Investigating the microbial “dark matter” of the human oral microbiome using cultivation-dependent and –independent approaches

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Karissa Lynn Cross

August 2019

ACKNOWLEDGEMENTS

This acknowledgements section is no doubt one of the most difficult sections for me to write because there have been so many people along my journey who have supported me and have helped make this possible. There are not enough words to address them all but for all of them I am beyond thankful.

First, I would like to thank my advisor Dr. Mircea Podar for his support and vision to encourage me to push the limits of what a PhD degree could mean. I thank all of the current and past Podar Lab members: Dr. Snehal Joshi, Shafer Belisle, Zamin Yang, Dr. Manasi Balachandran, Stephen Lebreux, Dr. Jim Campbell, Dr. Alisha Campbell, and Dr. Louie Wurch. I would also like to thank my committee Dr. Heidi Goodrich-Blair, Dr. Erik Zinser, Dr. Frank Loeffler, Dr. Stephen Kania, and formerly Dr. Igor Zhulin who have stood by my side year after year and offered their insight and support. Thank you to my collaborators both near and far throughout the years who have taught me how to listen and learn and have afforded me the opportunities to push my science beyond my comfort zone. Thank you to my friends both at ORNL and UTK whose companionship and love have no doubt made this possible.

A special thank you goes to William Ellis Baltz who I met my first year of graduate school and has continued to be by my side through not just the good days but the bad ones as well. I told him if he could make it through law school and I through graduate school at the same time, there is no limit to what we could accomplish together, and he continues to support my dreams steadfast. Also, I am thankful for his family Liana, Billy, and Sarah (and many more) who have made me feel welcomed and loved.

Lastly, and at the top of my list, I thank my family. My grandparents, Nancy and Lowell, raised my sister and I and have supported us through all of our ventures. With their support I have had the confidence to pursue my goals where no one in our family has yet gone. They have served as strong role models demonstrating where hard work and perseverance can take you, and about never giving up or backing down. To my sister Ashley, thank you for the many phone calls and advice as we transverse this life together. With every step you have given me the assurance and love to keep moving forward when at times it felt impossible and I look forward to where this life continues to take us.

ABSTRACT

The advancement of sequencing technology has provided a renewed appreciation for the diversity of life on earth with most environments being dominated by yet-uncultured microorganisms. Analysis of 16S rRNA gene sequencing libraries of the original Human Microbiome Project dataset revealed that approximately 70% of sequences belonged to bacteria that were uncultured in the laboratory and current metagenomic analysis of the human microbiome has proposed 77% of taxa lacked sequenced genomes. As the human microbiome is important for healthy host physiology, these findings emphasize just how little we currently know about the microorganisms that reside in the human body. Understanding the microorganisms that make up the normal human microbiome is the first logical step in determining, during dysbiosis and disease, how the community changes as a whole so as to permit investigation into preventative and/or treatment measures during disease. For example, many disease-associated lineages of bacteria in periodontal disease remain uncultured in the laboratory and therefore limited insight is available as to whether or not these organisms play a role in disease etiology or their presence is a secondary effect of the disease state. Therefore, cultivating more microorganisms from the human microbiome is critical as it enables hypothesis-driven experimentation into the roles of those organisms in human health and disease. The work presented here represents a two-fold approach for cultivation of bacteria from the human oral microbiome by both classical means and by the development of a novel, “reverse-genomics” approach. First, I identified a specific interaction between *Desulfobulbus oralis* and *Fusobacterium nucleatum* that resulted in the isolation of the first *Desulfobulbus* genus member from the human oral cavity and I characterized this bacterium both physiologically and genomically. Next, I investigated what causes the recalcitrance of some oral bacteria to cultivation and what contributes to their dependence on other bacteria, such as *F. nucleatum*. Lastly, we leveraged existing public sequence data to target specific TM7 bacteria from within the microbial community for both single-cell genomic characterization and cultivation. Collectively this work demonstrates how cultivation-dependent and cultivation-independent methodology can be combined to investigate the “dark microbial matter” in the human oral cavity.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE - Introduction

Publication Note

Portions of this introduction will be submitted as part of a mini-review for the journal *Environmental Microbiology*. Mircea Podar was invited for the review and all material contained within this chapter was written by K.L.C.

The Human Microbiome and Uncultured Bacteria

Microbes have been evolving for billions of years on Earth, well before the emergence of primates. Humans have co-evolved with our microbial counterpart – the human microbiome (1, 2). We may begin acquiring our microbiome even before birth (3-5) and homeostasis with our microbiota is pertinent to our health (6). Soon after the Human Genome Project was completed, the Human Microbiome Project (HMP) was conceptualized. The HMP sought to identify all of the microorganisms residing in/on the healthy human body so as to lay the groundwork for understanding, in part, their ecological roles in human health and disease (1). Using metagenomics and 16S rRNA gene analysis, the HMP began to elucidate the vast diversity of microorganisms in the human microbiome and identified specific groups of microorganisms dominating different niches (2). In doing so, a large collection of microbial reference genomes was produced that further exemplified just how little we may know about the microorganisms that live in/on the human body. These analyses have provided a robust baseline for what may constitute a healthy human microbiome and has set the stage for further studies looking at microbiome variations during different human states and diseases.

A healthy microbiome is important, in part, for to its abilities to provide inaccessible nutrients for host consumption and to provide a line of defense against foreign invaders. The field of clinical microbiology began with the culturing of microorganisms from the human body. Historically, the main focus was on microorganisms that cause disease, which led to Koch's postulate by which a pathogen must be isolated from a sick individual to be deemed the causative agent of a disease. This resulted in the development of media that supported growth of such pathogens. Conversely, when microorganisms were cultured from the environment, such media did not necessarily work due to organisms having different metabolic requirements and growth dynamics. By implementing new methods, such as the Winogradsky column, more diverse organisms could be cultured.

Recently we have begun to appreciate this in regard to understanding the microorganisms of the human microbiota, as 16S rRNA gene-based studies have correlated native microbial species to human disease with many being currently uncultured in the laboratory by standard culturing techniques (i.e. periodontitis (7)).

Preliminary analysis in 2012 by Fodor et al. of the HMP dataset identified about 70% of OTUs belonged to “uncultured” or “clone” designations (8). They further ranked OTUs as “most wanted” for whole genome sequencing and/or cultivation if they had less than 90% identity (16S rRNA gene V1-V3 and V3-V5 regions) to those existing in the GOLD-Human or HMP databases and noted that only between 1.7% and 10% of total HMP OTUs belonged to this “most wanted” list (8). It became evident that the human host harbors numerous uncultured microorganisms, however few belong to those phylogenetically novel bacterial lineages. More recent database mining at the 16S level has suggested between 45%-97% of human microbiome taxa may, in fact, be cultured (9). The advancement of sequencing and binning technology for metagenomics has currently allowed us to move away from just 16S gene level analysis and take the whole genome into consideration, which is important when considering the difficulty in assigning species and strain level taxonomy (10). A recent metagenomics analysis of over 150,000 metagenomes from the human microbiota identified 77% of their metagenome-assembled genomes (MAGs) previously belonged to microorganisms with no available genomes (either WGS of isolates or MAGs) (11). While great strides have been made in understanding “who is there” in the human microbiome, the physiology of most microbes, their interactions with each other, and their impact on human health and disease are just beginning to be uncovered. More reference genomes, and most importantly more cultured representatives, are required to advance from mere associations-based inferences so as to permit hypothesis driven experimentation at the interface of human health and disease.

The Oral Microbiome and Oral Diseases

In 2010, the Human Oral Microbiome Database (HOMD) was created, relying primarily on 16S rRNA gene sequences from clone libraries, that provides a comprehensive database to house phylogenetic and genomic information pertaining to the plethora of bacteria that reside in the oral cavity (12). At the time there was an

estimated 619 bacterial taxa from 13 phyla making up the human oral microbiota, 96% of them belonging to just 6 phyla, with approximately 34% of total taxa being uncultured (12, 13). Currently, the HOMD houses reference information for 770 microbial species and has been expanded (eHOMD) to also include microorganisms of the human aerodigestive tract (14). Of these approximately 770 microbial species, 62% have genomic data available, however 30% remain as uncultivated phylotypes in the oral cavity (www.homd.org v14.51)(14). There is still much to learn as to why these microorganisms remain uncultured and what role those microorganisms play in human health and disease.

Periodontitis is a common disease that results from host inflammatory responses to native oral bacteria (7, 15, 16) and has a complex etiology originating from within the human oral microbiome as a result of synergistic and dysbiotic interactions between members of the microbial community (17, 18). This inflammation results in swelling or bleeding of the gums, tooth loss, and bone and tissue loss (7, 19). The exact etiology of periodontal disease remains unclear due to the complexity of the changing oral microbial community (20). Specifically, periodontal diseases do not arise from a single, easily detectable pathogen but rather as a result of a native community shift favoring keystone species (21). These species are capable of provoking a proinflammatory response that can further alter the community and lead to dysbiosis (21).

The oral cavity provides a variety of niches, such as saliva, supragingival, and subgingival locations, for the growth of distinct organisms with varying metabolic requirements (i.e. aerobes vs. anaerobes) (22). In 2012, Griffen et al. used 454 pyrosequencing to compare the subgingival microbial community between healthy and chronic periodontal patients to understand how the microbial community may change as a result of disease (7). They discovered that 53 microbial species were significantly more abundant in health while 123 species were more abundant in disease (7). Of those identified to associate with disease, over half were currently uncultured, with entire phyla existing with no cultured representatives at the time (7). Likewise, using Illumina sequencing of 16S rRNA gene libraries of the subgingival microbiome, eight specific classes of bacteria were correlated with periodontitis including *Chloroflexi*, *Deltaproteobacteria*, *Spirochaetes*, and *Synergistetes* (23), each encompassing yet-cultured representatives (7). In addition to periodontal disease, another common disease

is dental caries which results in dental demineralization. A survey of the supragingival microbiome of twins with and without caries identified uncultured bacteria from the candidate phyla TM7 and GN02 present in the samples with a particular uncultured TM7 bacterium correlating with taxa associated with caries-positive subjects (24).

Cultivation-independent Studies

Although 16S rRNA gene studies can give us an idea about “who is there”, next-generation sequencing provides a more comprehensive genomic analysis where we can begin to make inferences about what these microorganisms may be doing based on their genomic content. Two methods to sequence the genomes of currently uncultured microorganisms from within complex microbial communities is by single-amplified genomics (SAGs) and metagenomic-amplified genomics (MAGs). Each approach has pros and cons and it ultimately depends on the research question being asked as to which approach works best for which situation (reviewed in (25, 26)). One benefit of SAGs is that the approach can be more targeted towards organisms of interest. Further, since sequenced DNA is from a single cell, this method provides species/strain level resolution (25, 26). Conversely, due to amplification bias associated with SAGs, MAGs tend to provide a more complete genome with even sequencing coverage as well as the opportunity to provide a larger microbiome-wide look at the genomic content of each sample set (25).

An interesting group of microorganisms that reside in the human oral microbiota belong to the “candidate phyla radiation” (CPR) (27), which includes TM7, GN02, and SR1 bacteria, deep-branching uncultured lineages. Single-cell genomics provided an initial view into some of these organisms in the oral cavity. The first genomic information for TM7 came from single-cell genomes in 2007 in which the phrase microbial “dark matter” was first coined (28). An early 16S rRNA gene clone study had previously investigated TM7 diversity in the human oral cavity (29) and found that TM7 rRNA gene abundance was higher in patients exhibiting periodontal disease, relative to total bacterial rRNA, with a specific TM7 subgroup being detected more in diseased than in healthy sites (29). That first glimpse at the genomic information contained within a TM7 genome generated initial hypotheses as to the metabolic potential for TM7 bacteria in the oral

cavity and recognition that they encoded Type IV secretion systems (28, 30, 31). Furthermore, a single-cell genome of an oral SR1 bacterium identified that this particular group of organisms harbored a unique genetic code by recoding the traditional UGA stop codon as a glycine (32).

In addition to phylogenetically novel microorganisms, single-cell genomics has helped elucidate the potential roles other uncultured disease-associated lineages of bacteria may play in the oral cavity. As stated previously, 16S rRNA gene studies revealed correlations between *Deltaproteobacteria* and *Chloroflexi* with periodontal disease (7, 23). Further single-cell genomic sequencing of these groups of organisms provided novel insights into their host adaptation and a putative role in disease and also provided hypotheses for cultivation in the laboratory (33, 34). Uncovering information about disease-associated microorganisms via single-cell genomics has been useful in furthering our understanding of health-associated species, as is the case with a previously uncultured health-associated *Tannerella* species (35).

As periodontal disease and caries have a polymicrobial etiology, metagenomic studies can resolve community-wide genomes to provide further insights into the whole community's functional potential (36). One of the first metagenomic studies to look at the differences in community composition in oral health and disease was in 2012. This study demonstrates that healthy individuals (those who had never suffered from dental caries) were significantly enriched in certain functions relating to antimicrobial peptides and quorum sensing (36). Based on that, it was hypothesized that some of the health-associated oral bacteria, specifically some streptococci, could act as "probiotics" by having an antagonistic effect on cariogenic bacteria (36). This led to cultivation of 249 isolates from healthy individuals, of which 16 were capable of inhibiting the growth of cariogenic bacteria, most being closely related to *Streptococcus oralis, S. mitis*, and *S. sanguis* (36).

