interestingly, GAP people were closer in their microbial profile to their first-degree relatives with chronic periodontitis than to others with the same diagnosis of aggressive periodontitis.[44]

In another Chinese study, Liu (GAP = 57, CP = 73;) quantified the levels of Tannerella fosythia, Porphyromonas gingivalis and Aggregatibacter actinomycetemcomitans using real-time PCR. Porphyromomas gingivalis was found in higher frequency and quantity in CP compared to GAP. Aggregatibacter actinomycetemcomitans was detected more in GAP but the difference was not statistically significant.[45]

In Japan, Tomita (GAP=20, CP=20) quantified Aggregatibacter actinomycetemcomitans, Porphyromonas gingivalis, Tannerella forsythia using real-time PCR. None of the target bacteria was detected in healthy individuals. Moreover, no significant difference was found in the prevalence or abundance of the 3 pathogens between the two diseased groups. That being said, Tannerella fosythia was found 4 folds higher in CP than in GAP.[46]

The microbiome is both individualized and universal

To understand the limitations of the previous studies, it would be helpful to examine these investigations in light of our current understanding of the microbiome associated with health and disease.

A major shift in our collective understanding of the microbiome occurred due to The Human Metagenome Project (HMP) which was a large multicenter study that investigated the microbiome of 242 healthy individuals across 18 body sites (supra- and subgingival plaque being among them). Both functional and organismal diversity were analyzed. It was found that at the human community level, each body site is usually inhabited by a few signature taxa, which makes the plurality of the inhabitants of that body site. These clades would constitute, on average, anywhere from 17% to 84% of the inhabitants of that body site. The most abundant taxa in supragingival plaque was found to be from the Actinomyces taxa, while the Prevotella dominates the subgingival plaque. Less dominant taxa were considered to be highly personalized, both at the body site level, and among different people.[47,48] One explanation to this diversity has been confirmed by ethnicity studies, which found that in the oral cavity, the bacterial profile of a person is highly individualized, with only 2% bacteria being present in all ethnicities. Nevertheless, certain clusters of bacteria were found to be highly specific to certain ethnicities (74-91% specificity).[37]

Functionally, it was found that across body parts, the pathways in the microbial community metagenome where much less variable than the organismal diversity. Nevertheless, each body part had specific and highly abundant general metabolic pathways that can characterize the healthy function of that body part. While a relatively

few genes were specifically present or absent in any one body habitat, over two thirds were differentially abundant in the body habitats. These pathways appear to be relatively stable over time.[47,48] The diversity of the microbial constituents along with the relative stability of the functional pathways could be due to the interchangeability of the microbial community members, each with the capability of providing the same functional pathway to the collective microbiome.[49]

The gradual disruption of the healthy profile of the microbiome is termed dysbiosis.[50] The current theory is that certain microorganisms in low abundance, "keystone species", are able to change the commensal microbial environment, and as such prepare the microbiome to be more virulent. Experimental evidence suggests that, in the oral cavity, P. gingivalis is a candidate keystone species.[51] Some explanation of the dysbiotic transformation of the microbiome has been recently published by Yost. In a prospective study, it was found that the stable sites in individuals with periodontitis have higher microbial activity than the sites in periodontally healthy individuals, even though they both were clinically similar in appearance. Moreover, it found that the members of the aforementioned yellow complex, *Streptococcus oralis, S. intermedius, S. mitis*, as well as *Veillonella parvula* and *Pseudomonas fluorenscens* were found to be actively transcribing virulence factors.[27] This suggests that the whole biofilm becomes virulent as opposed to only specific pathogens in it.

Based on the available evidence, we hypothesized that the microbiome associated with generalized aggressive periodontitis would differ significantly from both that associated with health and that of chronic periodontitis. We tested this hypothesis with the following specific aims:

1. Aim 1: To investigate whether generalized aggressive periodontitis is associated with a site-specific or global dysbiosis in the microbiome
2. Aim 2: To compare the functional diversity of the periodontal microbiome in subjects with chronic and aggressive periodontitis.

CHAPTER 2

Materials and methods

Study population and study design

The research protocol was a cross-sectional study. The recruitment of patients occurred in two different locations; The Ohio State University, and Louisiana State University, as two separate studies.

In the Ohio State University, patients that were receiving nonsurgical periodontal therapy as a part of their treatment in the Graduate Periodontology Clinics of The Ohio State University were selected from April 2014 to February 2016.

Inclusion criteria

Patients recruited in the research project had to be at least 8 years of age, be able and willing to provide consent for the study (or obtain assent from guardian), and be diagnosed with Generalized Aggressive Periodontitis according to the criteria of the 1999 American Academy of Periodontology workshop.

Data from twenty-five healthy never-smokers (attachment loss ≤ 1, probing pocket depths ≤ 3) and twenty-five never-smokers with generalized moderate to severe chronic periodontitis (attachment loss ≥ 5, probing pocket depths ≥ 5) were used as the comparison groups.

Exclusion criteria

Patients were excluded from the research project if they were pregnant, had taken antibiotics in the last 3 months, were medicated with immunosuppressant medications or bisphosphonates, or presented with signs of uncontrolled systemic disease or diabetes.

Informed consent was obtained from every participating patient before entering the study. The study protocol and the informed consent forms were approved by The Ohio State Institutional Review Board (2014H0020).

Study design

Patients were seen at the time of scaling and root planing and consent obtained prior to any clinical procedure. At that time, the study protocol was explained, the investigator determined the patient's eligibility, the patient was given the opportunity to ask questions, and written informed consent form was obtained. Demographic and medical information was obtained from the patient and the chart.

During the visit, an assessment of clinical indices was initially conducted, including gingival index (GI, Loe and Silness[52]), plaque index (PI, Silness and Loe[52]). After isolation with cotton rolls, supragingival plaque and calculus were carefully removed using curettes. Following this step, paper points were carefully inserted in each tooth sulcus of the pocket investigated. In all cases, paper points were left in the site for 15-30 seconds, and immediately placed in 100μL of RNAlater and temporarily stored in ice for the duration of the appointment. Special attention was given to keep the paper points isolated from any saliva, and for them to be transferred to the RNAlater media as

quickly as possible to avoid contamination. Each patient had 3 sites with deep probing depth (GAPD) and 3 sites with shallow probing depth (GAPH) sampled (PPD ≥ 3 mm).

After the visit, the samples were transferred from ice to a controlled freezer (-80℃) and left in this location until the next step.

DNA isolation

At this step, the paper points were removed from the freezer and placed in a new sterile 1.5 mL collection tube. 200 µL of PBS was added and the samples were agitated for 30 minutes to allow for a maximum of the DNA to transfer from the paper points to the liquid. After this step, the paper points were removed and placed in a small punctured collection tube, which was then placed into a 1.5 mL tube and centrifuged to separate the DNA from the paper points.

180 µL of ATL buffer and 40 µL of proteinase K were added to the eluent and vortexed for 15 seconds. The samples were incubated in a water bath at a temperature of 56℃ for a minimum of 2 hours. 200 µL of AL buffer was then added, the mix was agitated for 15 seconds and heated for 10 minutes in a dry bath at a temperature of 70℃. 200 µL of 100% ethanol was added to the mix, which was then agitated for 15 seconds.

The mix was then added to a QIAamp Spin Column and centrifuged for 1 minute. The filtrate was then discarded and 500 µL of AW1 washing solution was added to the spin column, centrifuged for 1 minute and the filtrate was again discarded. 500 µL of AW2 washing solution was then added to the same spin column, centrifuged for 3 minutes and the filtrate was discarded along with the 2.0mL collection tube.

The spin column was seated in a new 1.5 mL collection tube, 50 μL TE buffer with EDTA added to the filter incubated at room temperature for 5 minutes and centrifuged for 1 minute. The eluent was quantified in a NanoSpec ND 1000 Spectrophotometer and 100ng of DNA was used for whole genome shotgun sequences.

Illumina Whole-Genome Sequencing

With the help of the Illumina MiSeq paired-end platform (Illumina Inc., San Diego, California) located in a commercial facility (Molecular Research LP, Shallowater, Texas), Multiplex bacterial tag-encoded Whole Genome Shotgun Sequencing was performed. An Nextera DNA Sample Preparation Kit (Illumina) was then used according to manufacturer's instructions to complete library generation. Genomic DNA was then sheared enzymatically to obtain an average fragment size of 500 base pairs. Pooled libraries (12pM) were loaded to a 600 Cycles v3 Reagent cartridge (Illumina) and the sequencing was performed on Miseq (illumina).

Metagenomic analysis

The sequences were trimmed and filtered, and submitted electronically to the MG-RAST pipeline for quality processing and functional analysis. Sequences were functionally assigned by the MG-RAST server (Argonne National labs). Functional potential was compared between groups using KEGG (Kyoto Encyclopedia of Genes and Genomes) and SEED hierarchies.

Statistical analysis

The statistical analysis for functional potential comparisons were conducted using the Wald-test for inter- and intra-species diversity, and adjustment of the p-values through FDR (False Discovery Rates) methods using the statistical package DESeq2 within R.[53]

CHAPTER 3

Results

Clinical data

From April 2014 to February 2016, 17 patients completed the study at The Ohio State University and Louisiana State University. The demographic data for these 17 patients, as well as the demographic data of 25 samples of chronic periodontitis (CP) can be found in appendix I.

GAP - Core Metagenome in Diseased Sites

A total of 8998 genes were identified at sites with deep periodontal pocket depths in subjects with GAP. 57.3% (5159 genes) can be considered as part of the core metagenome of this disease, based on their presence in 80% of the samples (appendix B).

The largest category of the core metagenome was related to metabolism of carbohydrates, which occupies 14.4% of genes (748 genes). Within this category, utilization and synthesis of mono-, di-, oligo-, and polysaccharides constituted 32% of the carbohydrate metabolism genes, while 23.1% of the genes were related to central carbohydrate metabolism pathways such as the TCA cycle, pyruvate metabolism, and glycolysis and gluconeogensis. Utilization of sugar alcohols such as ethanolamine,

glycerols, and mannitol, and fermentation of sugars to lactate, acetone, and ethanol occupied 17.9% of this category.

Utilization of amino acids and proteins contributed to 14.2% of the core metagenome. 41.2% of these genes were related to protein biosynthesis, degradation, processing, folding, and modification, while the utilization, synthesis and degradation of the proteinogenic amino acids lysine, threonine, methionine, cysteine, alanine, serine, and glycine constituted 20% of the metagenome of that category.