Single-cell genomes and metagenomes can reveal the physiological potential of a microorganism, but the functions encoded in genomes are not always expressed at either the RNA or protein level. Further cultivation-independent approaches can harness technologies such as transcriptomics and proteomics to measure the relative abundance of mRNA transcripts or proteins which can serve as a representation of the metabolic

activity of specific organisms or the microbial community as a whole. As noted in the original HMP dataset, microbial community structure can change considerably in healthy individuals but on a functional level may remain consistent (6). A more robust analysis of the human oral microbiome using metatranscriptomics revealed that periodontal patients displayed conserved metabolic profiles during disease (37). Specifically, gene expression relating to lysine fermentation to butyrate and pyruvate fermentation were significantly more expressed in all diseased sites indicating that these functions may be important for influencing disease-associated microbial communities (37). RNAseq further provided species-level resolution and revealed that *Fusobacterium nucleatum* was responsible for lysine degradation to butyrate in those individuals, despite no significant change in *F. nucleatum* rRNA abundance, indicating that a shift in gene expression by specific organisms may contribute to the changing environment that results in oral diseases (37). Similarly, specific metabolic changes occur within the microbial community during the onset of gingivitis in which the community favors genes involved in pathways such as propanoate and butanoate metabolism (38).

Although the information gained from sequencing studies can provide useful insight into the functions of the human microbiota, it still remains a bioinformatics end-point. Further cultivation of currently uncultured bacteria can fuel hypothesis-driven studies about the role these microorganisms play in their environments. Isolation enables us to advance from disease correlations to understanding etiology and specific pathogenicity mechanisms so that effective treatment can be achieved.

Cultivation-dependent Studies

Some periodontal pathobionts have been identified but many disease-associated lineages remain uncultured (17). With the 770 bacterial taxa identified in the human oral cavity, 1/3rd represent yet-cultivated members (12, 13). Their recalcitrance to culturing under standard culture methods may be due to their dependencies on co-evolving bacterial community members, an idea supported by recent advances in culturing such organisms. For example, *Fretibacterium fastidiosum* (phylum Synergistetes) is dependent upon *Fusobacterium nucleatum* for growth in pure culture (39). Another example is cultivation of a human health-associated *Tannerella* species (40), first

characterized by single-cell genomics (35). That isolate, *Tannerella* BU063, requires 'helpers' such as *Prevotella intermedia* or *Propionibacterium acnes* to stimulate growth, with colonies of BU063 forming satellites around other oral species (40). Likewise, the first oral bacterium cultured from the phylum *Chloroflexi, Chloroflexi* taxon *Anaerolineae* bacterium HOT439, requires 'helper' bacteria, including *F. nucleatum*, for growth (41) and was isolated by adding siderophores to culture medium (41). These findings suggest a reliance of these organisms on a community lifestyle, where species interact and may depend upon each other, an aspect that needs to be considered when designing cultivation experiments.

In some instances, catering media variations to specific growth requirements enables successful cultivation of previously uncultured microorganisms (42, 43). Traditional media (such as LB or BHI) used to culture bacteria does not support the growth of a high diversity of oral bacteria (43). However, if steps are taken to develop a medium that meets the nutritional requirements of more diverse oral microorganisms then perhaps more organisms may be brought into culture. One such success is the development of "SHI medium" that supports the growth of high percentages of uncultured oral bacteria (43). Similarly, longer-term anaerobic cultivation aimed at identifying slow-growing and/or satellite colonies on diverse media arrays was successful in isolating novel *Bacteroidetes*, *Lachnospiraceae* and *Clostridiales* species (42).

Bacteria from the candidate TM7 phylum (more recently renamed "Saccharibacteria"(44)) have been recalcitrant to cultivation for many years. The first representative to be cultured, "TM7x," was isolated from the human oral cavity in 2015 (45). TM7x is an obligate epibiotic parasite on *Actinomycetes odontolyticus* XH001 (45, 46), confirming earlier hypotheses that these organisms may be heavily reliant on others for growth (47). TM7x masks a proinflammatory response when incubated with its host and bone marrow macrophages, as compared to host alone, and enhances biofilm formation when grown with its host (45, 48).

In addition to traditional approaches for culturing novel microorganisms of the human oral cavity, newer technologies have emerged to facilitate further cultivation of the uncultured. Many organisms in the human oral cavity have a biofilm lifestyle where species grow and interact in complex ways resulting in strict interdependencies where

growth may only occur in a community (49, 50). In-vitro biofilm models can be employed to grow stable, reproducible, communities that harbor uncultured microorganisms. One such example resulted in the growth of 51 uncultured oral taxa including a TM7 (51). Another example is the use of microfluidics, microdroplets, and artificial teeth to simulate the natural environment and allow species to interact as they would in the oral cavity to support their growth (52-54).

Sulfate Reducing Bacteria

As periodontitis leads to the formation of deep subgingival pockets, this allows proliferation of anaerobic microorganisms. Sulfate-reducing bacteria (SRB) have been detected in periodontal pockets (55) and implicated in other diseases such as ulcerative colitis (56). Inherently SRB could be involved in such destructive diseases due to the effects of their metabolic by-products (33, 57) such as hydrogen sulfide that results from dissimilatory sulfate reduction. SRB are not abundant members of the oral microbial community, with the most common being species of *Desulfovibrio*. Isolates of *Desulfovibrio fairfieldensis* have been associated with periodontitis (55) while culture-independent studies have linked the SRB genus *Desulfobulbus* with periodontal disease in the oral cavity (23, 58, 59).

Approximately eleven environmental *Desulfobulbus* species have been isolated and formally characterized to date (60-67). A single *Desulfobulbus* OTU exists within the human oral cavity (HOT041). It is generally present at abundances below 1% of the total microbial community (7, 58) but in periodontitis can reach levels as high as 2% (58). Environmental *Desulfobulbus* species are of particular interest due to their ability to methylate mercury producing a powerful neurotoxin (68). Should oral *Desulfobulbus* HOT041 be able to produce methylmercury this could have significant health implications. Cultivation of host-associated *Desulfobulbus* and its comparison to environmental *Desulfobulbus* would provide important insights into the evolution of this genus and in identifying whether anaerobic sulfate-reducers merely persist in disease sites or are capable of actively contributing to disease.

Candidate Phylum TM7 Bacteria

As mentioned previously, the human oral cavity is home to three candidate phyla lineages of bacteria: TM7, GN02, and SR1. The TM7 phylum (Candidate Division Saccharibacteria) is ubiquitous and is found in environments ranging from arctic regions (69), sediments (47, 70), wastewater (71), bioreactors (44, 72, 73), and the human body (28, 29, 45, 74, 75). The first identification of this bacterium was in 1996 when clone sequences were prepared from a peat bog in Germany (76, 77). However, their clone 'TM7' represented a solitary deep rooting phylogenetic lineage and no inferences could be made concerning its role in the environment. Hugenholtz et al. performed the first targeted study of the TM7 phylum in 2001 (73) in which they used new and existing 16S rRNA gene sequences from environmental samples to garner a more robust phylogenetic analysis of these bacteria. This study greatly expanded our view of TM7 phylogeny (73). The TM7 phylum still remains an enigmatic lineage of bacteria to this day with slow strides being made in their study.

Recently, deep-sequencing community analysis revealed TM7 bacteria to be enriched in disease samples relative to non-disease samples. Further genomic investigation identified potential virulence factors encoded in TM7 genomes, connecting genotype to a potential disease function (78). Generally present at below 1% of the healthy oral community (29, 70), TM7 abundance can increase in periodontal bacterial communities (29, 78). The advancement of whole genome sequencing has continued to provide a more robust look at TM7 bacteria (28, 70), identifying over a dozen lineages of TM7 in the oral cavity and noting the average size of TM7 genomes being close to the lower limit of life for free-living bacteria (44). Partial and completed TM7 genomes (28, 70) seem to suggest TM7 have limited metabolic potential even in free-living environments. Genome streamlining and reduction can occur as bacteria transition from free-living to host-associated environments. This often results in restricted host range and in some cases the acquisition of pathogenicity traits (79-81), but limited knowledge can be acquired from just sequencing analysis alone. More unique genomes representing specific and diverse OTUs from host-associated environments are desirable and without more cultured representatives, TM7 adaptation to the oral cavity and potential role in disease remains unknown.

Non-bacterial Diversity of the Human Oral Cavity

Many studies have focused on the highly diverse communities of bacteria that reside in the human microbiota. However, all branches of life are present including fungi, archaea, and viruses. This varies by body site where some specialized niches may harbor none (i.e. subgingival microbiome (24)). Of the HMP dataset, only 0.01% of sequences aligned to fungal genomes (82). The most abundant fungi identified in human stool samples is *Saccharomyces* (82) while for the skin it is *Malassezia* (83), and for the oral and vagina it is *Candida* (84, 85). Many fungi from the human microbiome have cultured representatives at the genus level but cultivation of these different groups is inconsistent due to variability regarding which species colonize which individuals and their potential transience (82). For example, in 2010 one of the first studies to look at uncultured fungi within the oral cavity identified 74 culturable and 11 non-culturable fungal genera (85). However 39 genera occurred only once among the 20 patients analyzed emphasizing how much variability occurs between individuals (85). Continued advancement in metagenomics and single-cell genomics can aid in further exploration of fungi that lack uncultured or unsequenced relatives (86, 87).

Archaea are an even more underexplored component of the human microbiome; with many uncultured representatives and none known to cause disease, although some associations have been made (88). The most abundant archaea detected and perhaps the best studied, are the methanogenic archaea (89) hypothesized to be keystone species that play crucial roles in overall community composition (reviewed in (90)). Of the methanogenic archaea, five have been isolated from either fecal samples or the oral cavity (90). *Methanobrevibacter oralis* is the most dominant oral archaeon (91). Methanogenic archaea form syntrophic relationships with other bacteria and these interspecies dependencies may result in the recalcitrance of other bacteria that depend on archaeal partner metabolic byproducts for their own metabolism (91). Despite some progress made in the cultivation of archaea, there is much unexplored diversity to still uncover.

Perhaps the most difficult components of the human microbiome to study is the viral portion (or the virome) in part because of the lack of common markers for viruses, the variability in their genomic content (i.e. DNA vs RNA viruses), and limited reference

material. The human microbiome harbors an estimated number of 10 viruses for every 1 bacterial cell (92), with many viruses residing in the oral cavity, gut, vagina, and skin. The majority are bacteriophages (93-96). The skin and oral cavity have relatively stable viromes (94, 95) rather than being highly transient at these exposed locations. Such sequencing analyses can detect virus-like-particles and provide insight into possible functions of the virus community.

As we continue to learn more about the bacterial portion of the human microbiome, we further appreciate the tight communities they form with each other as well as how their interactions with fungi, archaea, and viruses may help shape their interactions with the human host. More work is required to fill databases and culture collections to continue to advance our understanding of the diverse microorganisms of the human microbiome. As cultivation-dependent and -independent approaches continue to advance, no branch on the tree of life may remain untouched.

Research Objectives

Despite advances in both cultivation-independent and cultivation-dependent studies of the human microbiome, many microorganisms remain uncultured. This results in a lack of hypothesis-driven experimentation regarding their role in human health or disease. The work described in this dissertation collectively aims to overcome this barrier and bring previously uncultured microorganisms into laboratory cultures by refining traditional methodologies and also developing new approaches **(Figure 1.1)**.

In **Chapter 2** I aimed to isolate and characterize the human oral *Desulfobulbus*, a sulfate reducing bacterium for which single cell genomic data were previously obtained in our laboratory (33). This particular organism has been linked to periodontal disease and single-cell genomic analysis has identified potential virulence factors that may play a role in disease progression. Oral *Desulfobulbus* has evaded cultivation and it has been unknown whether this bacterium merely persists in subgingival pockets or actively contributes to a proinflammatory environment. Chapter 2 describes the isolation and characterization of *Desulfobulbus oralis* sp. nov. and provides evidence of the potential virulence of this organism.

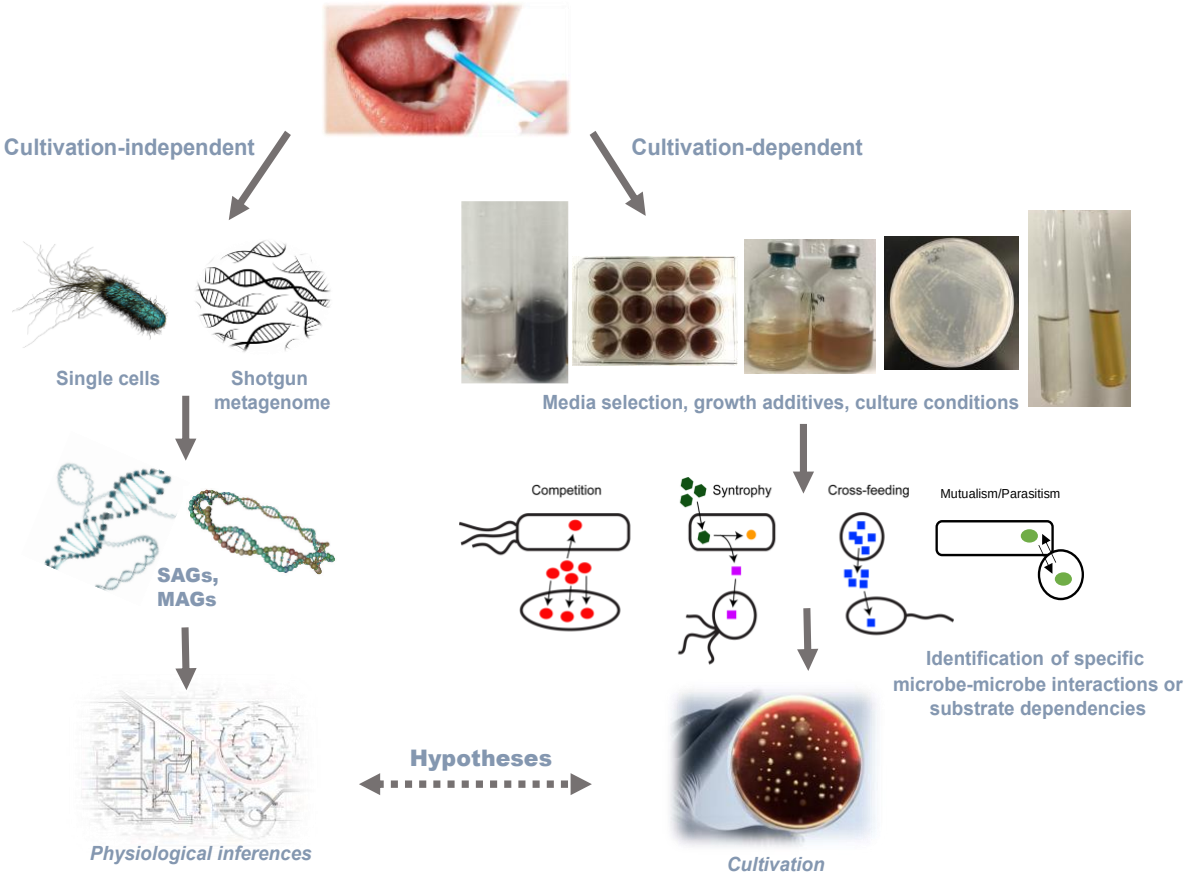


Figure 1.1. Graphical abstract outlining cultivation-dependent and cultivation-independent approaches for the study of human microbiome samples. Methods such as single cell genomics and metagenomics can elucidate the potential functions microorganisms have in their environments (left) and can be used to formulate hypotheses about how to culture them. Coupling this with traditional microbiology (right) we can uncover specific species interactions that cannot be inferred from sequencing alone.

During the isolation of *D. oralis* in Chapter 2, I uncovered clues as to why this organism has been recalcitrant to cultivation. *D. oralis* is dependent on *Fusobacterium nucleatum* by-products for growth. **Chapter 3** investigates this dependency by using metabolomics facilitated by mass spectrometry to understand the metabolic by-products of *F. nucleatum*. As other bacteria have been shown to also have this dependency, identifying metabolites produced by *F. nucleatum* may assist in the cultivation of other currently uncultured microorganisms of the human microbiome.

Finally, to begin to understand other oral lineages that lack cultured representatives, **Chapter 4** details a “reverse-genomics” cultivation approach using TM7 bacteria as a test case. TM7 bacteria are phylogenetically diverse in the oral cavity (over a dozen distinct OTUs) with only one representative cultured to date (a single Group 1 TM7 out of the 6 identified clades). We demonstrate the utility of this approach by culturing three new TM7 species and by sequencing 23 single-cell TM7 genomes.

CHAPTER TWO - Insights into the evolution of host association through the isolation and characterization of a novel human periodontal pathobiont, *Desulfobulbus oralis*

Publication Note

The following chapter is modified from a manuscript published in © American Society for Microbiology in the journal mBio, volume 9, issue 2, year 2018, DOI 10.1128/mBio.02061-17. Other co-authors include P. Chirania, W. Xiong, C.J. Beall, J.G. Elkins, R.J. Giannone, A.L. Griffen, A.M. Guss, R.L. Hettich, S.S. Joshi, E.M. Mokrzan, R.K. Martin, I.B. Zhulin, E.J. Leys, M. Podar.

Author's contributions.

E.M.M., A.L.G., and E.J.L. provided oral samples and original media recipes and cultivation feedback.

K.L.C. did laboratory cultivation experiments, *D. oralis* isolation, and physiological characterization.

K.L.C. analyzed cultures by FISH, qPCR, and SEM.

K.L.C. worked with R.K.M. to perform HPLC with input from J.G.E.

K.L.C. led whole genome sequencing efforts, annotation, and analysis with M.P.

S.S.J. led preparation of the Circos image.

K.L.C. supervised P.C. for a summer rotation project to perform the oral epithelial proinflammatory assay.

W.X., R.J.G. and R.L.H. performed proteomics and mass spectrometry analysis.

J.G.E. and A.M.G. provided intellectual contributions concerning central metabolism and reconstruction.

I.B.Z. performed the MiST database analysis.

C.J.B. performed the metatranscriptomics mapping.

M.P. oversaw all aspects of the project and participated in data interpretation.

K.L.C. and M.P. wrote the manuscript.

Abstract

The human oral microbiota encompasses representatives of many bacterial lineages that have not yet been cultured. Here we describe the isolation and characterization of previously uncultured *Desulfobulbus oralis*, the first human-associated representative of its genus. As mammalian-associated microbes rarely have free-living

close relatives, *D. oralis* provides opportunities to study how bacteria adapt and evolve within a host. This sulfate-reducing deltaproteobacterium has adapted to the human oral subgingival niche by curtailing its physiological repertoire, losing some biosynthetic abilities and metabolic independence, and by dramatically reducing environmental sensing and signaling capabilities. The genes that enable free-living *Desulfobulbus* to synthesize the potent neurotoxin methylmercury were also lost by *D. oralis*, a notably positive outcome of host association. However, horizontal gene acquisitions from other members of the microbiota provided novel mechanisms of interaction with the human host, including toxins like leukotoxin and hemolysins. Proteomic and transcriptomic analysis revealed that most of those factors are actively expressed, including in the subgingival environment, and some are secreted. Similar to other known oral pathobionts, *D. oralis* can trigger a pro-inflammatory response in oral epithelial cells, suggesting a direct role in the development of periodontal disease.

Importance

Animal-associated microbiota likely assembled as a result of numerous independent colonization events by free-living microbes followed by co-evolution with their host and other microbes. Through specific adaptation to various body sites and physiological niches, microbes have a wide range of contributions, from beneficial to disease causing. *Desulfobulbus oralis* provides insights into genomic and physiological transformations associated with transition from an open environment to a host-dependent life style and the emergence of pathogenicity. Through a multi-faceted mechanism triggering pro-inflammatory response, *D. oralis* is a novel periodontal pathobiont. Even though culture-independent approaches can provide insights into the potential role of the human microbiome “dark matter”, cultivation and experimental characterization remain important to studying of the roles of individual organisms in health and disease.

Introduction

Periodontitis is one of the most prevalent human diseases around the world. Moderate stages affect more than half of adults while the severe form, involving tooth loss, occurs in 10-20% of the population (97, 98). Periodontitis has a complex

polymicrobial etiology with current models implicating synergistic and dysbiotic interactions between members of the gingival microbial community (including species present at low abundance) as the disease onset (16, 21, 99, 100). Plaque biomass accumulation followed by changes in the composition of the microbiota lead to and advance gum inflammation (gingivitis). Further increase in microbial biomass, exacerbated by poor hygiene and genetic predisposition sustain additional shifts in the microbial community composition, which becomes increasingly anaerobic in deepening periodontal pockets as periodontitis develops (100). At that stage, microbial signaling, with some species playing keystone disease roles, interferes with the normal host immune defense and causes a persistent inflammatory response, resulting in tissue and bone loss (101, 102).

Among the first bacteria implicated in the etiology of periodontal disease were *Porphyromonas gingivalis, Treponema denticola* and *Tannerella forsythia* (103). More recent microbiome sequencing and disease association studies suggest that a much larger number of species (from over 700 that constitute the human oral microbiome) likely play a role in the development of oral disease, highlighting that periodontitis has a complex, polymicrobial etiology (20, 78, 104, 105). Some species are potentially beneficial, contributing to a healthy oral homeostasis, while others are associated with various stages of gingivitis and periodontitis, potentially acting as keystone species and/or virulent components of the disease-inducing microbiota (100). Studies with pure cultures, defined communities *in vitro* and in animal models have revealed pro-inflammatory characteristics and interspecies signaling that result in community shifts and a host response leading to tissue destruction (106, 107). However, approximately one-third of the oral microbial species-level taxa are still uncultured, including dozens of disease-associated lineages (108). To better understand the complex interactions between oral bacteria and their role in the transition from health to disease, cultivation remains a critical initial step (108).

Based on rRNA relative sequence abundance, *Desulfobulbus* sp. HOT41 (Deltaproteobacteria) has been one of the uncultured bacteria strongly associated with advanced periodontitis (104, 105). Deltaproteobacteria have low taxonomic diversity in the human microbiome and other species have been linked, based on relative

abundance, to either oral disease (*Desulfomicrobium orale* and *Desulfovibrio fairfieldensis*) (109, 110) or to intestinal and gynecological disorders (*Desulfovibrio piger* and *Desulfovibrio intestinalis*) (111-113). As with other oral species, it has remained unclear if those organisms play a role in the etiology of periodontitis or their increased abundance is a consequence of the disease state, with deep subgingival pockets favoring the proliferation of strict anaerobes. No host-associated species of *Desulfobulbus* has been cultivated so far, although multiple free-living relatives have been isolated and characterized from aquatic and sediment samples (60-62, 65-67, 114). All free-living *Desulfobulbus* species are sulfate reducers, strict anaerobes, largely prototrophic, and use a variety of electron donors and acceptors. Adaptation to a host-associated lifestyle provides relative environmental stability and close interaction with other host-recruited species. These factors often lead to genome reduction, metabolic co-dependence/specialization, and sometimes emergence of pathogenicity through lateral gene transfer (114-116). Previous single-cell genomic data suggested *Desulfobulbus* sp. HOT41 is capable of sulfate reduction (33), which we used here to selectively enrich for, isolate, and characterize that organism as the first host-associated member of the *Desulfobulbus* genus. Its physiological, genomic, and immune modulatory characteristics provide new insights into adaptation to the human host-associated lifestyle and the emergence of pathogenicity.

Results and Discussion

Enrichment cultures and isolation of a human oral *Desulfobulbus* in pure culture. We established culture enrichments of sulfate-reducing bacteria using human oral subgingival samples collected from an individual with chronic periodontitis. Upon repeated passage in sulfate-lactate based medium we obtained a stable, simple community composed of *Fusobacterium nucleatum*, a common human oral bacterium, and *Desulfobulbus* sp. HOT41. The composition of this co-culture was characterized using small subunit rRNA amplicon sequencing (Illumina MiSeq), fluorescence *in situ* hybridization (FISH), and quantitative PCR (qPCR) revealing that *F. nucleatum* outnumbered *Desulfobulbus* sp. HOT41 by five to ten-fold in stationary phase (**Figure 2.1a,b**). During time course co-cultivation experiments, the growth of *Desulfobulbus* sp.

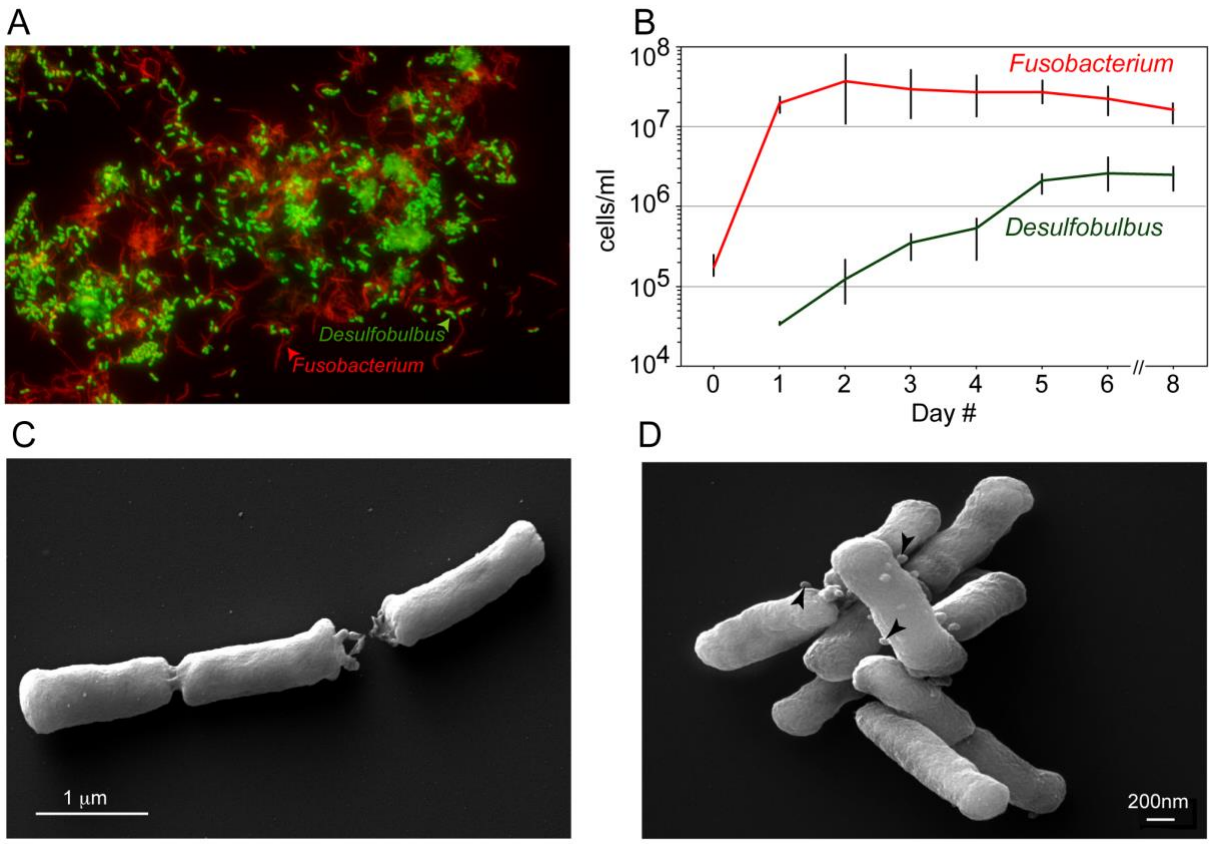


Figure 2.1. *Desulfobulbus* sp. strain HOT041 (*D. oralis*) in coculture with *Fusobacterium nucleatum* and in pure culture. **a.** FISH using fluorescent oligonucleotide probes specific for Deltaproteobacteria (green) and universal Bacteria (red). **b.** Growth of *Desulfobulbus* sp. strain HOT041 and *F. nucleatum* in coculture monitored by species-specific qPCR (with error bars based on three biological replicates). **(c and d).** Scanning electron micrographs of the *D. oralis* isolate. The arrowheads point to membrane vesicles.