The third most abundant category of genes (12.2%) are clusters with hypothesized functions based functional coupling, but whose functions are not known (clustering-based subsystems).

GAP vs Health

9621 genes were identified in healthy patients and deep sites of patients with GAP. Of those genes, 4062 genes were differentially abundant in either condition (Wald test, $p>0.05$). After p-value adjustment for multiple comparison testing, 2027 were significantly more abundant in GAP, compared to 1519 genes in healthy patients. This suggests qualitative functional diversity in GAP compared to health (figure 1-2). 1388 (68.47%) genes were in the core metagenome of GAP.

When the genes that are significantly abundant in the two states are compared (figure 3), the majority of them were related to the central metabolic functions in the microbiome such as carbohydrate, protein and amino acid metabolism. Within this group, the largest contributor in both conditions were related to carbohydrates, which contribute 15.1% and 14.6% to the total metagenome in GAP and in healthy sites, respectively, with

the genes related to the central carbohydrate metabolism and its auxiliary pathways being the most significant contributors (figure 4). The total functional capability for carbohydrate utilization is depleted compared to health (268 GAP, 305 in health), especially in regards to utilization of di and oligosaccharides. The capability for fermentation, and metabolism of one-carbon such the case of methanogenesis, however, is enriched in GAP. The second largest contributor to the central metabolic functions is protein metabolism, which contribute to 14.7% and 16% in GAP and healthy, respectively.

Of the genes that are related to the virulence lifestyle, some genes in GAP appear to have an overall enrichment of both quality and quantity. For example: membrane transport (figure 5), a more diverse genetic makeup exists in GAP, in terms of quality (healthy: membrane secretion type II, IV, ABC transporters. GAP: membrane secretion type II, VI, VIII, ABC transporter, protein translocation across cytoplasmic membrane). Likewise, with regards to motility, there are 63 differentially enriched genes were found in GAP that were related to flagella, all of which are from 0.9-5.8 logfold change, compared to 1 gene in health, which is concerned with bacterial chemotaxis. Iron acquisition and metabolism, surprisingly, is more diverse in health compared to diseased sites (54 unique genes vs 46, abundance higher in health than in disease). Finally, in some categories, such as in phages, there is a distinct shift in function between the two conditions (figure 6), such as the enrichment of transposable elements in GAP compared to health (15 vs 0; in diseased sites; the majority of which are related to conjugative transposon in the order *Bacteroidales*) while 35 unique genes were found in health with regards to phages, prophages, compared to 15 in the diseased sites.

GAPD- GAPH - Global disease

10183 genes have been identified in patients with generalized aggressive periodontitis in both shallow and deep sites. When the shallow and deep sites were compared, 1725 genes were differentially abundant in either condition (Wald test, $p>0.05$), which represents %28.72 of the genes. When the p-value was adjusted for multiple comparison analysis, 1 gene only, belonging to the clustering based subsystem, reached statistical significance. As such, the microbiome in the deep and shallow sites is considered as functionally identical (figure 7).

GAP vs CP

A total of 9741 genes were identified in sites with disease in subjects with GAP and CP. Of those genes, 1982 were differentially abundant in either condition (Wald test, $p>0.05$), which represents 20.3% of the genes. After p-value adjustment for multiple comparison testing, 333 were significantly more abundant in GAP, compared to 761 genes in CP patients. (figures 9-10). 247 of those genes are part of the core metagenome of GAP, representing 74.1% of the identified abundant genes in this disease.

Of the genes that are significantly abundant in either condition (figure 11), the majority of them were related to the central metabolic functions in the microbiome such as carbohydrate, protein and amino acid metabolism. The largest contributor to the differences in their metagenome in both conditions is, again, the processes related to the metabolic potential for the two disease states. Genes related to carbohydrates are slightly

more depleted in GAP, while protein metabolism is slightly more enriched in GAP compared to CP.

Of the genes that are related to virulence, CP had a more diverse genetic capability than GAP. The production capability of cell wall and capsule components in CP is more diverse than in GAP (CP 77 vs. GAP 29); more gram positive and negative cell wall components, and capsular and extracellular polysaccharides are found in CP than in GAP. Moreover, membrane transport genes (35 in CP, 8 in GAP) are more abundant in CP, especially those related ECF class transporters, and ATP-dependent efflux . In relation to phages and transposable elements (39 in CP, 7 in GAP), GAP can be characterized by its enrichment of genes related to integrons, and the aforementioned conjugative transposons in Bacteriodales. In CP, phage capsid proteins, entry and exit, and packaging are more abundant. The capability of GAP to resist stresses is depleted compared to CP (GAP - 10 genes, CP - 13 genes). Capability of GAP to resist oxidative, osmotic, acid, and heat/cold challenges is depleted compared to CP. However, periplasmic stress response is enhanced in GAP compared to CP. The resistance to antibiotics and toxic compounds is more diverse in CP than in GAP. CP has the capability to produce more proteins related to zinc resistance, and fluoroquinolones, and beta- lactamase, while GAP is more capable to produce genes related to methicillin resistance, and multidrug resistance efflux pumps. Finally, the capability of the GAP metagenome for motility and chemotaxis is enriched compared to CP (42 in GAP, 2 in CP), with GAP being capable to produce factors related to the aforementioned bacterial chemotaxis, flagellar motility, flagellum.

CHAPTER 4

Discussion

The functional potential of the subgingival microbiome, and its shift from health and disease has been examined by several investigators recently.[24,54-56] These studies have focused on the functional differences in health and disease[54-56], and after treatment.[24] It is assumed from those studies that the diseased state is that of chronic periodontitis. To our knowledge, this is the first study that looks at the functional difference between generalized aggressive periodontitis and health, and as well as comparing the generalized aggressive form to the chronic form of periodontitis.

The shift from health states to diseased states is characterized by a change in the diversity in the genetic capability in the microbiome.[54] This shift results from a gain and loss of certain functions that give the diseased microbiome its virulence factors. The microbiome in chronic periodontitis has been described as a global disease in the mouth, with the metagenome in the shallow and the diseased sites being virtually indistinguishable, and that the disease state is due to the overexpression of certain genes in the metagenome.[27] This investigation reveals that generalized aggressive periodontitis is also a globalized condition in the mouth, where the shallow and deep sites are functionally identical. It remains to be examined whether the difference between the shallow and deep sites is also due to the heightened transcription rate in disease.

The transition of the subgingival microbiome from health to GAP results in a microbiome that begins to organize and isolate itself from the outside stressors, most likely toward tissue and intracellular environments. Intercellular invasion of the microbiome in GAP has been confirmed previously through histology[57], which undercuts the importance of the microbiome in GCF. Moreover, both A.a. and P.g. have the capability of intracellular invasion of buccal epithelial cells[58], and it remains to be seen whether such intracellular invasion is possible in the cells of the sulcular epithelium. Investigating the abundance of the membrane secretion pathways (II, V, VI) in GAP could describe the molecules by which the microbiome could invade those cells. The matching of these membrane transport system, as well as the other functional genes, with their species would shed light on what the true function of those virulence associated genes are. Interestingly, it has been found that fimbrial biogenesis can be classified as a membrane secretion system type II.[59]

As mentioned previously, the microbiome in GAP is also characterized by high capacity for motility, and intracellular invasion. In light of this, the capacity of the outer layer of GAP to resist outside stressors, and may be of lesser importance to the survival of the microbiome. This is evident during the comparison between GAP and CP, in which the capability of the disease to resist outside stressors and create a favorable environment in the GCF compartment is depleted compared to CP.

CHAPTER 5

Conclusion

The results of this study shows the distinct functional differences in the state of health and disease in generalized aggressive periodontitis. These pockets, despite looking clinically healthy, have the same functional capacity as the deeper pockets in those patients. This confirms that, just like the chronic variant, generalized aggressive periodontitis is a globalized condition in the mouth.

These results bring interesting findings, both in terms of our understanding of the pathogenesis of the disease, as well as shedding light on novel methods by which diagnosis, and prognosis of the diseases can be done.

REFERENCES

1. American Academy of Periodontology. Glossary of periodontal terms. 2001.

2. Eke PI, Dye BA, Wei L, Slade GD, Thornton-Evans GO, Borgnakke WS, et al. Update on Prevalence of Periodontitis in Adults in the United States: NHANES 2009 to 2012. JOP. 2015 May;86(5):611–22.

3. Armitage GC. Development of a classification system for periodontal diseases and conditions. Annals of Periodontology. 1999;4(1):1–6.

4. Sgolastra F, Gatto R, Petrucci A, Monaco A. Effectiveness of Systemic Amoxicillin/Metronidazole as Adjunctive Therapy to Scaling and Root Planing in the Treatment of Chronic Periodontitis: A Systematic Review and Meta-Analysis. JOP. 2012 Oct;83(10):1257–69.

5. Sgolastra F, Petrucci A, Gatto R, Monaco A. Effectiveness of Systemic Amoxicillin/Metronidazole as an Adjunctive Therapy to Full-Mouth Scaling and Root Planing in the Treatment of Aggressive Periodontitis: A Systematic Review and Meta-Analysis. JOP. 2012 Jun;83(6):731–43.

6. Karoussis IK, Kotsovilis S, Fourmousis I. A comprehensive and critical review of dental implant prognosis in periodontally compromised partially edentulous patients. Clin Oral Implants Res. 2007 Dec;18(6):669–79.

7. Lang N, Bartold PM, Cullinan M, Jeffcoat M, Mombelli A, Murakami S, et al. Consensus report: aggressive periodontitis. annals of periodontology / the American Academy of Periodontology. Am Acad Periodontology; 1999;4(1):53–3.

8. Löe H, Brown LJ. Early onset periodontitis in the United States of America. JOP. 1991 Oct;62(10):608–16.

9. Brown LJ, Albandar JM. Early-onset periodontitis: progression of attachment loss during 6 years. Journal of …. 1996.

10. Susin C, Haas AN, Albandar JM. Epidemiology and demographics of aggressive periodontitis. Periodontology 2000. 2014 Jun;65(1):27–45.

11. Saxby MS. Juvenile periodontitis: an epidemiological study in the west Midlands of the United Kingdom. Journal of Clinical Periodontology. Blackwell Publishing Ltd; 1987 Nov 1;14(10):594–8.

12. Armitage GC, Cullinan MP. Comparison of the clinical features of chronic and aggressive periodontitis. Periodontology 2000. 2010 Jun;53(1):12–27.

13. American Academy of Periodontology Task Force Report on the Update to the 1999 Classification of Periodontal Diseases and Conditions*. JOP. 2015 Jul;86(7):835–8.

14. Cigarette smoking and alveolar bone height in subjects with a high standard of oral hygiene. 1987 Sep;14(8):466–9.

15. Lang NP, Bartold PM, Cullinan M. Aggressive periodontitis.(Consensus report). Ann Periodontol; 1999.

16. Page RC, Schroeder HE. Pathogenesis of Inflammatory Periodontal Disease; a Summary of Current Work. 1976. 1 p.

17. Tatakis DN, Kumar PS. Etiology and pathogenesis of periodontal diseases. Dent Clin North Am. 2005 Jul;49(3):491–516–v.

18. Kumar PS, Mason MR. Mouthguards: does the indigenous microbiome play a role in maintaining oral health? Front Cell Infect Microbiol. 2015 May 6;5:5721–9.

19. Kumar PS, Leys EJ, Bryk JM, Martinez FJ, Moeschberger ML, Griffen AL. Changes in periodontal health status are associated with bacterial community shifts as assessed by quantitative 16S cloning and sequencing. Journal of Clinical Microbiology. 2006 Oct;44(10):3665–73.

20. Benítez-Páez A, Belda-Ferre P, Simón-Soro A, Mira A. Microbiota diversity and gene expression dynamics in human oral biofilms. BMC Genomics. 2014;15:311.

21. Paster BJ, Olsen I, Aas JA, Dewhirst FE. The breadth of bacterial diversity in the human periodontal pocket and other oral sites. Periodontology 2000. 2006;42:80–7.

22. Kittichotirat W, Bumgarner RE, Chen C. Evolutionary Divergence of Aggregatibacter actinomycetemcomitans. Journal of Dental Research. 2015 Dec 23;95(1):94–101.

23. 16S rDNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. 2007 Sep;45(9):2761–4.

24. Shi B, Chang M, Martin J, Mitreva M, Lux R, Klokkevold P, et al. Dynamic Changes in the Subgingival Microbiome and Their Potential for Diagnosis and Prognosis of Periodontitis. mBio. 2015 Feb 17;6(1):e01926–14–11.

25. Socransky SS, Haffajee AD, Cugini MA, Smith C, Kent RL. Microbial complexes in subgingival plaque. Journal of Clinical Periodontology. 1998 Feb;25(2):134–44.

26. Griffen AL, Beall CJ, Campbell JH, Firestone ND, Kumar PS, Yang ZK, et al. Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. ISME J. Nature Publishing Group; 2011 Dec 15;6(6):1176–85.

27. Yost S, Duran-Pinedo AE, Teles R, Krishnan K, Frias-Lopez J. Functional signatures of oral dysbiosis during periodontitis progression revealed by microbial metatranscriptome analysis. Genome Med. 2015 Apr 27;7(1):1659–20.

28. Teles R, Teles F, Frias-Lopez J, Paster B, Haffajee A. Lessons learned and unlearned in periodontal microbiology. Periodontology 2000. 2013 Jun;62(1):95–162.

29. Shaddox LM, Huang H, Lin T, Hou W, Harrison PL, Avila G, et al. Microbiological characterization in children with aggressive periodontitis. Journal of Dental Research. 2012 Oct;91(10):927–33.

30. Tonetti MS, Mombelli A. Early-onset periodontitis. Annals of Periodontology. 1999.

31. Takeuchi Y, Umeda M, Ishizuka M, Huang Y, Ishikawa I. Prevalence of periodontopathic bacteria in aggressive periodontitis patients in a Japanese population. JOP. Am Acad Periodontology; 2003;74(10):1460–9.

32. Mayorga-Fayad I, Lafaurie GI, Contreras A, Castillo DM. Microflora subgingival en periodontitis crónica y agresiva en Bogotá, Colombia: un acercamiento epidemiológico. Biomédica. 2007.

33. Nguyen-Hieu T. Microbial sampling process can change results of microbiological analysis in periodontitis diagnosis. A minireview. J Invest Clin Dent. 2012 Nov 27;4(3):151–9.

34. Haffajee AD, Socransky SS. Effect of sampling strategy on the false-negative rate for detection of selected subgingival species. Oral Microbiol Immunol. 1992 Feb;7(1):57–9.

35. Laksmana T, Kittichotirat W, Huang Y, Chen W, Jorgensen M, Bumgarner R, et al. Metagenomic analysis of subgingival microbiota following non-surgical periodontal therapy: a pilot study. Open Dent J. 2012;6:255–61.

36. Faveri M, Mayer MPA, Feres M, de Figueiredo LC, Dewhirst FE, Paster BJ. Microbiological diversity of generalized aggressive periodontitis by 16S rDNA clonal analysis. Oral Microbiol Immunol. 2008 Apr;23(2):112–8.

37. Mason MR, Nagaraja HN, Camerlengo T, Joshi V, Kumar PS. Deep Sequencing Identifies Ethnicity-Specific Bacterial Signatures in the Oral Microbiome. Biswas I, editor. PLoS ONE. 2013 Oct 23;8(10):e77287–7.

38. Riep B, Edesi-Neuss L, Claessen F, Skarabis H, Ehmke B, Flemmig TF, et al. Are putative periodontal pathogens reliable diagnostic markers? Journal of Clinical Microbiology. American Society for Microbiology; 2009 Jun;47(6):1705–11.

39. Nibali L, D'Aiuto F, Ready D, Parkar M, Yahaya R, Donos N. No association between A actinomycetemcomitans or P gingivalis and chronic or aggressive periodontitis diagnosis. Quintessence Int. 2012 Mar;43(3):247–54.

40. Rescala B, Rosalem W Jr., Teles RP, Fischer RG, Haffajee AD, Socransky SS, et al. Immunologic and Microbiologic Profiles of Chronic and Aggressive Periodontitis Subjects. JOP. 2010 Sep;81(9):1308–16.

41. Lourenço TGB, Heller D, Silva-Boghossian CM, Cotton SL, Paster BJ, Colombo APV. Microbial signature profiles of periodontally healthy and diseased patients. Journal of Clinical Periodontology. 2014 Sep 22;41(11):1027–36.

42. Gajardo M, Silva N, Gómez L, León R, Parra B, Contreras A, et al. Prevalence of periodontopathic bacteria in aggressive periodontitis patients in a Chilean population. JOP. 2005 Feb;76(2):289–94.

43. Casarin RCV, Del Peloso Ribeiro É, Mariano FS, Nociti FH Jr, Casati MZ, Gonçalves RB. Levels of Aggregatibacter actinomycetemcomitans, Porphyromonas gingivalis, inflammatory cytokines and species-specific immunoglobulin G in generalized aggressive and chronic periodontitis. J Periodont Res. 2010 Aug 20;45(5):635–42.

44. Li Y, Feng X, Xu L, Zhang L, Lu R, Shi D, et al. Oral microbiome in chinese patients with aggressive periodontitis and their family members. Journal of Clinical Periodontology. 2015 Nov 14;42(11):1015–23.

45. Liu J, Zhao J, Li C, Yu N, Zhang D, Pan Y. Clinical and microbiologic effect of nonsurgical periodontal therapy on patients with chronic or aggressive periodontitis. Quintessence Int. 2013;44(8):575–83.

46. Tomita S, Komiya-Ito A, Imamura K, Kita D, Ota K, Takayama S, et al. Prevalence of Aggregatibacter actinomycetemcomitans, Porphyromonas gingivalis and Tannerella forsythia in Japanese patients with generalized chronic and aggressive periodontitis. Microbial Pathogenesis. Elsevier Ltd; 2013 Aug 9;61-62(C):11–5.

47. Consortium THMP. Structure, function and diversity of the healthy human microbiome. Nature. Nature Publishing Group; 2012 Jun 4;486(7402):207–14.

48. Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, et al. Metabolic Reconstruction for Metagenomic Data and Its Application to the Human Microbiome. Eisen JA, editor. PLoS Comp Biol. 2012 Jun 13;8(6):e1002358–17.

49. Jorth P, Turner KH, Gumus P, Nizam N, Buduneli N, Whiteley M.

Metatranscriptomics of the Human Oral Microbiome during Health and Disease. mBio. 2014 Feb 25;5(2):e01012–14–e01012–14.

50. Berezow AB, Darveau RP. Microbial shift and periodontitis. Periodontology 2000. 2010 Dec 7;55(1):36–47.

51. Hajishengallis G, Liang S, Payne MA, Hashim A, Jotwani R, Eskan MA, et al. Low-Abundance Biofilm Species Orchestrates Inflammatory Periodontal Disease through the Commensal Microbiota and Complement. Elsevier Inc; 2011 Nov 17;10(5):497–506.

52. Löe H. The gingival index, the plaque index and the retention index systems. JOP. 1967.

53. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014 Dec 5;15(12):31–21.

54. Li Y, He J, He Z, Zhou Y, Yuan M, Xu X, et al. Phylogenetic and functional gene structure shifts of the oral microbiomes in periodontitis patients. ISME J. Nature Publishing Group; 2014 Mar 27;8(9):1879–91.

55. Wang J, Qi J, Zhao H, He S, Zhang Y, Wei S, et al. Metagenomic sequencing reveals microbiota and its functional potential associated with periodontal disease. Sci Rep. 2013 May 15;3:1–10.

56. Duran-Pinedo AE, Chen T, Teles R, Starr JR, Wang X, Krishnan K, et al. Community-wide transcriptome of the oral microbiome in subjects with and without periodontitis. ISME J. 2014 Mar 6;8(8):1659–72.

57. Christersson LA, Albini B, Zambon JJ, Wikesjö UM, Genco RJ. Tissue localization of Actinobacillus actinomycetemcomitans in human periodontitis. I. Light, immunofluorescence and electron microscopic studies. JOP. 1987 Aug;58(8):529–39.

58. Rudney JD, Chen R, Sedgewick GJ. Intracellular Actinobacillus actinomycetemcomitans and Porphyromonas gingivalis in buccal epithelial cells collected from human subjects. Infection and Immunity. 2001 Apr;69(4):2700–7.

59. Henderson IR, Navarro-Garcia F, Desvaux M, Fernandez RC, Ala'Aldeen D. Type V Protein Secretion Pathway: the Autotransporter Story. Microbiology and Molecular Biology Reviews. 2004 Dec 8;68(4):692–744.