HOT41 always exhibited a lag phase but growth continued for days after *F. nucleatum* reached stationary phase. As initial experiments to isolate *Desulfobulbus* sp. HOT41 in pure culture by serial dilutions were unsuccessful, we hypothesized that the organism required metabolites produced by *F. nucleatum* in the co-culture and, because microscopic analysis did not reveal a stable physical interaction between the two organisms, that *Desulfobulbus* sp. HOT41 could acquire them directly from the co-culture medium. Such dependence on metabolic co-factors produced by other community members, including *Fusobacterium*, has been described for other oral bacteria as well (39, 41, 117). We therefore supplemented fresh culture medium with filter-sterilized medium from stationary-phase co-cultures (cell-free spent medium, CFS) and, by applying three successive rounds of dilution to extinction, we obtained pure cultures of *Desulfobulbus* sp. HOT41. Purity was determined microscopically by FISH and by deep rRNA amplicon and genome sequencing. The isolate was named *Desulfobulbus oralis*, its SSU rRNA being 99.6% identical to *Desulfobulbus* sp. HOT41 clone R004 (118).

Physiological characterization of *D. oralis*. *D. oralis* is Gram-negative, non-motile and non-sporulating, with rod-shaped cells 1-2 µm in length and 0.3 µm in diameter. Cells usually occur singly or in pairs although sometimes form longer chains. In scanning electron micrographs we observed the presence of outer membrane vesicles (25-50 nm in diameter) on some cells (**Figure 2.1c,d**). Such vesicles, common in Gram-negative organisms including known oral pathobionts, have been shown to play important roles in interspecies interaction and to contain factors that trigger an inflammatory response (119, 120).

For routine cultivation, we grew *D. oralis* anaerobically at 37°C in a minimal defined liquid medium with lactate as the sole carbon and electron source and sulfate as the electron acceptor, supplemented with 5% (v/v) *F. nucleatum* CFS. No growth occurred in the presence of trace levels of oxygen or in the absence of *F. nucleatum* CFS. We found that the soluble, growth-promoting component(s) of CFS could not be inactivated by heating or protease treatment and could not be substituted by defined supplements such as vitamins, amino acids, volatile fatty acids, siderophores or complex nutrients (yeast extract, peptone, brain-heart infusion), individually or in combinations. Therefore, the nature of the critical component(s) provided to *D. oralis* by *F. nucleatum* remain unknown

but such requirement has been documented for other human oral bacteria as well, including the *Chloroflexi* sp. HOT-439 (41) and *Fretibacterium fastidiosum* (39). Within the oral cavity, this type of interspecies dependence has become evident and is likely the main reason for recalcitrance to cultivation of many yet-uncultured species that, likewise, have lost the ability to synthesize a number of compounds and therefore must rely on other members of the microbiota (108, 121). CFS from a type strain of *F. nucleatum* (ATCC 23726) also supported the growth of *D. oralis*, indicating its physiological dependence is not restricted to the strain it co-enriched with, although we have not extended that analysis to additional strains or species.

D. oralis is quite limited in the types of electron donors and acceptors it can utilize as an energy source for respiratory growth. Lactate and pyruvate are the only electron donors capable of being respired, while only sulfate or thiosulfate served as electron acceptors, out of all combinations tested (**Table A2.1**) (all additional material appears in the Chapter Appendix). Lactate was utilized with stoichiometric (1:1) production of acetate as the metabolic byproduct (**Figure 2.2a**) and the release of sulfide. Respiratory growth on pyruvate was very weak. We also tested *D. oralis* for its ability to ferment organic substrates in the absence of an electron acceptor (sulfate respiration). Growth was supported by pyruvate and weakly by amino acids but not by lactate, glucose, mucin, or peptone. Pyruvate was fermented to acetate and propionate in a 2:1 ratio (**Figure 2.2b**). In the subgingival space, sulfate could originate both from the blood and serum, where its concentration is relatively low (0.15mM) (122), but also a result of the connective tissue and bone breakdown during periodontal disease progression (122, 123). Lactate and pyruvate are released by other members of the oral microbiota and are important nutrient sources within the microbial oral biofilm (e.g. (124)). Therefore, we expect that *D. oralis* grows *in vivo* both by lactate-sulfate respiration and by pyruvate fermentation, depending on the local microenvironment and disease state. *D. oralis* expands therefore the list of confirmed sulfate reducers in the oral cavity that includes species of *Desulfovibrio* and *Desulfomicrobium*. Sulfate reducers have long been associated with periodontal disease and their abundance correlated with disease severity (125). The availability of nutrients and especially sulfate may be limiting factors in their proliferation, as these organisms are specialized users of that terminal electron acceptor (57).

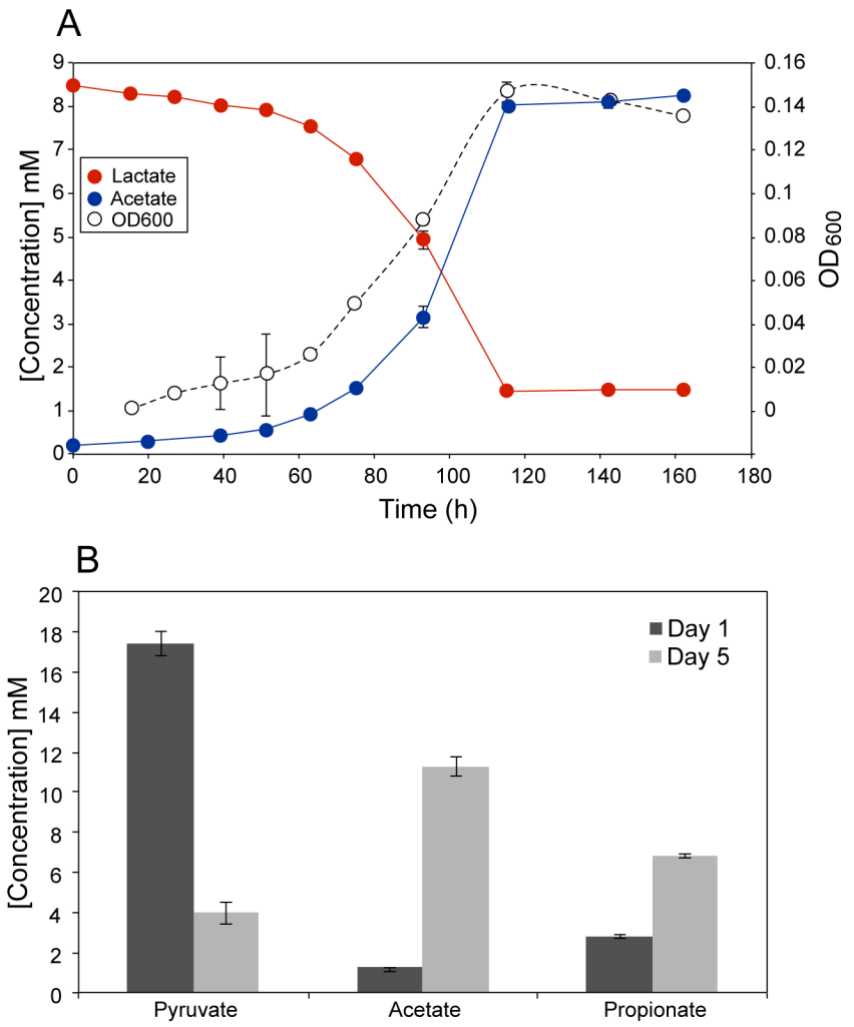


Figure 2.2. Growth of *D. oralis* by respiration and fermentation. **a.** Time course cultivation using sulfate and lactate as the electron donor-acceptor pair, with the release of acetate. **b.** Substrate utilization and metabolic by-products of pyruvate fermentation. The error bars are based on three culture replicates.

Comparative physiology and emergence of host association within the *Desulfobulbus* genus. Host-association leads to a reduced metabolic versatility, as the chemical and biological environment is relatively stable. This was evident as we compared the ability of several *Desulfobulbus* species to utilize various growth substrates (**Table A2.1**). All of the free-living isolates are able to utilize a wider range of substrates as electron donors and acceptors and for fermentation than *D. oralis*. Notably, propionate respiration and lactate fermentation, common traits among all free-living *Desulfobulbus*, have been lost in *D. oralis.* Within the *Desulfobulbus* genus, host association is not restricted to humans, as closely related uncultured species have been detected in the oral microbiota of domesticated dogs (126) and cats (127). We performed a phylogenetic analysis using full length SSU rRNA gene sequences, which suggested that mammalian host association was likely the result of an ancestral oral colonization by a free living *Desulfobulbus*, most closely related to *D. oligotrophicus* **(Figure 2.3)**. Additional sequences from other mammalian lineages as well as cultivation of other host-associated *Desulfobulbus* would be required to better understand the co-evolution of these bacteria with their hosts.

Genome sequencing, annotation, and metabolic reconstruction. We obtained the complete genome sequence of *D. oralis* using long read sequencing (PacBio). The assembly resulted in a single contig (>600 fold coverage), which was circularized as a single chromosome of 2,774,417 base pairs, with a G+C content of 60%. We further used mapping of Illumina short reads to correct any potential sequence and assembly errors. Gene prediction and annotation was performed in both Genbank and IMG and was supplemented with annotations based on Pfam, KEGG and eggNOG (128) (**Dataset S2.1**). The *D. oralis* genome encodes 2,395 proteins (GenBank), 51 tRNAs, and has 2 rRNA operons. A slightly larger number of proteins were predicted with IMG (2442) and we identified the differences as small hypothetical proteins and mobile genetic elements (transposases).

When compared to genomes of environmental *Desulfobulbus* species (**Table A2.1**), the genome of *D. oralis* is 30-50% smaller. There is also limited gene synteny based on comparison with the only closed genome of an environmental *Desulfobulbus* (*D. propionicus*) (**Figure 2.4**) suggesting numerous chromosomal rearrangements

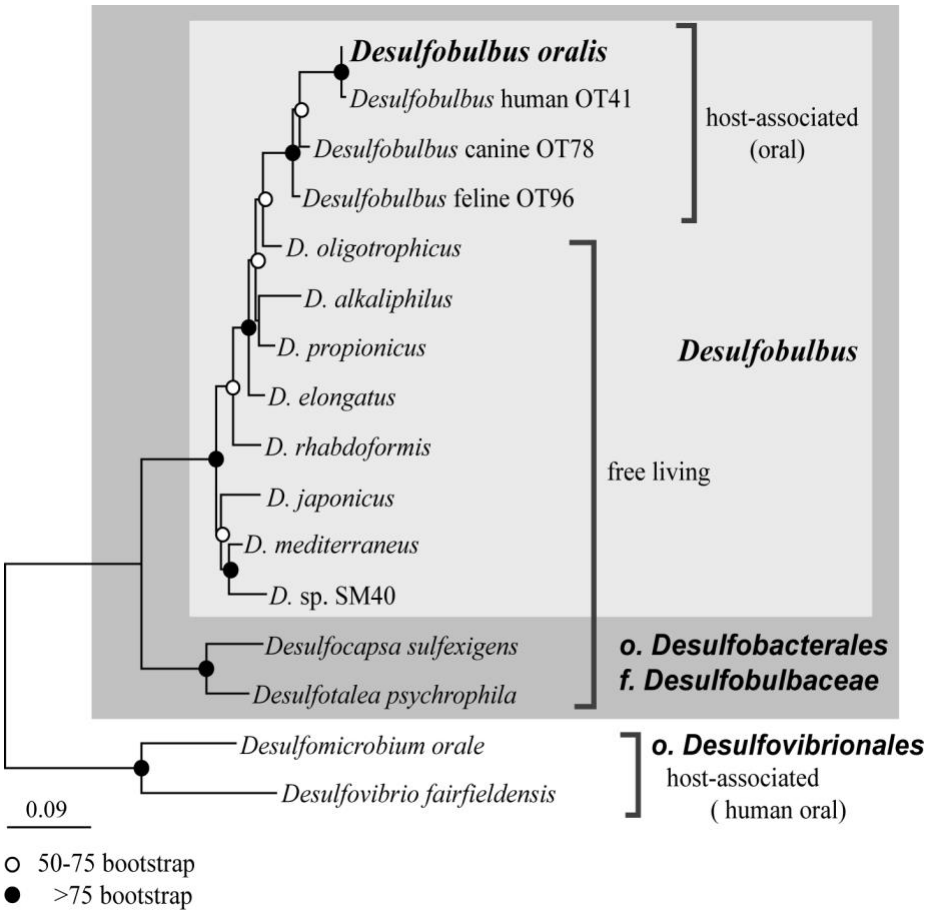


Figure 2.3. Maximum likelihood phylogeny of *Desulfobulbus* and related Deltaproteobacteria. The tree is based on full-length SSU rRNA gene sequences. Node support codes are based on 100 bootstrap replicates.

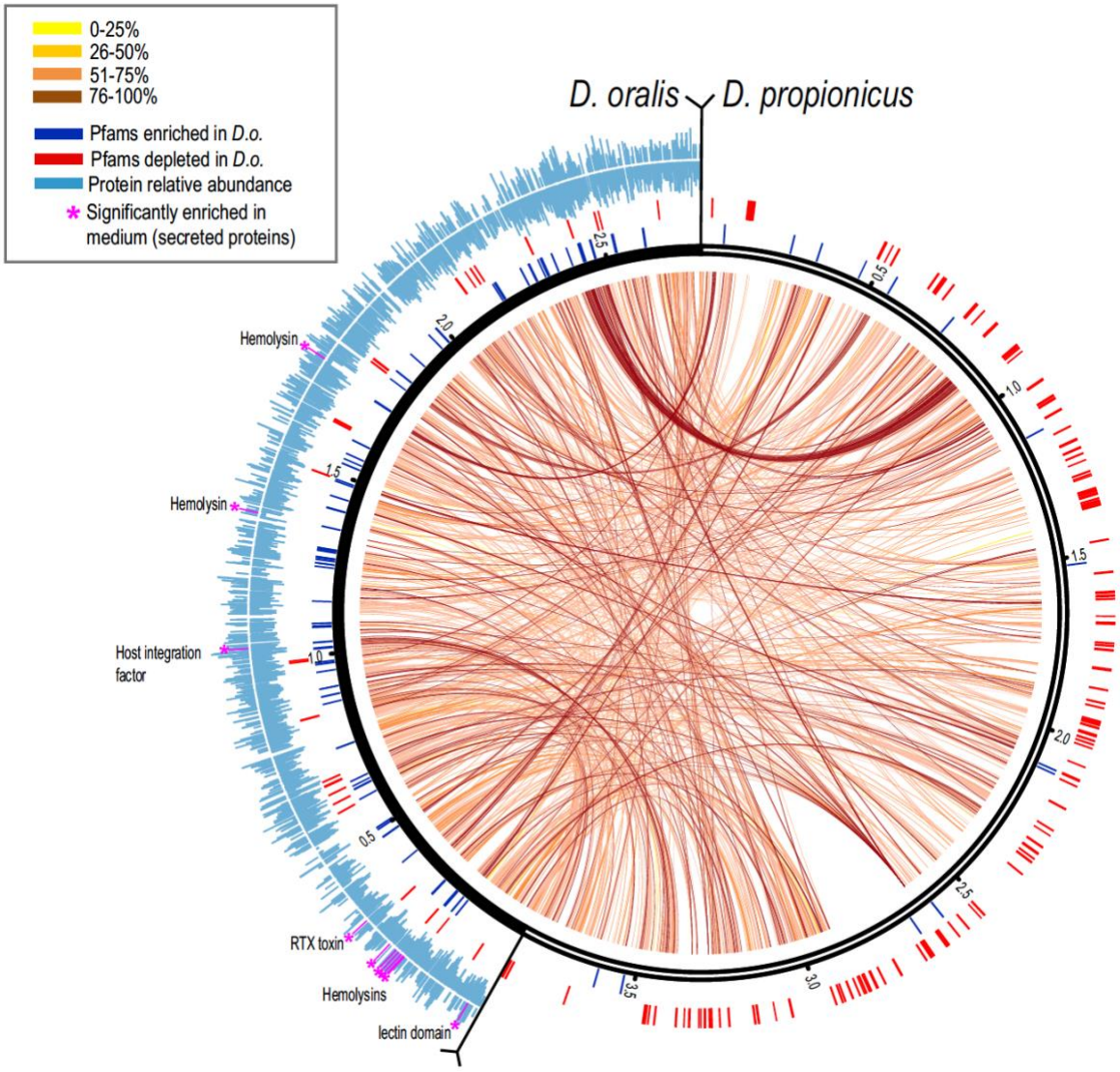


Figure 2.4. Genomic comparison of *D. oralis* and *D. propionicus*. Matching orthologs are connected by internal arc lines, the color scale indicating the degree of pairwise sequence identity at the protein level. External tick lines indicate genes in each genome that contain protein domains enriched (blue) or depleted (red) in *D. oralis* (based on data in Data Set S2). The outer histograms (light blue) indicate the relative abundance of each detected protein in *D. oralis* cells (bottom) or culture medium (top). Secreted proteins significantly enriched in the medium are identified and indicated by purple asterisks and lines.

associated with genomic fragmentation. For metabolic reconstruction, we combined assignments from IMG, KEGG's BlastKOALA, and MetaCyc, with similarity searches, domain architecture analyses, sequence alignments, and phylogenetic reconstruction for comparisons of enzymes and pathways with other sulfate reducers and oral bacteria.

As physiological characterization of *D. oralis* revealed streamlining and loss of some of the free-living *Desulfobulbus* respiratory and fermentative capabilities, the first objective in metabolic reconstruction was to predict the bioenergetics and carbon flow associated with respiratory versus fermentative growth in the oral environment. The dissimilatory sulfate reduction pathway (129), common to all *Desulfobulbus* species, could be reconstructed based on the *D. oralis* genome and is proposed to follow transformations represented schematically in **Figure 2.5**. Sulfate enters the cell via a SulP permease and is activated by adenylation via a sulfate adenylyl transferase (Sat). The resulting pyrophosphate is hydrolyzed by a manganese-dependent inorganic pyrophosphatase (ppaC), which drives the otherwise endergonic activation step. The adenosine 5'-phosphosulfate is converted to sulfite by APS reductase (Apr), a heterodimeric iron–sulphur flavoprotein complex, which receives electrons from a membrane-bound oxidoreductase complex (QmoABC) that interacts with the quinone pool as part of the energy conservation cycle (130). Sulfite is further reduced to sulfide by the dissimilatory sulfite reductase (Dsr), a large enzyme complex that includes cytoplasmic, membrane anchored and periplasmic subunits including cytochromes b and c, quinone interacting domains, and iron sulfur centers that mediate electron transfer and proton translocation (**Figure 2.5**). Another energy-conserving membrane complex present in *D. oralis* shares homology with the membrane-bound NADH dehydrogenase complex Nuo and is encoded by an eleven-gene cluster. However, the NuoEFG subunits, responsible for oxidizing NADH, are missing, so the complex may oxidize a different electron donor, such as reduced ferredoxin, similar to other sulfate reducers (131). The Nuo complex is complete in *D. propionicus*.

A cytoplasmic Hdr–Flx complex that can perform an alternative mode of energy conservation, flavin-based electron bifurcation (FBEB)(132) is present in *D. oralis* but absent in free living *Desulfobulbus*. The complex appears to have originated from a *Desulfovibrio* by horizontal gene transfer (HGT) and is most closely related to that of the

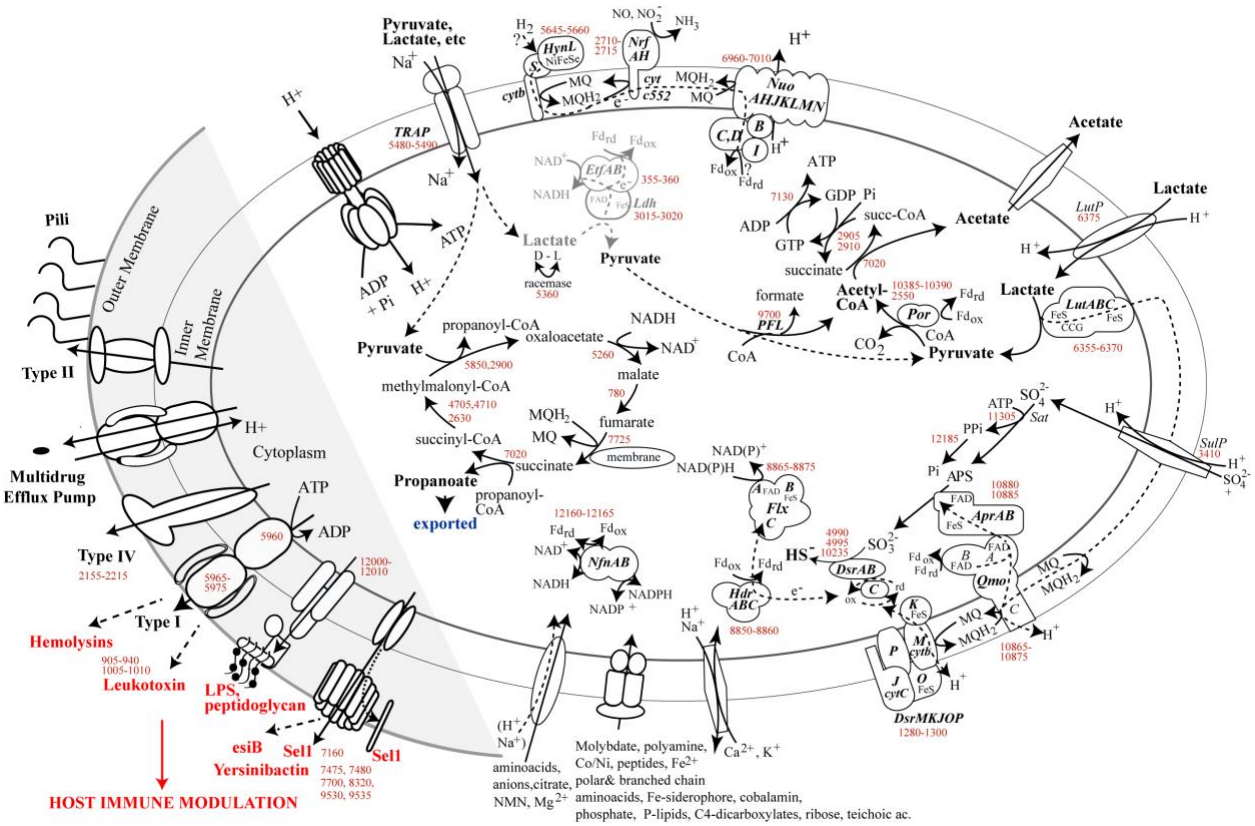


Figure 2.5. Metabolic reconstruction of *D. oralis*, with emphasis on membrane processes, energy conservation, and pathogenicity potential. Selected genes encoding enzymes, complexes, and other proteins proposed to function in such cellular processes are indicated by numbers in red, corresponding to GenBank genes in Data Set S1. The putative EtfAB-Ldh complex is depicted in gray.

human oral *D. fairfieldensis*. The FlxABC complex has FAD, NAD(P)-binding sites, and an iron-sulfur center, representing a distinct type of NAD(P)H dehydrogenase, which was proposed to oxidize NAD(P)H using both ferredoxin and a high redox disulfide center (DsrC) as electron acceptors (133). The reverse reaction, NAD+ recycling to NADH, can be coupled in environmental *Desulfovibrio* with aldehyde conversion to ethanol during pyruvate fermentation, using electrons coming from reduced ferredoxin through HdrABC (133). While we did not detect alcohol production by *D. oralis*, a related aldehyde detoxification process could occur through an NADPH-dependent alcohol dehydrogenase (CAY53_10125), most closely related to genes from human and animal-associated Actinobacteria. The interconversion between reduced and oxidized forms of NAD(H) and NADP(H) is performed by NfnAB, also an electron-bifurcating, ferredoxin-dependent flavoprotein. Another complex with a role in detoxification is the cytochrome c-containing nitrite reductase, NrfAH. This membrane complex, present in pathogenic bacteria, detoxifies nitric oxide produced by host cells as a defense mechanism against bacterial invasion (134).

Oxidation of lactate to pyruvate during sulfate respiration is catalyzed by a peripheral membrane associated multisubunit lactate dehydrogenase (LutABC) that uses the menaquinone pool as the electron acceptor. The complex is encoded in the same operon with a lactate permease (LutP) and appears to have been acquired from host associated *Desulfovibrio*, lacking close homologues in environmental *Desulfobulbus*. Lactate is an important shared substrate among many oral bacteria and the acquisition of the Lut system by *D. oralis* likely reflects an important adaptation to this environment following colonization. Pyruvate is further decarboxylated to acetyl-CoA by a pyruvate:ferredoxin oxidoreductase, generating reduced ferredoxin. *D. oralis* has lost both phosphate acetyl transferase and acetate kinase and we hypothesize that acetate release occurs via CoA transfer by a succinate:acetate-CoA transferase, that was likely acquired from a host-associated *Campylobacter* (**Figure A2.1**), followed by trans-phosphorylation by succinate thiokinase and nucleoside diphosphate kinase (**Figure 2.5**). This pathway is consistent with the observed 1:1 stoichiometry of lactate conversion to acetate (**Figure 2.2a**). It would result in the production of one ATP per lactate via substrate level phosphorylation and an additional amount of ATP via sulfate respiration

that depends on the number of protons translocated across the cytoplasmic membrane and the coupling number of the F0F1 ATPase.