Appendix A
Demographic Data

Table 1 - Demographic data of the sampled patients

	GAP	CP	Health
Age	22.93 ± 7.5	58 ± 2	56 ± 3
%Male	52	72	89
PD	6.9 ± 1.1	6.0 ± 1.5	3.2 ± 1.7
%BOP	100	100	10 ± 4

Appendix B

Core Metagenome in diseased sites of GAP

Table 2: Core metagenome in GAP

Level 1	Total	Level 2	Total
Amino Acids and Derivatives	431	Lysine, threonine, methionine, and cysteine	101
		Arginine; urea cycle, polyamines	77
		Aromatic amino acids and derivatives	70
		Branched-chain amino acids	66
		Alanine, serine, and glycine	46
		Glutamine, glutamate, aspartate, asparagine; ammonia assimilation	31
		Histidine Metabolism	19
		Proline and 4-hydroxyproline	16
		Other amino acids and derivatives	5
Carbohydrates	748	Central carbohydrate metabolism	173
		Monosaccharides	126
		Di- and oligosaccharides	92
		Other carbohydrates	73
		Sugar alcohols	68
		Fermentation	66
		One-carbon Metabolism	39
		CO2 fixation	32
		Organic acids	31
		Polysaccharides	22
		Aminosugars	19
		Glycoside hydrolases	7
Cell Division and Cell Cycle	46	Other cell division and cell cycle	46
Cell Wall and Capsule	268	Capsular and extracellular polysacchrides	104
		Gram-Negative cell wall components	73
		Other cell division and cell cycle	46
		Gram-Positive cell wall components	23
		Cell wall of Mycobacteria	6
Cofactors, Vitamins, Prosthetic	359	Folate and pterines	148
		Tetrapyrroles	70
		NAD and NADP	35

continued

Table 2 continued

Groups, Pigments		Other cofactors, vitamins, prosthetic groups	24
		Quinone cofactors	23
		Riboflavin, FMN, FAD	17
		Biotin	14
		Coenzyme A	13
		Pyridoxine	12
		Lipoic acid	3
DNA Metabolism	177	DNA repair	80
		DNA replication	39
		Other DNA metabolism	20
		DNA uptake, competence	19
		CRISPs	15
		DNA recombination	4
Dormancy and Sporulation	18	Other dormancy and sporulation	18
Fatty Acids, Lipids, and Isoprenoids	144	Fatty acids	56
		Isoprenoids	46
		Phospholipids	25
		Other fatty acids, lipids, and isoprenoids	12
		Triacylglycerols	5
Iron acquisition and metabolism	107	Other iron acquisition and metabolism	97
		Siderophores	10
Membrane Transport	185	Other membrane transport	80
		ABC transporters	29
		Protein translocation across cytoplasmic membrane	18
		Protein and nucleoprotein secretion system, Type IV	14
		Protein secretion system, Type II	11
		Sugar Phosphotransferase Systems, PTS	10
		Uni- Sym- and Antiporters	9
		Protein secretion system, Type VI	8
		Protein secretion system, Type V	3
		Protein secretion system, Type I	1
		Protein secretion system, Type III	1
		Protein secretion system, Type VIII (Extracellular nucleation/precipitation pathway, ENP)	1
Metabolism of Aromatic Compounds	33	Peripheral pathways for catabolism of aromatic compounds	13
		Metabolism of central aromatic intermediates	11
		Anaerobic degradation of aromatic compounds	6
		Other metabolism of aromatic compounds	3

continued

Table 2 continued

Motility and Chemotaxis	85	Flagellar motility in Prokaryota	69
		Other motility and chemotaxis	16
Nitrogen Metabolism	45	Other nitrogen metabolism	45
Nucleosides and Nucleotides	134	Purines	56
		Pyrimidines	42
		Other nucleosides and nucleotides	22
		Detoxification	14
Phages, Prophages, Transposable elements, Plasmids	90	Phages, Prophages	55
		Transposable elements	23
		Pathogenicity islands	7
		Plasmid related functions	3
		-	2
Phosphorus Metabolism	44	Other phosphorus metabolism	44
Potassium metabolism	22	Other Potassium metabolism	22
Protein Metabolism	303	Protein biosynthesis	161
		Protein degradation	57
		Protein processing and modification	47
		Protein folding	23
		Selenoproteins	15
Regulation and Cell signaling	97	Other regulation and cell signaling	54
		Programmed Cell Death and Toxin-antitoxin Systems	23
		Quorum sensing and biofilm formation	10
		Regulation of virulence	10
Respiration	164	Electron donating reactions	61
		Other respiration	38
		Electron accepting reactions	36
		ATP synthases	18
		Sodium Ion-Coupled Energetics	11
RNA Metabolism	274	RNA processing and modification	248
		Transcription	25
		Other RNA metabolism	1
Secondary Metabolism	10	Plant Hormones	4
		Biologically active compounds in metazoan cell defence and differentiation	2
		Aromatic amino acids and derivatives	1
		Bacterial cytostatics, differentiation factors and antibiotics	1
		Lipid-derived mediators	1

continued

Table 2 continued

		Plant Alkaloids	1
Stress Response	142	Oxidative stress	5
		Heat shock	22
		Osmotic stress	20
		Other stress response	16
		Acid stress	5
		Detoxification	14
		Periplasmic Stress	5
		Cold shock	2
Sulfur Metabolism	46	Organic sulfur assimilation	21
		Other sulfur metabolism	16
		Inorganic sulfur assimilation	9
Virulence, Disease and Defense	143	Resistance to antibiotics and toxic compounds	77
		Other virulence, disease, and defense	39
		Adhesion	9
		Bacteriocins, ribosomally synthesized antibacterial peptides	6
		Detection	6
		Invasion and intracellular resistance	5
		Toxins and superantigens	1
Clustering-based subsystems	633	Other clustering-based subsystems	372
		Clustering-based subsystems	22
		Fatty acid metabolic cluster	21
		Biosynthesis of galactoglycans and related lipopolysacharides	13
		Cell Division	13
		Cytochrome biogenesis	11
		Probably GTP or GMP signaling related	9
		Protein export?	9
		Two related proteases	9
		Carbohydrates	8
		Ribosomal Protein L28P relates to a set of uncharacterized proteins	8
		Translation	8
		Isoprenoid/cell wall biosynthesis: PREDICTED UNDECAPRENYL DIPHOSPHATE PHOSPHATASE	7
		Monosaccharides	126
		Lysine, threonine, methionine, and cysteine	101
		proteosome related	6
		Ribosome-related cluster	6

continued

Table 2 continued

		TldD cluster	6
		Methylamine utilization	5
		Nucleotidyl-phosphate metabolic cluster	5
		Oxidative stress	5
		Putative asociate of RNA polymerase sigma-54 factor rpoN	5
		Tricarboxylate transporter	5
		CRISPRs and associated hypotheticals	4
		DNA polymerase III epsilon cluster	4
		Hypothetical in Lysine biosynthetic cluster	4
		Hypothetical lipase related to Phosphatidate metabolism	4
		Hypothetical protein possible functionally linked with Alanyl-tRNA synthetase	4
		Carotenoid biosynthesis	3
		Phosphate metabolism	3
		recX and regulatory cluster	3
		Sulfatases and sulfatase modifying factor 1 (and a hypothetical)	3
		Three hypotheticals linked to lipoprotein biosynthesis	3
		tRNA sulfuration	3
		Chemotaxis, response regulators	2
		D-tyrosyl-tRNA(Tyr) deacylase (EC 3.1.-.-) cluster	2
		Hypothetical associated with RecF	2
		Probably organic hydroperoxide resistance related hypothetical protein	2
		Probably Ybbk-related hypothetical membrane proteins	2
		Putative GGDEF domain protein related to agglutinin secretion	2
		Putrescine/GABA utilization cluster-temporal,to add to SSs	2
		Pyruvate kinase associated cluster	2
		Recombination related cluster	2
		Shiga toxin cluster	2
		Choline bitartrate degradation, putative	1
		Chromosome Replication	1
		DNA metabolism	1
		Flagella protein?	1
		Hypothetical Related to Dihydroorate Dehydrogenase	1
		Lysine Biosynthesis	1
		Molybdopterin oxidoreductase	1
		Probably Pyrimidine biosynthesis-related	1

continued

Table 2 continued

		Sarcosine oxidase	1
Miscellaneous	409	Plant-Prokaryote DOE project	376
		Other miscellaneous	33
Photosynthesis	2	Other photosynthesis	2

Appendix C
Comparison of Deep Sites in GAP with Healthy Patients

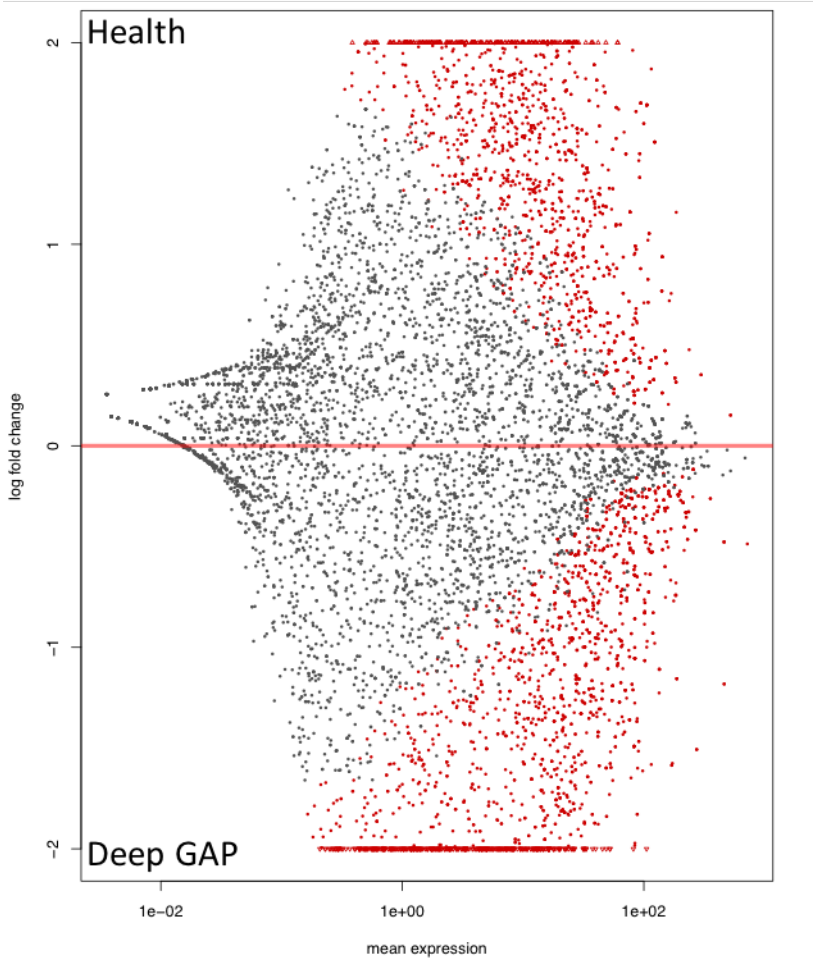

Figure 1 – logfold change in genes between deep sites (GAP) and health. Red dots represent significantly abundant genes (P≥0.05; Wald-test, FDR adjustment)

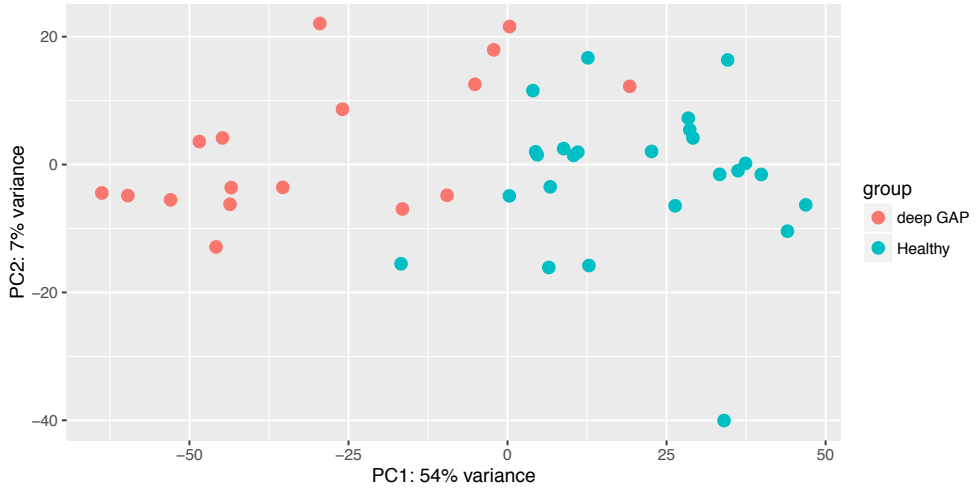

Figure 2 – PCA comparison between GAP and healthy patients.

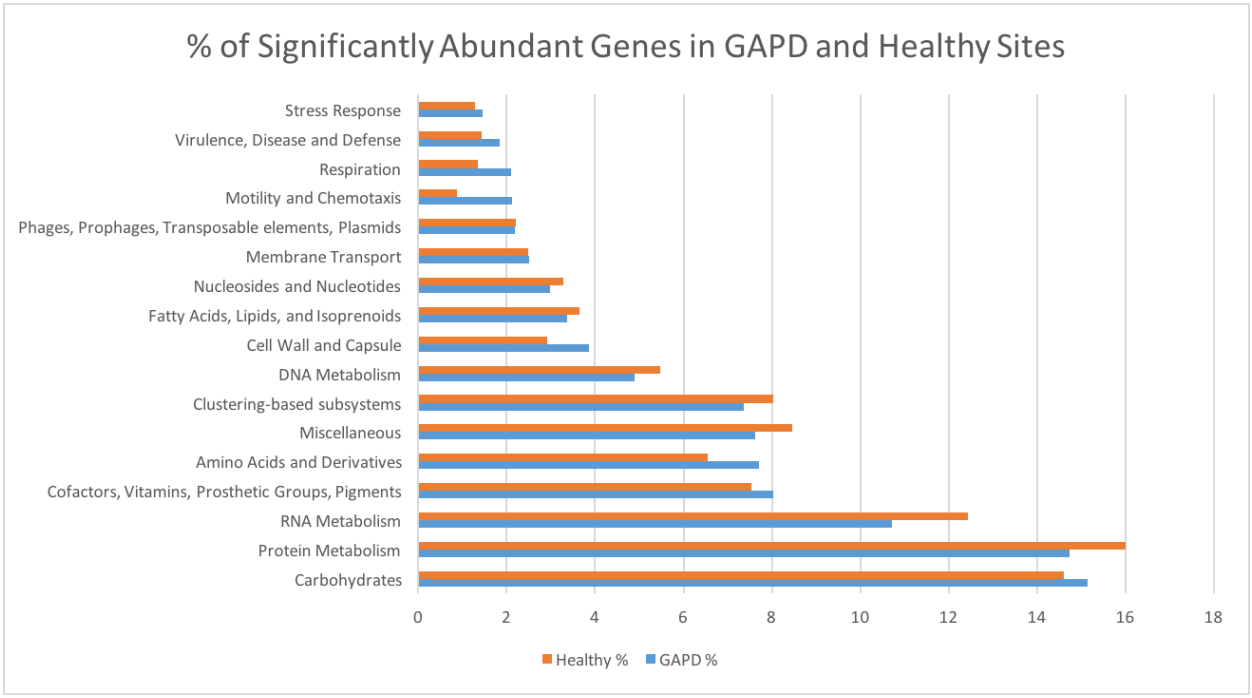

Figure 3 – comparison of GAP and healthy patients for the statistically significant genes as a fraction of their total. Only genes occupying 1% or more shown here

Appendix D
GAP and Health

Table 3: Composition of statistically significant genes in GAP and Health as a percentage of their whole

level1	level2	GAP AVG	GAP STDEV	Health AVG	Health STDEV
Carbohydrates	Central carbohydrate metabolism	0.044844058	0.003970112	0.042508453	0.004375016
	One-carbon Metabolism	0.024113997	0.001603548	0.026006976	0.003891372
	Fermentation	0.019988291	0.002224415	0.01489299	0.002974447
	Monosaccharides	0.014385848	0.00193677	0.012600605	0.00226214
	Di- and oligosaccharides	0.014192072	0.002058469	0.016853442	0.002858538
	CO2 fixation	0.007810726	0.000728702	0.008016848	0.001151148
	Organic acids	0.007634439	0.000755776	0.006406819	0.001348671
	Polysaccharides	0.00609234	0.000932249	0.006406176	0.000934815
	Sugar alcohols	0.005872963	0.001005617	0.006381105	0.001122468
	Aminosugars	0.005061295	0.000927526	0.005417736	0.001744858
	Glycoside hydrolases	0.001481311	0.000340409	0.000647652	0.000318513
Protein Metabolism	Protein biosynthesis	0.085312912	0.005327484	0.10457927	0.010107436
	Protein degradation	0.030011005	0.00166249	0.025073996	0.002690724
	Protein processing and modification	0.015550648	0.000703323	0.015996113	0.002432144
	Protein folding	0.008317158	0.000579213	0.007742746	0.001606086
	Selenoproteins	0.00813191	0.001424451	0.006728577	0.00160069
RNA Metabolism	RNA processing and modification	0.097489905	0.008195182	0.112901851	0.009901477
	Transcription	0.009766972	0.000594068	0.011504962	0.001419734
Cofactors, Vitamins, Prosthetic Groups, Pigments	Folate and pterines	0.026390484	0.000924003	0.027608376	0.001813022
	Tetrapyrroles	0.024234636	0.005131993	0.018390298	0.003150241
	Riboflavin, FMN, FAD	0.007997049	0.001121664	0.007887121	0.001343256
	NAD and NADP	0.007576357	0.000779844	0.007262464	0.001046334
	Biotin	0.005495757	0.00100596	0.005517486	0.001762248
	Coenzyme A	0.004911941	0.000408782	0.005412905	0.000877175
	Pyridoxine	0.003462805	0.000390119	0.00312757	0.000898651
	Quinone cofactors	0.000279345	0.000134569	0.000194049	0.000128981
Amino Acids and Derivatives	Lysine, threonine, methionine, and cysteine	0.02987797	0.003406327	0.025664234	0.003571661
	Glutamine, glutamate, aspartate, asparagine;	0.01197363	0.001255435	0.012825373	0.002664051

continued

Table 3 continued

	ammonia assimilation				
	Aromatic amino acids and derivatives	0.009194873	0.000753106	0.006795463	0.001595475
	Alanine, serine, and glycine	0.00895808	0.001533849	0.00851892	0.001750576
	Branched-chain amino acids	0.007503421	0.001759952	0.004457933	0.001850326
	Arginine; urea cycle, polyamines	0.004344551	0.000675972	0.00357856	0.000990057
	Histidine Metabolism	0.004188339	0.001210911	0.00260821	0.001013093
	Proline and 4-hydroxyproline	0.001022775	0.00017202	0.00097639	0.000518605
Miscellaneous	Plant-Prokaryote DOE project	0.076125463	0.002708525	0.084677841	0.006324063
Clustering-based subsystems	Fatty acid metabolic cluster	0.010071492	0.001510115	0.012666959	0.001139291
	Cell Division	0.010029231	0.000898716	0.013249607	0.001511273
	Clustering-based subsystems	0.004576949	0.00066912	0.005054034	0.001046436
	Probably GTP or GMP signaling related	0.003924633	0.000698197	0.004478691	0.001219779
	Isoprenoid/cell wall biosynthesis: PREDICTED UNDECAPRENYL DIPHOSPHATE PHOSPHATASE	0.003577738	0.000412499	0.003014254	0.000825664
	Hypothetical protein possible functionally linked with Alanyl-tRNA synthetase	0.003314987	0.000842579	0.00438759	0.000818764
	Ribosomal Protein L28P relates to a set of uncharacterized proteins	0.003147785	0.000355045	0.004068064	0.000871713
	Carbohydrates	0.002700248	0.000627284	0.003828212	0.001229157
	Lysine, threonine, methionine, and cysteine	0.002568641	0.000274656	0.003708942	0.00057511
	Shiga toxin cluster	0.002300682	0.000463371	0.002862629	0.000940578
	Translation	0.002220197	0.000254432	0.002471574	0.000763966
	Oxidative stress	0.002015469	0.00077505	0.000372676	0.000394178
	Biosynthesis of galactoglycans and related lipopolysacharides	0.001959703	0.000367025	0.001588339	0.000663688
	Recombination related cluster	0.001912799	0.000460107	0.001918503	0.001215986
	Nucleotidyl-phosphate metabolic cluster	0.001864671	0.000377591	0.001439033	0.000434377
	Tricarboxylate transporter	0.001733824	0.000527181	0.001131828	0.000799455
	Putative asociate of RNA polymerase sigma-54 factor rpoN	0.00170794	0.000731023	0.001065438	0.000463151
	Protein export?	0.001700011	0.000263534	0.001541607	0.000472263