For free-living *Desulfobulbus*, lactate can also be fermented. An electron-bifurcating, cytoplasmic lactate dehydrogenase-flavoprotein complex (Ldh-Etf) oxidizes lactate to pyruvate by coupling the endergonic NAD^+-dependent oxidation of lactate to the exergonic electron transfer from reduced ferredoxin to NAD+ resulting in the oxidation of one ferredoxin and the production of 2 NADH (135). For every three pyruvate produced from lactate, one is converted to acetate, generating reduced ferredoxin and an ATP by substrate level phosphorylation, while two more pyruvate molecules are reduced to propionate. This would result in an expected product ratio of 2:1 propionate to acetate (136). Even though we did not detect fermentative growth on lactate, *D. oralis* has a potential Ldh-Etf complex encoded in its genome. It is therefore possible that it may ferment lactate *in vivo*, and that the inability to do so under pure culture conditions is due to insufficient protein expression. In fact, based on whole cell proteomic analysis (detailed below) we found that the Ldh complex subunit was present at extremely low levels (~0.001% of the proteome), two orders of magnitude below that of the Etf subunits. In the absence of lactate fermentation, *D. oralis* can nevertheless ferment pyruvate through the propanoate cycle. In this case, for every three pyruvate molecules, two can be converted to two acetate molecules with the generation of two ATP and two reduced ferredoxins, while one additional pyruvate serves as an electron acceptor and is reduced to propionate. This results in the expected ratio of 1:2 propionate to acetate during pyruvate fermentation, consistent with our observed results (**Figure 2.2b**).

D. oralis has incomplete pathways for the biosynthesis of several amino acids (serine, threonine, cysteine, tyrosine, and phenylalanine) as well as for biotin, molybdenum cofactor, and cobalamin. Amino acid and vitamin auxotrophies are common among bacteria that live within complex communities such as the oral microbiota. However, *D. oralis* has additional, yet unidentified dependencies; simple supplementation with amino acids, peptides and/or vitamins did not relieve the requirement for the *Fusobacterium* culture filtrate. Nucleotide biosynthesis and conversion appears fully functional and *D. oralis* can also synthesize a variety of fatty acids and sugars for membrane, peptidoglycan, and lipid A assembly. Future elucidation of this nutritional

requirement may enable more rapid isolation of other fastidious microorganisms from human microbiomes.

Genomic fingerprints for adaptation to a host-associated environment and potential pathogenicity. *Desulfobulbus oralis* has a reduced genome which displays limited conserved gene synteny when compared to its closest sequenced free-living relative (**Figure 2.4**), as well as evidence for both gene loss and gene gain. Aside from becoming dependent on exogenous amino acids and vitamins, gene loss has affected hydrogenases, ferredoxin oxidoreductases, cytochromes, the dissimilatory nitrate reductase, and pyruvate kinase. This restricts *D. oralis* from utilizing a wider range of electron donors and acceptors and fermentation substrates that are not present at high concentrations in the oral environment (nitrate, short chain fatty acids, and alcohols). As very few organisms from the oral microbiota are capable of sulfate respiration, the restricted energy metabolism in *D. oralis* may reflect its adaptation to a metabolic niche in which there are few competitor species, with no selection to maintain metabolic versatility typical of free living *Desulfobulbus* species.

Following host colonization, cohabitation with bacteria not present in open environments presents the opportunity for horizontal gene transfer and acquisition of physiological traits that are absent in free-living species. Horizontal gene transfer is an important factor influencing the human oral microbiota and its individual species' positive and negative interactions with the host (137, 138). Based on sequence similarity and phylogenetic analyses, we found evidence for multiple horizontal gene transfer events. Sixty percent of the *D. oralis* genes have their closest homologs in the genomes of free-living *Desulfobulbus* and potentially can be regarded as inherited from the ancestral lineage that colonized the mammalian oral environment (**Dataset S2.1**). Approximately 200 predicted protein-encoding genes (~8% of the genome) have no distinguishable homologues, most of them being quite short and without recognizable protein domains. Out of the remaining genes, >250 (10%) are most closely related to homologues from other Deltaproteobacteria and have the oral *Desulfovibrio fairfieldensis* and *Desulfomicrobium orale* as the top hits, suggesting they might have been acquired from such physiologically-related lineages. Among those genes are clusters encoding for a variety of transporters and energy conservation complexes (lactate dehydrogenase,

glycerol-3-phosphate dehydrogenase, heterodisulfide reductase). Other potential horizontally acquired genes include transporters, drug and antibiotic efflux pumps, hydrolases, glycosyl transferases involved in cell membrane biosynthesis, CRISPR system proteins, and other enzymes and regulators from a wide range of oral bacteria, including *Actinomyces, Streptococcus, Porphyromonas, Tanerella, Prevotella, Neisseria,* and *Campylobacter* (**Figure A2.1**). As far as gene absence, it was notable to find that the *hgcA* and *hgcB* genes, present in all environmental *Desulfobulbus* identified so far (139), have been lost in *D. oralis*. Those genes encode a corrinoid iron-sulfur protein (HgcA) and a ferredoxin (HgcB) that are responsible for the microbial synthesis of methyl mercury, a potent neurotoxin (140). While the biological role of HgcA-B is still unknown, the finding that *D. oralis* has lost that biochemical potential represents a highly significant positive outcome of adaptation to the oral microbiota, especially considering mercury amalgam tooth fillings continue to be used and are present in many individuals.

To get a more quantitative estimate on gene family loss and gain in the *D. oralis* genome relative to its free-living relatives, we performed functional abundance profile analyses in IMG (**Dataset S2.2**). At the protein domain (Pfam) level, several categories of transposases and other mobile elements (integrases and endonucleases) are highly enriched in *D. oralis* as compared to other *Desulfobulbus*. Phylogenetically, many of those genes appear to be related to *Desulfomicrobium* and may reflect multiple genetic exchanges between these two related sulfate reducers in the oral microbiota. At the gene family level (COG), aside from mobile elements, COG families involved in membrane lipid biosynthesis and transporters are enriched. Considering that *D. oralis* relies on exogenous amino acids, it is not unexpected to see that amino acid and peptide ABC transporters are at higher abundance than in other *Desulfobulbus*, and some of the genes were likely acquired from other oral bacteria by HGT. Four genes (9825, 9840-9855) that encode members of an ankyrin repeat protein family that may be involved in the interaction with human cells (141), absent in free-living *Desulfobulbus*, were also apparently acquired from oral *Synergistetes* (**Figure A2.1**).

With respect to gene loss, the most pronounced difference between *D. oralis* and its free-living relatives were observed in the category of signal transduction and gene regulation (**Dataset S2.2**). Indeed, this functional category shows the steepest

dependence on the total number of genes in bacteria and archaea (142); however, *D. oralis* presents significant deviation from the general trend. While the number of genes in *D. oralis* is only 30% smaller than that in a free-living *D. propionicus*, the number of signal transduction genes in *D. oralis* is 4.5 times smaller: 72 genes (**Dataset S2.3**) versus 320 genes in *D. propionicus* (data from the MiST database (143)). Such deviations were proposed to reflect particular biological phenomena as well as environmental conditions in a microbial habitat (144). Thus, we explain a disproportional loss of signal transduction capabilities in *D. oralis* as a direct result of adaptation to a much more stable niche, where constant temperature, pH, and set of nutrients eliminate the need to monitor and respond to various environmental factors. The absence of swimming motility in *D. oralis* is yet another conspicuous example of adaptation to a new environment via gene loss. While free-living representatives of *Desulfobulbus* contain a full complement of flagellar genes and chemotaxis machinery (more than 50 genes), *D. oralis* lost all of these genes. However, it retained type IVa pili (TFP) that are essential for active translocation across solid surfaces, including PilT, a marker gene for TFP-based motility, as well as two regulatory chemosensory systems that belong to TFP class (**Dataset S2.4**). Thus, despite the loss of flagella and chemotaxis, *D. oralis* is capable of actively navigating toward beneficial microenvironments within the oral cavity.

As the relatively streamlined metabolism of *D. oralis* does not suggest it could provide the human cells with vitamins or other beneficial factors, we searched for potential genes that could mediate evading the host defense system and a potential role in pathogenesis. One such candidate gene family encodes eight proteins containing the Sel1 domain, characteristic of solenoid proteins involved in a variety of signal transduction processes (145). Three of the *D. oralis* Sel1 domain proteins are related to proteins that are involved in the bacterial interaction with human cells including evasion of lysosomal host defense mechanisms, such as the *Legionella* LpnE and LidL (146) and yet uncharacterized proteins from opportunistic pathogens like *Neisseria, Haemophilus* and *Kingella* (147-150) (**Figure A2.1**). Another relative of those proteins are EsiB (*Escherichia coli* secretory immunoglobulin A-binding protein), a pathogenic *E. coli* marker (151) confirmed to have host immune modulatory activity by interacting with secretory IgA, thus leading to inhibition of neutrophil activation and evasion of host defense mechanisms

(152, 153). This ultimately allows *E. coli* to avoid clearance by the host's immune system and permits further infection. A similar role could be employed by *D. oralis* to avoid negative effects of host immune responses in periodontal disease progression allowing for its proliferation and increase in abundance. *D. oralis* also encodes genes for leukotoxin/hemolysin production and export through type I secretion systems. Leukotoxins belong to the RTX toxin family, broadly distributed among Gram-negative bacteria and with diverse biological functions. Leukotoxin produced by the periodontal pathobiont *Aggregatibacter actinomycetemcomitans* are lethal to human monocytes and human T lymphocytes *in vitro* (154, 155) and are a main virulence factor within deep periodontal pockets in aggressive periodontitis (156-158). Genes encoding other RTX family homologues include adenylate cyclase/hemolysin, which is the major virulence factor of *Bordetella* (159-161), as well as an outer membrane translocation complex involved in RTX protein processing and export (157). Based on these findings, we propose that leukotoxin/hemolysin systems of *D. oralis*, acquired as three distinct operons (**Dataset S2.5**), could play roles in the putative pathogenicity of *D. oralis* by contributing to host immune evasion, manipulation, and neutralization of defense mechanisms.

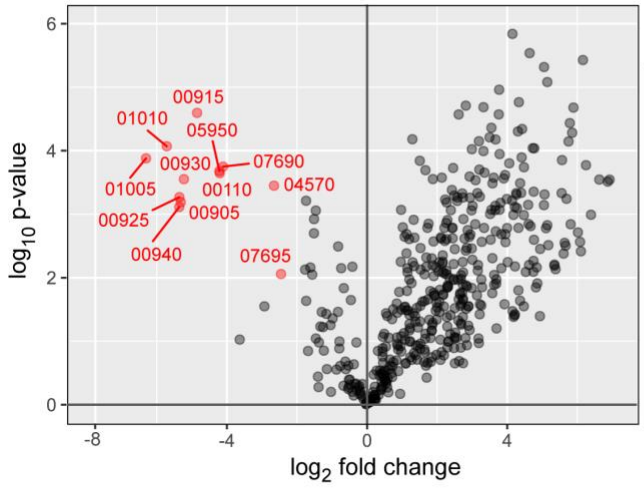
Proteomic analysis of *D. oralis* and its pathogenicity potential. While genome analysis revealed both native and acquired physiological potential, including the possibility that *D. oralis* may play an active role in the etiology of periodontal disease, simple gene presence is not sufficient evidence to assume that every potential gene product is actually expressed. We therefore applied a proteomic approach to determine the relative amounts of every expressed protein within the cell, and whether specific proteins are released into the extracellular space. Because the main focus was on proteins that may contribute to pathogenesis, some of which were predicted to be membrane- anchored while others secreted into the extracellular space, we collected both cells and filtered culture supernatant. Three separate culture replicates were performed and the samples for analyses were collected in late log phase.

One of the primary metabolic byproducts of *D. oralis* sulfate respiration, hydrogen sulfide, is well-known to trigger an inflammatory response at the leukocyte-endothelium interface, acting as a primary signaling molecule and synergizing with the pro-inflammatory action of virulence factors (162, 163). As we wanted to test the synthesis

and potential pro-inflammatory action of potential virulence factors encoded by *D. oralis*, we grew the cells fermentatively, in the absence of sulfate. The whole-cell proteome and the soluble proteins released into the medium were fragmented and analyzed by bottom-up, two-dimensional liquid chromatography tandem mass spectrometry (2D-LC-MS/MS). A total of 1628 proteins were detected, representing 70% of the predicted proteome (**Datasets S2.1 and S2.6**), which is similar to coverage that has been previously obtained on other organisms with reduced genomes (164, 165). Considering that proteins containing multiple transmembrane domains are difficult to detect in tryptic digests (here we detected only 40% of the proteins predicted to contain more than 1 transmembrane domain) we conclude that a very large fraction of the *D. oralis* genetic potential (>80%) was expressed under the tested condition. We also detected most of the putative virulence factors (hemolysins, leukotoxin, Sel1 proteins) and the proteins involved in their processing and export. In terms of dynamic range, we detected proteins spanning five orders of magnitude in relative abundance, from 2% of the cell total cellular proteome (sulfite reductase) to 2e-5% (a type IV transporter subunit). The most abundant proteins were determined to be enzymes involved in energy conversion and central metabolism, protein translation, chaperones, regulators, and several transporters. While the high-level of proteins involved in respiratory growth may appear surprising, it has been shown in other sulfate reducers grown under fermentative conditions that the sulfate reduction pathway is constitutively expressed (166). This indicates that the sulfate reduction capacity of *D. oralis* is actively maintained and, when periodontal sulfate levels are adequate, it may be the primary energy source.