continued

Table 3 continued

	Ribosome-related cluster	0.001632383	0.000363493	0.001329125	0.000667358
	Chromosome Replication	0.001337328	0.00031042	0.001848361	0.0004864
	TldD cluster	0.001189969	0.000232884	0.000912841	0.000423999
	tRNA sulfuration	0.000949697	0.000273279	0.001308602	0.000569001
	Three hypotheticals linked to lipoprotein biosynthesis	0.000836897	0.00016622	0.001133671	0.000200347
	CRISPRs and associated hypotheticals	0.000827336	0.000247043	0.000970828	0.000296438
	Carotenoid biosynthesis	0.000814801	0.000211484	0.00096563	0.000533166
	DNA polymerase III epsilon cluster	0.000685443	0.000231996	0.000323302	0.000257776
	D-tyrosyl-tRNA(Tyr) deacylase (EC 3.1.-.-) cluster	0.000613233	0.000215801	0.000625714	0.000346876
	Two related proteases	0.000611313	0.000126898	0.0003256	0.000220623
	Cytochrome biogenesis	0.000590957	0.000231882	0.000541844	0.000372504
	Monosaccharides	0.000467006	0.000112583	0.000249693	0.000183659
	Flagella protein?	0.00032135	0.000179712	0.000108139	0.000148189
	Sulfatases and sulfatase modifying factor 1 (and a hypothetical)	0.000254115	0.000138205	0.000187272	0.000191242
	Methylamine utilization	0.000179544	0.000110761	0.000127005	8.8942E-05
	Sarcosine oxidase	0.000179544	0.000110761	0.000127005	8.8942E-05
	recX and regulatory cluster	0.000175153	0.00013138	0.000127885	0.000138335
	Pyruvate kinase associated cluster	0.000166628	9.47136E-05	0.000115484	0.000117377
	Probably Ybbk-related hypothetical membrane proteins	0.000150913	5.95236E-05	1.85419E-05	4.05467E-05
	Hypothetical in Lysine biosynthetic cluster	0.00014612	9.82894E-05	0.000130393	0.000115296
	Hypothetical lipase related to Phosphatidate metabolism	5.10727E-05	4.42615E-05	2.75175E-05	4.34795E-05
	Chemotaxis, response regulators	3.42427E-05	2.33239E-05	2.56606E-05	4.48589E-05
	Putative GGDEF domain protein related to agglutinin secretion	2.87974E-05	2.2213E-05	1.50873E-05	3.87734E-05
DNA Metabolism	DNA repair	0.026979259	0.001660073	0.029487881	0.004155581
	DNA replication	0.018505303	0.001667179	0.021986398	0.002239318
	CRISPs	0.001786622	0.000372756	0.001763933	0.000478108
	DNA uptake, competence	0.001177692	0.000213462	0.00073876	0.000330383
	DNA recombination	0.000526265	9.75123E-05	0.000657016	0.00033193

continued

Table 3 continued

Cell Wall and Capsule	Gram-Negative cell wall components	0.021624274	0.003235724	0.014704217	0.005657932
	Capsular and extracellular polysacchrides	0.010611627	0.001161904	0.007476728	0.002035542
	Cell wall of Mycobacteria	0.004103474	0.000607401	0.005194657	0.000689289
	Gram-Positive cell wall components	0.002259662	0.00032047	0.001680193	0.000534938
Fatty Acids, Lipids, and Isoprenoids	Fatty acids	0.015989444	0.002003559	0.018446111	0.003544269
	Isoprenoids	0.014747781	0.001798662	0.015024886	0.001930453
	Phospholipids	0.003030232	0.000400685	0.003031151	0.000727639
Nucleosides and Nucleotides	Purines	0.015676226	0.001373682	0.018863873	0.003502888
	Pyrimidines	0.011598502	0.000866251	0.011300277	0.002343067
	Detoxification	0.002446628	0.000243222	0.002645379	0.000565687
Membrane Transport	ABC transporters	0.009573092	0.001904831	0.008297966	0.001486999
	Protein translocation across cytoplasmic membrane	0.006745742	0.000647912	0.007602533	0.001270975
	Protein secretion system, Type VI	0.004133769	0.00073233	0.005104805	0.001510207
	Protein secretion system, Type V	0.001975117	0.001253362	0.001577343	0.001082541
	Protein secretion system, Type II	0.001399741	0.000257342	0.001422142	0.000460472
	Uni- Sym- and Antiporters	0.001018326	0.000208007	0.000659921	0.000379783
	Protein secretion system, Type VIII (Extracellular nucleation/precipitation pathway, ENP)	0.000106762	5.45848E-05	8.19249E-05	0.000102479
	Protein and nucleoprotein secretion system, Type IV	7.59602E-05	6.59455E-05	1.32614E-05	2.81866E-05
Phages, Prophages, Transposable elements, Plasmids	Phages, Prophages	0.010095552	0.00124057	0.011347872	0.001273807
	Transposable elements	0.008781031	0.003047399	0.007065314	0.003852535
	Pathogenicity islands	0.002765709	0.000438444	0.00361287	0.00082205
	-	0.000119587	8.3035E-05	3.85911E-05	6.28518E-05
Motility and Chemotaxis	Flagellar motility in Prokaryota	0.02120679	0.008809105	0.008760606	0.00624271
Respiration	Electron donating reactions	0.007731682	0.000735014	0.004712614	0.002234944
	ATP synthases	0.005879688	0.001416407	0.003105329	0.001313651
	Electron accepting reactions	0.004255997	0.000382361	0.004006055	0.000815977
	Sodium Ion-Coupled Energetics	0.003209521	0.000836236	0.00173872	0.001164642
Virulence, Disease and Defense	Resistance to antibiotics and toxic compounds	0.015575265	0.00316478	0.011116091	0.001861343
	Detection	0.001482428	0.000273957	0.001749597	0.00051248
	Adhesion	0.000888689	0.000183578	0.001078551	0.000425931
	Bacteriocins, ribosomally	0.000525629	0.00034621	0.000303432	0.000398592

continued

Table 3 continued

	synthesized antibacterial peptides				
	Invasion and intracellular resistance	2.47586E-05	2.96503E-05	8.31306E-06	2.06734E-05
Stress Response	Oxidative stress	0.007715306	0.000832054	0.006847427	0.001702424
	Heat shock	0.005245206	0.000693301	0.005103509	0.000558209
	Periplasmic Stress	0.001087627	0.000336429	0.000452542	0.000338038
	Osmotic stress	0.000433775	0.000222965	0.000283195	0.000175665
	Acid stress	4.87474E-05	4.29777E-05	2.61143E-05	4.24376E-05
	Detoxification	3.19721E-05	2.08644E-05	1.31418E-05	2.37941E-05
Metabolism of Aromatic Compounds	Peripheral pathways for catabolism of aromatic compounds	0.003635265	0.000577297	0.004513986	0.001785778
	Anaerobic degradation of aromatic compounds	0.001883113	0.000366801	0.001006804	0.000590269
	Metabolism of central aromatic intermediates	0.000502541	0.000195357	0.000305844	0.000169873
Regulation and Cell signaling	Regulation of virulence	0.003091922	0.000394667	0.004052434	0.0009115
	Quorum sensing and biofilm formation	0.002146904	0.000542855	0.002768609	0.000479538
	Programmed Cell Death and Toxin-antitoxin Systems	0.000728094	0.00026332	0.000717803	0.000352078
Sulfur Metabolism	Organic sulfur assimilation	0.000704346	0.000201049	0.000443267	0.000308787
	Inorganic sulfur assimilation	4.50508E-05	3.58173E-05	5.13531E-06	2.11229E-05
Secondary Metabolism	Biologically active compounds in metazoan cell defence and differentiation	0.000646517	0.000174183	0.000365623	0.000269666
Iron acquisition and metabolism	Siderophores	0.00025362	0.000118315	0.000126565	0.000156075

Appendix E

Comparisons between statistically significant genes GAP and Health

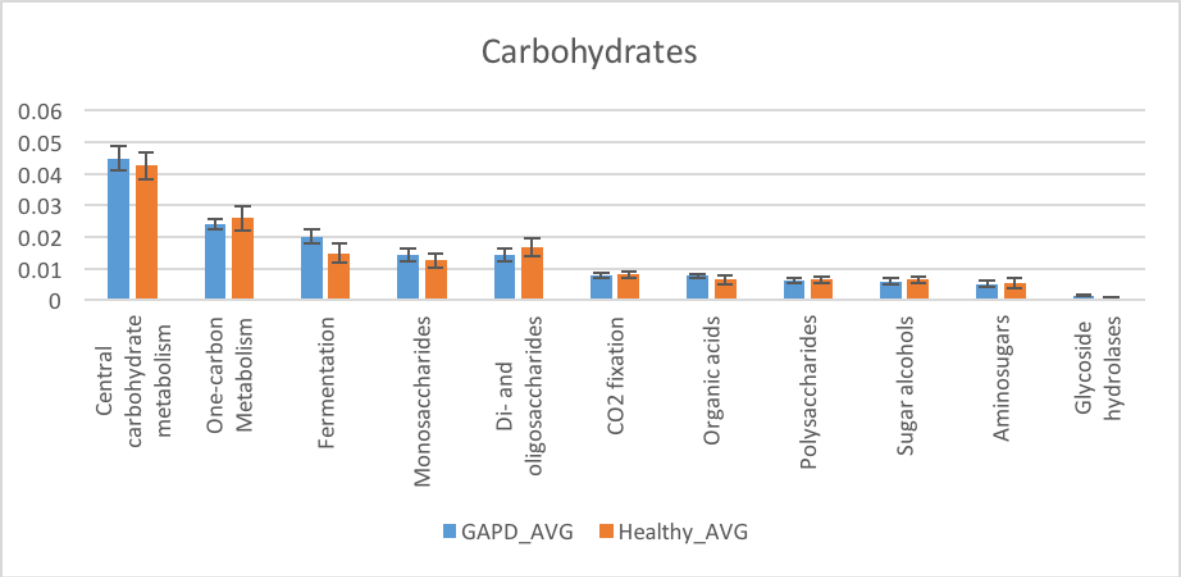

Figure 4 – Carbohydrates - Functional capability of GAP vs Health

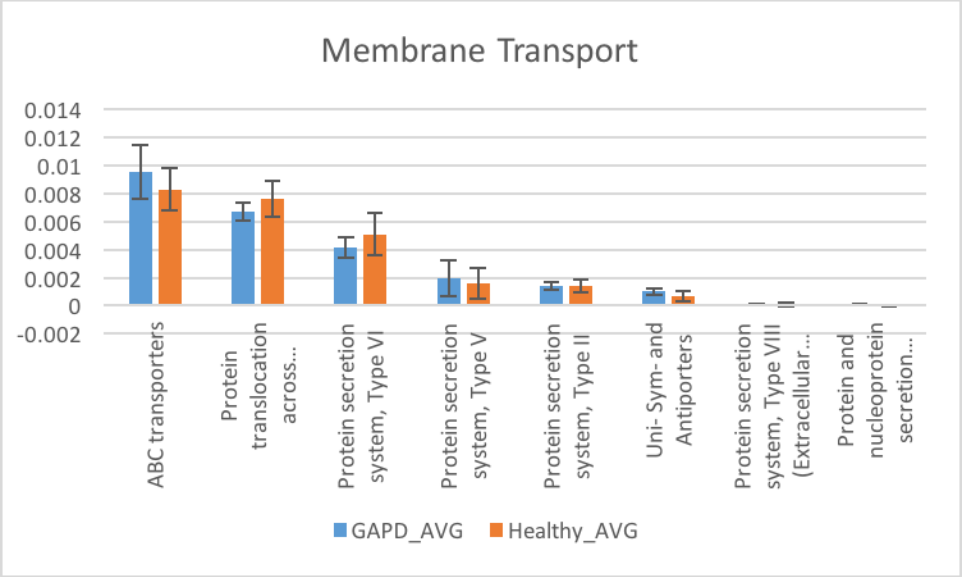

Figure 5 – Membrane Transport - Functional capability of GAP vs Health

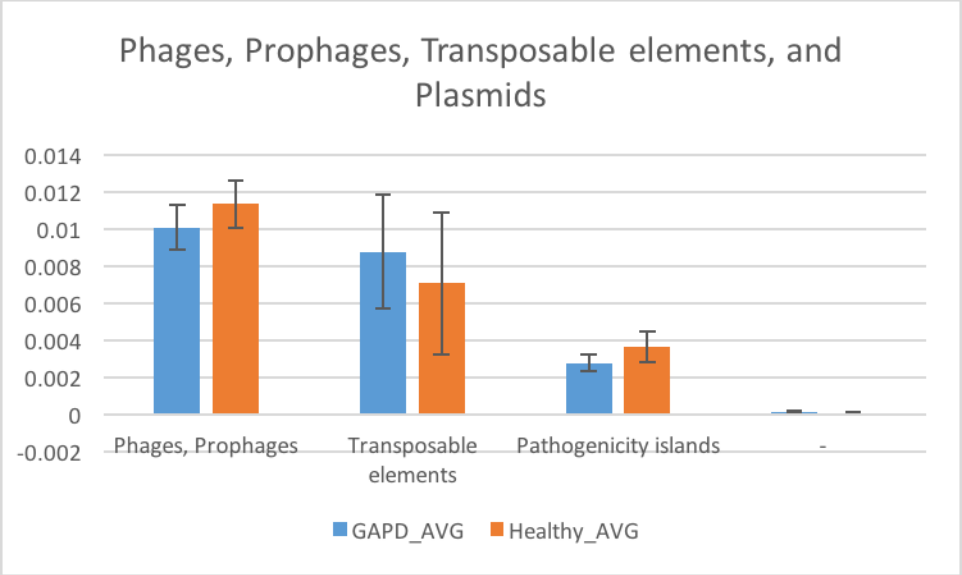

Figure 6 – Phages, prophages, transposable elements, plasmids - Functional capability of GAP vs Health

Appendix F

Comparison of Deep and Shallow Sites in GAP

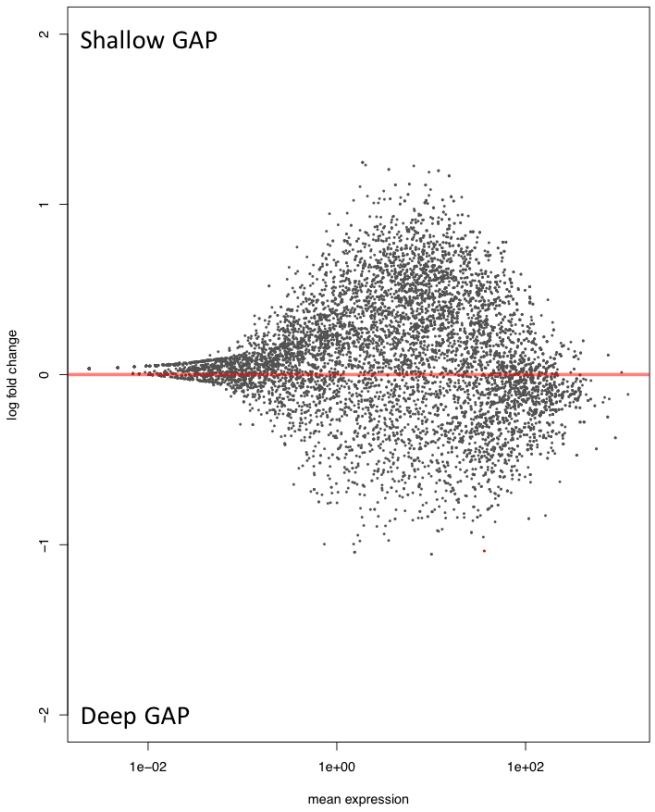

Figure 7 – logfold change in genes between deep and shallow sites of GAP.

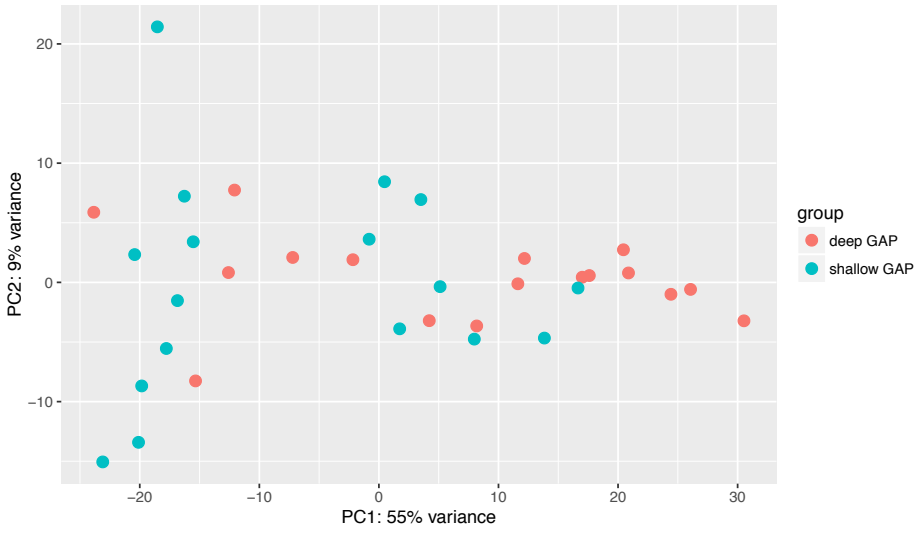

Figure 8 – PCA comparison between deep and shallow sites of GAP.

Appendix G

Comparison Between deep sites in generalized aggressive periodontitis (GAP) and chronic periodontitis (CP)

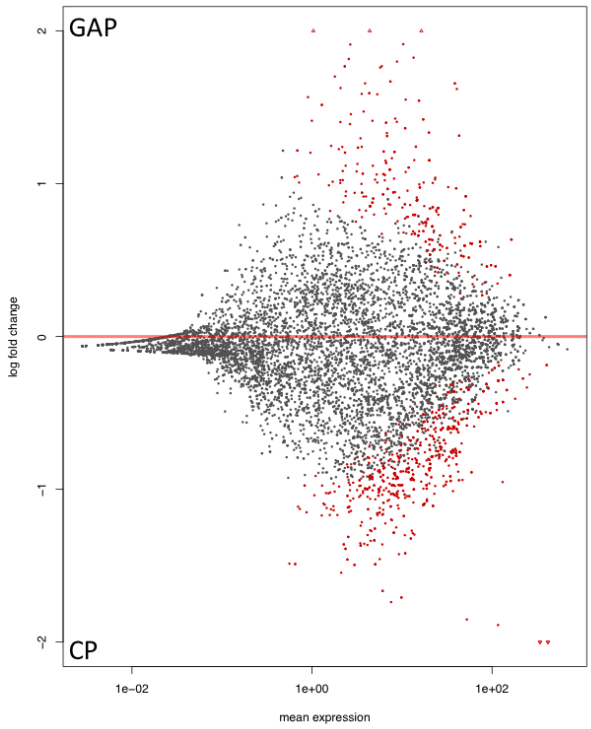

Figure 9 – logfold change in genes between GAP and CP. Red dots represent significantly abundant genes (P≥0.05; Wald-test, FDR adjustment)