To identify if any of the putatively expressed virulence factors were secreted, we compared the proteome of whole cells against their secretome. We identified 12 proteins that were significantly enriched in the culture medium relative to their abundance in whole cells (**Figures 2.4 and 2.6, Dataset S2.6**). With the exception of two, a DNA binding factor and an unknown protein, all others were predicted hemolysin family proteins and were enriched between 40-100 fold in the culture medium, suggesting that the cells actively secreted them. *In vivo*, such secreted hemolysins could diffuse through the epithelial layer, causing cytotoxicity, lysis, and modulation of the host immune response.

a



Protein	Closest relative	HOMD blast	env. Dsb.
00110 Hemolysin	Rhodobacter capsulatus	5e-33	1.4 e-24
00905 Haemolysin	Campylobacter ureolyticus	7e-60	none
00915 Haemolysin	Neisseria. oral taxon 014	4.e-11	none
00925 Haemolysin	Campylobacter ureolyticus	1.7e-81	none
00930 Haemolysin	Ottowia oral taxon 894	1e-64	2e-18
00940 Haemolysin	Ottowia oral taxon 894	4.4e-94	7 e-24
01005 rtx toxin RtxA	Neisseria meningitidis	1.7e-41	1.7 e-27
01010 Hemolysin	Abiotrophia defectiva	4.2e-89	1.4 e-14
04570 integration factor	Desulfobulbus elongatus	2.3e-48	2.3e-48
05950 Hemolysin	unc. bacterium EKD36112	3e-162	4e-26
07690 Hemolysin	unc. bacterium EKD36112	3.8e-162	3.4e-26
07695 unknown	no homologues	none	none

b

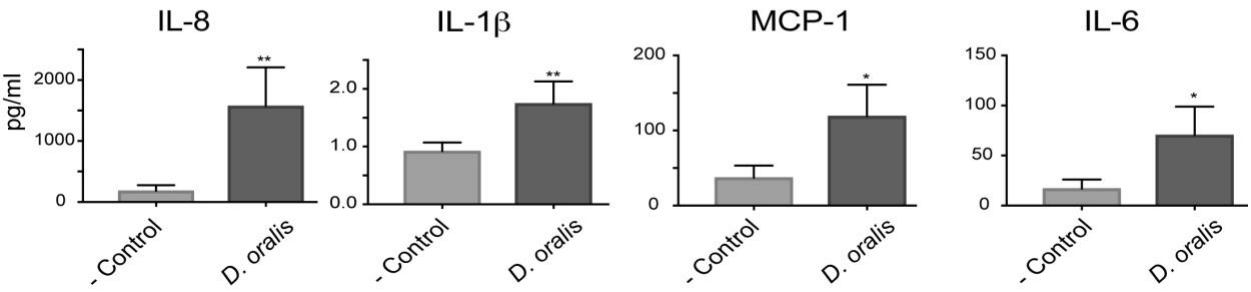


Figure 2.6. Proteomics and cytokine stimulation implicating *D. oralis* in proinflammatory responses. **a.** *D. oralis* proteins that are secreted from the cell (indicated in red) based on enrichment analysis. Only proteins that had at least a 4-fold enrichment and P<0.01 (based on three biological replicates) were considered. The table on the right identifies those secreted proteins and compares them to their closest homologues in the human oral microbiota or the free-living *Desulfobulbus*. **b.** Cytokine stimulation in oral epithelial keratinocytes by *D. oralis* cells. Keratinocytes were incubated with *D. oralis* cells or with the medium control for 48h, and cytokine stimulation was measured using a multiplex bead assay. Each bar represents the mean +/- standard deviation of the mean from triplicate measurements. Asterisks indicate results significantly different from the keratinocyte medium-only control: *, P<0.05; **, P<0.01.

***D. oralis* cytokine stimulation in oral epithelial keratinocytes.** During the initial stages of periodontal disease there is an acute inflammatory response to some of the oral microbiota resulting in the recruitment of polymorphonuclear leukocytes stimulated by host chemokines and/or cytokines, such as IL-8 (167). As inflammation progresses from acute to chronic, an increase in the production of pro-inflammatory mediators such as IL-1β, IL-6, IL-8, and TNF-α by various host cells, including oral epithelial keratinocytes, results in the recruitment of additional immune cells such as macrophages, T-cells, and B-cells (167). Various recognized periodontal pathobionts trigger distinct types of inflammatory responses (168-170) and some (e.g. *Porphyromonas gingivalis*) can modulate that by interfering with chemokine production (171).

Based on the finding that *D. oralis* secretes and has surface associated proteins (e.g., leukotoxin, hemolysins, esiB-Sel1) that have been shown in other oral bacteria to be linked to an inflammatory response in periodontitis, we analyzed its ability to induce pro-inflammatory cytokine production *in vitro*. We used a human immortalized oral keratinocyte cell line (OKF6/TERT2), which has been previously used as a model to study the interaction of oral bacteria with human epithelial cells (169, 172). Cells and filtered medium from the same *D. oralis* cultures used for proteomic analyses were applied to epithelial cells in culture and the production of 13 human inflammatory cytokines/chemokines was measured after 48 hours using bead-based immunoassays. We detected a significant stimulation of IL-1β release by the keratinocytes exposed to *D. oralis* culture filtrate while IL-1β, IFN-α, IFN-γ, MCP-1, IL-6, IL-8, and IL-18 were stimulated when keratinocytes were exposed to *D. oralis* cells (**Figure 2.6 and Figures A2.2 and A2.3**). The difference in pro-inflammatory responses between culture medium and whole cells are due to protein concentration differences, as the enriched proteins in the medium (secreted leukotoxins/hemolysins) and the bacterial cell surface agonists (e.g. Sel1 proteins, surface pili, LPS) may activate different pathways. The strongest stimulation was observed for IL-8 ($p<0.01$). IL-8 is one of the most important chemokines associated with periodontitis, being elevated at the disease site and attracting neutrophil polymorphs that further increase pro-inflammatory mediators (167). The observed release of the other cytokines/chemokines is expected to stimulate B- and T-cell differentiation as well as attracting macrophages and natural killer cells to the site of

infection. The absence of stimulation of IL-10 and IL-12, both anti-inflammatory cytokines, does not support a potential beneficial role of *D. oralis*. On the other hand, the complex response triggered by *D. oralis* in epithelial cells indicates that this bacterium is not merely a colonizer in the low oxygen deep periodontal pockets but actively participates in inflammation. During periodontal disease progression the gingival extracellular matrix breaks down and glycosaminoglycans (GAGs) are released (123, 173). *D. oralis* encodes several sulfatases that could liberate inorganic sulfate from GAGs potentially playing an active role in matrix destruction (174). In addition, the hydrogen sulfide produced by *D. oralis* during sulfate respiration can lead to apoptosis in keratinocytes (163, 175) and further elicits a pro-inflammatory response in oral cells (176, 177). These features illustrate the capacity of *D. oralis* to create a dysbiotic oral environment and actively contribute to the pathogenesis of periodontal disease in a novel and compounded way, not described for other oral pathogens so far. These results emphasize that cultivation of novel bacteria, only associated with disease based on sequence data, is important to begin to fully understand their role in the oral cavity.

Expression of *D. oralis* pathogenicity genes in periodontitis. A recent clinical metatranscriptomic study identified Deltaproteobacteria as one of the bacterial lineages that increase in prevalence and transcriptional activity in periodontitis (178). We therefore aimed at determining the functional profile of *D. oralis* and specifically the relative abundance of its pathogenicity gene transcripts, based on this RNAseq data generated from oral clinical samples. We determined that in the datasets from three individuals with periodontitis, the sequences that mapped to *D. oralis* protein encoding genes (~3000-25,000 reads) represent 0.1-0.5% of the total number of microbial reads. Even though the overall coverage of the *D. oralis* transcriptome by the RNAseq datasets was shallow (<two fold), across the three datasets >80% of the predicted gene transcripts were detected (**Dataset S2.1**), with ~65% of the genes represented by at least two reads. This correlates with the overall number of proteins we detected in *D. oralis* grown in pure laboratory cultures and supports the hypothesis that most genes in this bacterium are expressed both *in vitro* and in the natural subgingival environment. While the coverage level was too low to enable a robust global comparison of the transcriptional versus proteome profile, we analyzed the relative levels of mRNAs encoding putative pro-inflammatory response-

inducing proteins. Several of the hemolysin-type genes were amongst the top 50 most abundant transcripts (0.2-1% of the mapped reads), with four of them being part of a highly expressed operon (CAY53_915-940) (**Dataset S2.5**). This indicates that such factors, which have been linked to inflammatory response in other oral bacteria, are being actively produced by *D. oralis* in the subgingival space and likely play a role in the development and progression of periodontal disease.

Description of *Desulfobulbus oralis* sp. nov. *Desulfobulbus oralis* sp. nov. (o.ra'lis. N.L. masc. adj. *oralis* of the mouth). Cells are non-motile, non-spore-forming Gram-negative rods 1-2 µm in length occurring singly, in pairs, and sometimes chains. Growth occurred at 37°C with optimum pH 7-7.1 under strictly anaerobic conditions, with cysteine as a reducing agent. *D. oralis* incompletely oxidizes lactate and pyruvate to acetate, using sulfate or thiosulfate as electron acceptors and releasing sulfide. The organism does not utilize citrate, propionate, acetate, hydrogen, formate, succinate, glucose, fructose, ethanol, propanol, fumarate or malate as electron donors or sulfite and nitrate as acceptors. In the absence of electron acceptors, pyruvate and amino acids are fermented while lactate, peptone, mucin, and dextrose are not. Growth requires the addition of 5% *Fusobacterium nucleatum* cell-free spent medium (CFS).

D. oralis type strain ORNL was isolated from a human subgingival sample from a patient with periodontitis at Ohio State University College of Dentistry, Columbus, Ohio. The genome is represented by a circular chromosome of 2.77 Mbp and with G+C content of 59.76 mol%.

Materials and Methods

Clinical sample collection. Subgingival biofilm and crevicular fluid was obtained from a patient with periodontitis at the Ohio State University College of Dentistry. Informed consent was obtained, under a protocol approved by the Ohio State University IRB (2007H0064). Following removal of supergingival plaque with gauze, sterile endodontic paper points were inserted into the gingival sulcus and left for one minute. Afterwards they were placed into Liquid Dental Transport Medium (Anaerobe Systems, Morgan Hill, CA) and used for inoculation into primary enrichment medium.

Microbial cultivation. Primary oral enrichments were passaged (2% v/v) in an ATCC 1249 modified Baar's medium containing (per liter): 2g $MgSO_4 \cdot 7H_2O$, 1g $(NH_4)_2SO_4$, 0.5g K_2HPO_4, 1g $CaCl_2 \cdot 2H_2O$, 5g sodium citrate, 1.17g sodium propionate, 1.17g sodium acetate, 1.5mL L(+)-lactic acid (85% w/v, Sigma), 1g yeast extract, and 0.5 mg resazurin. The medium was degassed with N_2/CO_2 (80/20, v/v), sterilized for 15 minutes at 121°C followed by addition of sterile degassed cysteine (1mM final concentration) and 10mL $Fe(NH_4)_2(SO_4)_2$ 5% (w/v) and pH adjustment to 7.2 with 8M NaOH. Cultures were maintained in 15 mL glass tubes fitted with butyl rubber stoppers, under N_2/CO_2 (80/20, v/v) at 37°C for 1-2 weeks. Cell growth resulted in the release of hydrogen sulfide, which formed a black iron sulfide precipitate and was also detected using lead acetate indicator paper. Stocks were preserved by freezing at -80°C, after adding DMSO to 10% (v/v). Cell-free spent medium supernatant (CFS) was prepared by filtering cultures through two 0.22µm Nalgene™ Rapid-Flow™ sterile filters and degassing with N_2 (100%). Pure cultures were obtained by three successive dilutions to extinction transfers in liquid medium. For culture enrichments of *D. oralis*, the medium was supplemented with 25% (v/v) co-culture CFS. Pure cultures of *D. oralis* were grown in a modified Lactate-Baar basal medium containing (L^{-1} distilled water): 2g $MgSO_4 \cdot 7H_2O$, 1g $(NH_4)_2SO_4$, 10mM MOPS, 10mM HEPES, 10mM sodium lactate, 0.1% yeast extract, 2.5mM cysteine, and 0.5 mg resazurin 0.5%. The pH of the media was adjusted to 7-7.1 with 8M NaOH, degassed with N_2 (100%) and autoclaved for 30 minutes at 121°C. After cooling, sterile K_2HPO_4 (0.5M, pH 7) was added to 3mM final and CFS of *F. nucleatum* grown in brain-heart-infusion broth (BHI) supplemented with 0.2% glucose was added to 5% (v/v). In subsequent experiments we determined that the addition of yeast extract was not necessary for growth.

Cellular and molecular analyses. Fluorescence *in situ* hybridization (FISH) with the oligonucleotide probes EUB338 (Bacteria) and Delta495a (Deltaproteobacteria) was performed as previously described (33). For DNA isolation we used proteinase K digestion followed by phenol-chloroform extraction (164). PCR-amplification of the 16S rRNA gene was carried out using the universal primer set 27F (5'-AGAGTTTGATYMTGGCTCAG-3') and 1492R (5'-TACGGYTACCTTGTTACGACTT-3') or a specific *D. oralis* set DsbF (5'-AGTTAGCCGGTGCTTCCT-3') and 1492R.