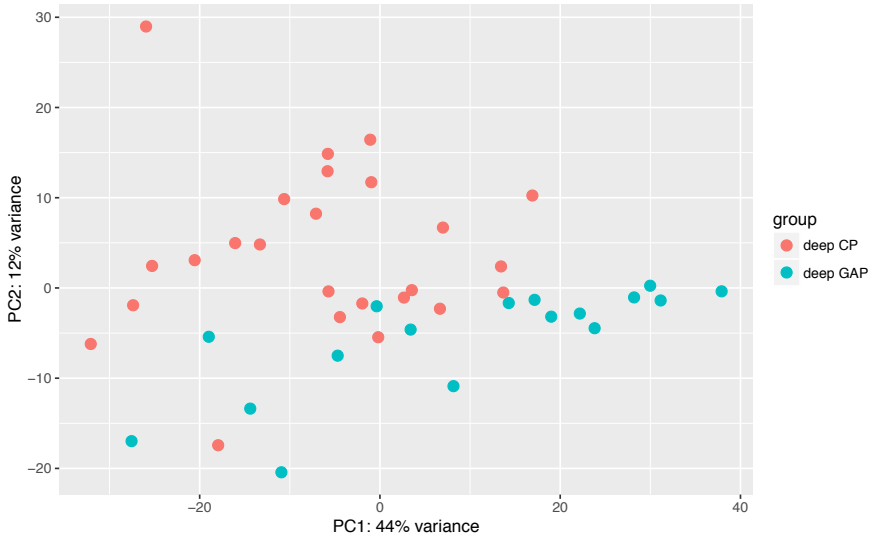

Figure 10 – PCA comparison between GAP and CP

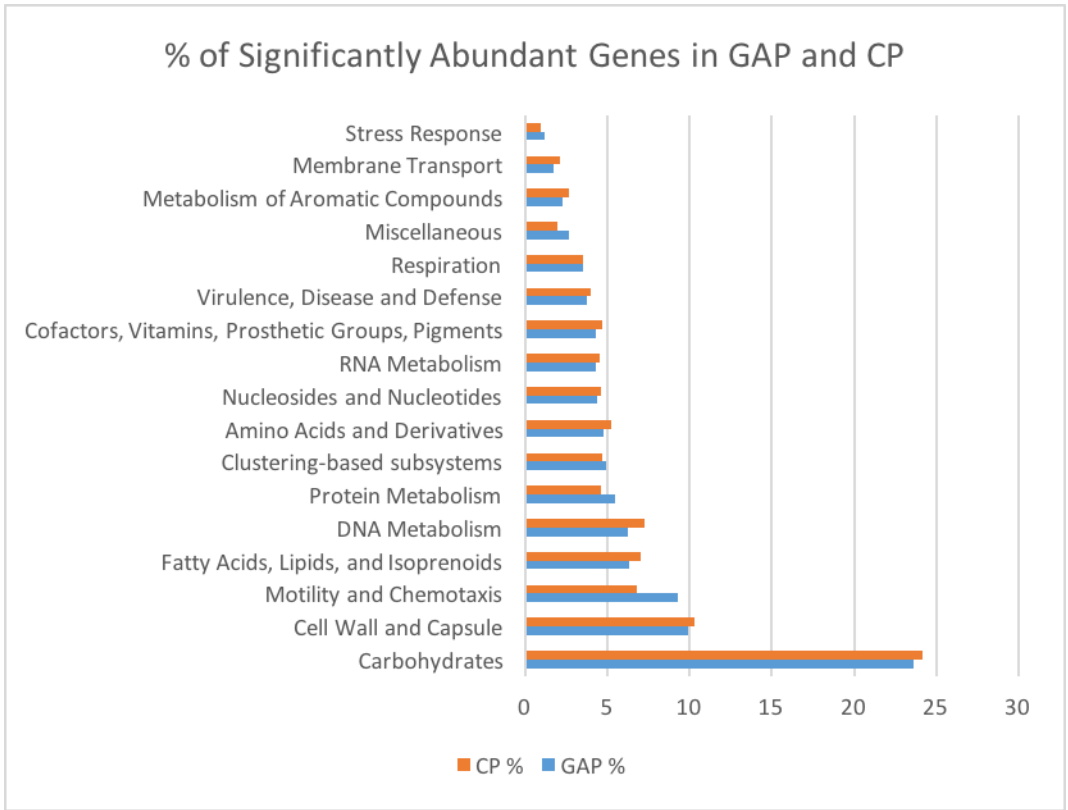

Figure 11 – comparison of GAP and CP patients for the statistically significant genes as a fraction of their total. Only genes occupying 1% or more shown here

Appendix H
GAP and CP

Table 4: Composition of statistically significant genes in GAP and CP as a percentage of their whole

level1	level2	GAP AVG	GAP STDEV	CP AVG	CP STDEV
Amino Acids and Derivatives	Alanine, serine, and glycine	0.016811781	0.002551866	0.017957558	0.007098682
	Arginine; urea cycle, polyamines	0.002994924	0.001309775	0.00146714	0.000862366
	Glutamine, glutamate, aspartate, asparagine; ammonia assimilation	0.027380124	0.00342813	0.032786306	0.010067709
	Proline and 4-hydroxyproline	0.000494799	0.000285857	0.000205757	0.000246
Carbohydrates	Aminosugars	0.008823208	0.001753347	0.009286076	0.003377454
	CO2 fixation	0.019331364	0.004169874	0.022196035	0.006340913
	Central carbohydrate metabolism	0.112919407	0.019285304	0.117162934	0.024697852
	Di- and oligosaccharides	0.01014923	0.004304208	0.007899753	0.004810601
	Fermentation	0.0107546	0.00252022	0.014391671	0.00734407
	Glycoside hydrolases	0.001496555	0.000795289	0.00137072	0.001441574
	Monosaccharides	0.03823996	0.003469442	0.038689722	0.011020533
	One-carbon Metabolism	0.003708689	0.001927946	0.003641143	0.002304363
	Organic acids	0.003708689	0.001927946	0.003641143	0.002304363
	Polysaccharides	0.016652365	0.003601603	0.015451392	0.005054959
	Sugar alcohols	0.010261465	0.004874919	0.007699423	0.003412058
Cell Wall and Capsule	Capsular and extracellular polysacchrides	0.025817302	0.003787072	0.027960525	0.007716975
	Gram-Negative cell wall components	0.061490558	0.006856763	0.062749194	0.015237252
	Gram-Positive cell wall components	0.011846134	0.001995115	0.012320872	0.003402864
Clustering-based subsystems	Biosynthesis of galactoglycans and related lipopolysacharides	0.011936844	0.002473013	0.015237226	0.006331885
	Flagella protein?	0.002362247	0.001215368	0.001919129	0.001498986
	Isoprenoid/Cell Wall Biosynthesis: Predicted Undec Diphosphate Phosphatase	0.008644211	0.001588574	0.009287494	0.00293081
	Methylamine utilization	0.001329569	0.000722251	0.000900077	0.000841129
	Monosaccharides	0.00333991	0.000575335	0.002349261	0.001675721
	Nucleotidyl-phosphate metabolic cluster	0.007967289	0.002420347	0.004123584	0.003533657
	Oxidative stress	0.001222253	0.000608455	0.000774187	0.000761161
	Probably Ybbk-related hypothetical membrane proteins	0.006245956	0.001535779	0.007921814	0.003058107
	Pyruvate kinase associated cluster	0.001274076	0.000731831	0.001112757	0.00198221
	Sarcosine oxidase	0.001329569	0.000722251	0.000900077	0.000841129
	Two related proteases	0.00333991	0.000575335	0.002349261	0.001675721

continued

Table 4 continued

Cofactors, Vitamins, Prosthetic Groups, Pigments	Biotin	0.022412161	0.004968859	0.026659132	0.007551213
	Coenzyme A	0.002493842	0.002149406	0.002169685	0.001516988
	Folate and pterines	0.010176172	0.002642448	0.010122399	0.004626405
	NAD and NADP	0.000744495	0.000629805	0.00035417	0.000564606
	Tetrapyrroles	0.007368583	0.002676761	0.007535568	0.003345923
DNA Metabolism	DNA repair	0.024734015	0.002788197	0.026930177	0.007840287
	DNA replication	0.037804116	0.009820985	0.045322613	0.012094898
Fatty Acids, Lipids, and Isoprenoids	Fatty acids	0.051110388	0.012819317	0.058459793	0.015722662
	Isoprenoids	0.002057585	0.000893951	0.001467604	0.001439486
	Phospholipids	0.009926292	0.00258945	0.010254213	0.003192975
Membrane Transport	ABC transporters	0.001316719	0.000491111	0.001058022	0.001165975
	Protein translocation across cytoplasmic membrane	0.015608357	0.002479108	0.019971285	0.004616321
Metabolism of Aromatic Compounds	Peripheral pathways for catabolism of aromatic compounds	0.022412161	0.004968859	0.026659132	0.007551213
Miscellaneous	Plant-Prokaryote DOE project	0.026182749	0.003720247	0.0194559	0.003502271
Motility and Chemotaxis	Flagellar motility in Prokaryota	0.092899263	0.03674605	0.067483554	0.029712652
Nucleosides and Nucleotides	Purines	0.017076271	0.00355883	0.016238503	0.005697313
	Pyrimidines	0.026715262	0.004111198	0.029889281	0.007776249
Phages, Prophages, Transposable elements, Plasmids	-	0.000878286	0.000564623	0.000171477	0.000279514
	Phages, Prophages	0.003571297	0.00136222	0.002794901	0.002086374
	Transposable elements	0.001685765	0.001293703	0.000954185	0.001005108
Protein Metabolism	Protein biosynthesis	0.010995244	0.001057812	0.008827522	0.004305189
	Protein degradation	0.019534178	0.003753906	0.015807317	0.00343408
	Protein folding	0.006166044	0.002607962	0.005702778	0.00474944
	Protein processing and modification	0.003571538	0.000947042	0.003322424	0.00277853
	Selenoproteins	0.014082815	0.003166607	0.012655155	0.005527843
RNA Metabolism	RNA processing and modification	0.033049782	0.004924277	0.035454216	0.007766228
	Transcription	0.010173089	0.002291217	0.009748663	0.003653734
Regulation and Cell signaling	Programmed Cell Death and Toxin-antitoxin Systems	0.001409828	0.000369675	0.000944248	0.000779902
	Quorum sensing and biofilm formation	0.000296064	0.000555148	0.000138851	0.000574745
	Regulation of virulence	0.006600152	0.001575887	0.005424283	0.002504768
Respiration	Electron accepting reactions	0.007724947	0.0042065	0.008233525	0.004345602
	Electron donating reactions	0.024603275	0.00457795	0.024181997	0.00652924
	Sodium Ion-Coupled Energetics	0.002854358	0.000892881	0.002861867	0.001621448
Stress Response	Osmotic stress	0.001329569	0.000722251	0.000900077	0.000841129
	Oxidative stress	0.004600892	0.002182701	0.002266443	0.002071781
	Periplasmic Stress	0.005953193	0.001875261	0.00598064	0.004041784
Sulfur Metabolism	Organic sulfur assimilation	0.000469654	0.000269269	0.000368973	0.000696528

continued

Table 4 continued

Virulence, Disease and Defense	Resistance to antibiotics and toxic compounds	0.037538612	0.005451919	0.039477194	0.009962676