Amplification with Lucigen 98HotStart Supermix used the following thermal profile: 94°C for 3 minutes, followed by 32 cycles of 94°C for 20 seconds, 60°C for 30 seconds, 72°C for 90 seconds, and a final extension of 72°C for 5 minutes. Following purification, the PCR products were sequenced directly using Sanger chemistry (ABI 3730) or were first cloned into pCR™2.1-TOPO (Invitrogen, Carlsbad CA, USA). Sequence data was analyzed using Geneious v.10 (179) and confirmed the *in silico*-inferred primer specificity. To quantitatively determine the co-culture composition, we developed a qPCR assay using 16S rRNA gene primers specific for *F. nucleatum* (5'-CAGCGTTTGACATCTTAGGA-3' and 5'-ATCGTAGGCAGTATCGCAT-3') and for *D. oralis* (5'-GGTGGTGCCTTCTTTGAA-3' and 5'-TTCCTCCGGTTTGAC ACC-3'). Amplification was with BioRad iQ SYBR Green Supermix in a CFX96™ Real-Time C1000 thermal cycler (Bio-Rad, Hercules, CA), using a plasmid-cloned 16S rDNA insert and the following thermal profile: 95°C for 3 minutes, followed by 40 cycles of 95°C for 20 seconds, 66°C for 30 seconds. Sanger sequencing confirmed primer specificity for each species and primers pair.

For scanning electron microscopy cells from 20 ml of culture were collected onto a 0.1µm PVDF syringe filter, washed with phosphate buffered saline (PBS, pH 7), fixed with 3% glutaraldehyde in PBS for 1 hour at room temperature, followed by three 10 minutes wash steps with PBS. Fixed cells were resuspended in 2% osmium tetraoxide in PBS for 1 hour, collected by centrifugation at 10,000 x g for 5 minutes, and washed three times in water (10 minutes each). On the final wash, cells were deposited onto a 3X4mm silicon chip (Ted Palla, Inc, CA) for 15 minutes followed by dehydration in an ethanol series (50, 70, 95, 100%), 10 minutes each. The cells were dried in a Ladd critical-point dryer then gold-coated using an SPI sputter coater and examined on a Zeiss Auriga Focused Ion Beam Scanning Electron Microscope.

Physiological characterization. Optimum pH was determined in Lactate-Baar medium containing 5% *F. nucleatum* CFS, and 15mM MES, MOPS, and/or HEPES, with the pH adjusted between 5 and 8 at 0.5 unit intervals (MES for pH 5-6, MES + MOPS for pH 6.5, MOPS for pH 7, MOPS + HEPES for pH 7.5, and HEPES for pH 8). Growth was monitored for one week at 37°C by daily measurements of optical density at 600nm. Utilization of electron donors was determined in Lactate-Baar medium containing 5% *F.*

nucleatum CFS and each substrate at 10mM. Substrates tested include citrate, propionate, acetate, lactate, formate, fumarate, succinate, malate, pyruvate (all as sodium salts), glucose, fructose, ethanol, and propanol. Hydrogen utilization was determined in the presence of sodium acetate and H_2/CO_2 (80/20, v/v). Utilization of electron acceptors was determined with sodium lactate (10mM) in minimal sulfate-free medium (L^{-1} distilled water): 2g $MgCl_2$, 1g NH_4Cl, 1g yeast extract, 10mM MOPS, 10mM HEPES, 3 mM K_2HPO_4, 2.5mM cysteine, 0.5 mg resazurin and 5% *F. nucleatum* CFS and containing 20mM sodium sulfate, sulfite, thiosulfate or nitrate. Fermentative capabilities were determined in the same basal medium used for the electron acceptor test, however lacking addition of electron donor-acceptor pairs. Substrates tested for fermentation include sodium lactate (20mM), sodium pyruvate (20mM), peptone (0.1%), mucin (0.25%), synthetic amino acids supplements (Y1501, Sigma)(0.2%), or dextrose (0.2%).

F. nucleatum CFS susceptibility to inactivation was tested by treatment (2 hours at 37°C) with proteinase K (0.2 mg/mL) or incubation at 80°C for 14 hours prior to supplementation into Lactate-Baar medium at 5% v/v. For testing potential replacement of CFS with defined amino acids, vitamins, and siderophores, yeast synthetic dropout media (amino acids) were added to 0.2% and ATCC vitamins to 1% in Lactate-Baar media prior to inoculation. Siderophore mixture contained a 1:1:1 mix of pyoverdines, 2,3-dihydroxybenzoic acid, and ferrichrome and were added to a final concentration of 1μg/ml. The cultures were monitored by optical density and microscopic examination for two weeks. Very weak growers were transferred to fresh media and monitored for growth up to 1 month. Growth was considered positive if an increase in OD_{600} occurred after three consecutive passages. All experiments were performed in triplicate.

To determine metabolic byproducts of lactate respiration and pyruvate fermentation, soluble analytes were measured via high-performance liquid chromatography (HPLC) using a Waters Breeze 2 system (Waters Corp., Milford, MA, USA) equipped with a refractive index detector (model 2414) and an Aminex HPX-87H column (Bio-Rad Laboratories). Sulfuric acid (5mM) was used as the mobile phase at a flow rate of 0.6 mL/min. Peak areas and retention times were compared against known standards.

Genomic sequencing and analysis. *D. oralis* genomic DNA isolated using the phenol-chloroform method described (164) was used to construct of a large insert (20kb) library for sequencing on a PacBio RS II system (Institute for Genome Sciences, University of Maryland School of Medicine). We generated 154,149 reads, with N50 length of 20.2 kbp and median of 14 kbp, which were assembled into one polished contig 2,796,768 bp in length, with 668 mean coverage, using PacBio SMRT Portal. We also sequenced in-house a small insert *D. oralis* genomic library on an Illumina MiSeq platform. Approximately 4 million reads (2x250 nt) were mapped on the PacBio contig using Geneious v.10. (179). Individual indel discrepancies were resolved based on both read coverage and inspection of their effect of disruption/truncation of predicted ORFs in that region. Contig end overlaps were identified which resulted in circularization of the initial contig into one chromosome of 2,774,417 bp. The finished genome was submitted to GenBank for gene prediction and annotation using NCBI PGAP (accession number CP021255) and to JGI IMG (IMG Genome ID 2711768589). In addition to the annotations provided by the NCBI and IMG platforms, which included Pfam, COG, TIGRFam, KO, EC, secretion signals and transmembrane domains, we also analyzed the predicted proteins and their functions through MetaCyc, eggNOG and TransportDB. For metabolic reconstruction we combined the use of the IMG networks, KEGG-BlastKOALA, and MetaCyc-Pathway Tools with individual gene searches (blast and hmmsearch), sequence alignments, and phylogenetic reconstruction using homologues from related sulfate reducers, acetogens, or human microbiome bacteria. Sequence alignments were performed using Muscle and maximum likelihood phylogenetic reconstructions were based on PhyML or FastTree, all implemented in Geneious. Genomic alignments were displayed using Circos (180).

Pro-inflammatory cytokine secretion assays. Late log cultures of *D. oralis* grown under pyruvate fermentation condition (50 mL, three replicates) were chilled on ice, and spun by centrifugation (4°C 10,000 x g for 10 minutes). The cell pellets were flash-frozen at -80°C and the spent medium was passed through a 0.22µm Millipore filter and frozen at -20°C.

Immortalized human oral keratinocytes (OKF6/TERT-2) were grown on keratinocyte-SFM 1X medium (Invitrogen, Carlsbad, CA) supplemented with epidermal

growth factor (0.3 ng/mL), bovine pituitary extract (50 µg/mL), 1 IU/mL penicillin and 1 µg/mL streptomycin to reach ~50% confluence, in 12-well tissue culture plates (Corning Inc., Corning, NY, USA) pre-coated with 3 mg/mL collagen (Stemcell Technologies, Vancouver, Canada). To each well we added 100µl of (a) *D. oralis* filtered spent medium (100µl), (b) *D. oralis* cells re-suspended in keratinocyte medium, (c) fresh *D. oralis* medium, or (d) fresh keratinocyte medium (negative controls), each in triplicate for every triplicate *D. oralis* biological sample. Keratinocytes were incubated for 48 hours at 37°C, after which media samples were collected, clarified by centrifugation, and frozen in aliquots at -20°C. Cytokine concentrations were measured with the LEGENDplex™ human inflammation panel (13-plex)(BioLegend®, San Diego CA, USA) following manufacturer's instructions, using a Cytopeia Influx flow cytometer (BD, Franklin Lakes, NJ, USA). Statistical analysis was performed using GRAPHPAD PRISM Version 7. One-way analysis of variance (ANOVA nonparametric) with Tukey correction was used to investigate significant differences between independent groups of data. Statistical differences were considered significant at $p < 0.05$.

Whole cell and secreted proteome data acquisition. Samples from the same three replicate *D. oralis* cultures used for the pro-inflammatory assay were also used for proteomic analysis, conducted as detailed in (181). Cell pellets were resuspended in 4% sodium deoxycholate (SDC) in 100 mM ammonium bicarbonate (ABC) and ultrasonically disrupted. The crude protein extract was clarified by centrifugation, reduced with 10 mM dithiothreitol (DTT), and cysteine residues blocked with 30 mM iodoacetamide (IAA). The proteins were then collected on top of a 10 kDa cutoff spin column filter. *D. oralis* spent culture medium samples (~12 mL) were concentrated on a 5 kDa filter. Collected proteins were washed with ABC and digested with sequencing-grade trypsin (Sigma). Peptides were collected by centrifugation, acidified to 1% formic acid (FA) followed by extraction with ethyl acetate and concentrated. The peptide mixture was analyzed by automated, two-dimensional liquid chromatography tandem mass spectrometry (2D-LC-MS/MS) system (181) using Ultimate 3000 RS system in-line with Q Exactive Plus (QE+) mass spectrometer (ThermoFisher Scientific).

Proteomic analyses. All MS/MS spectra were searched against a constructed FASTA database containing all proteins predicted from the sequenced genome of

Desulfobulbus oralis, concatenated with common contaminants and reversed-decoy sequences using Myrimatch version 2.2 (182). Peptides were identified and proteins inferred using IDPicker version 3.1 (183) with an experimental false discovery rate (FDR) <1% at the peptide level. Peptide abundances were derived in IDPicker by extracting precursor intensities/area-under-the-curve from chromatograms with the following parameters: retention time of ±90 s and mass tolerance of ±5 ppm. Protein abundances were calculated by summing together intensities of all identified peptides and normalized by their respective protein lengths. Protein abundances were log2-transformed and median-centered across all samples in InfernoRDN version 1.1 (184).

To investigate enriched secreted proteins, we performed Student's t-tests between the cell pellet and the spent medium (supernatant) samples using Perseus software platform (185). We limited the comparison to those proteins observed in all supernatant experiments to better focus on proteins that could be secreted and/or induce inflammatory response. Proteins with missing values were imputed from a normal distribution (width = 0.3, shift = 1.8). A protein was considered as significantly enriched in the supernatant if it passed the threshold of Benjamini-Hochberg FDR < 0.01 and log2 fold-change greater than 2 in the comparison of supernatant versus cell pellet.

Metatranscriptomic data analysis. MiSeq data for 10 healthy subjects and 6 with periodontitis (178) was downloaded from MG-RAST (http://metagenomics.anl.gov/linkin.cgi?project=5148). The reads were mapped to the *D. oralis* genome with bowtie2 v. 2.2.6 (with parameters: --very-sensitive-local --no-unal -X 1000 --score-min G,20,28 --no-mixed --rg-id). The resulting sam files were converted and sorted with samtools 0.1.19, alignments were inspected with IGV 2.3.55, and mapping was quantitated with htseq-count v. 0.6.1p1. Code used is available at https://github.com/cliffbeall/Doralis-metatrans.

Funding Information and Disclaimers

This research was fully funded by grant R01DE024463 from the National Institute of Dental and Craniofacial Research of the U.S. National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Acknowledgements

We acknowledge the University of Tennessee Advanced Microscopy and Imaging Center for instrument use, and for scientific and technical assistance with scanning electron microscopy. We acknowledge the Genomics Resource Center of the University of Maryland Institute for Genome Science for PacBio sequencing and the Genomics Core at the University of Tennessee Knoxville for Sanger sequencing. We thank Steve Allman, Zamin Yang, Dawn Klingeman and Ann Wymore for assistance with flow cytometry, sequencing, and tissue culture. We also thank Drs. Jim Rheinwald, Anna Mandinova and Kristina Todorova for providing the OKF6/TERT-2 cell line and advising with its cultivation.

Chapter Two Appendix

Table A2.1. *D. oralis* physiological characteristics compared to related species of the *Desulfobulbus* genus.

Characteristic	*D. oralis*	*D. oligotrophicus*	*D. propionicus*	*D. elongatus*	*D. rhabdoformis*	*D. japonicus*
Source	Human subgingival sample	Municipal sewage sludge	Anoxic freshwater ditch	Industrial digester	Water-oil separations system	Estuarine sediment
Cell shape	Rod-shaped	Oval-rods	Lemon-shaped	Rod-shaped	Rod-shaped	Rod-shaped
Motility	−	+	−	+	−	+
Growth conditions:						
Temperature (°C)	37°C	35°C	39°C	35°C	31°C	35°C
pH	7.1	7.6	7.2	7.0	6.8-7.2	6.7
Utilization of electron donors:						
Lactate	+	+	+	+	+	+
Pyruvate	+	+	+	+	+	+
Propionate	−	+	+	+	+	+
Acetate	−	−	−	−	−	−
Malate	−	−	−	−	+	+
Ethanol	−	−	+	+	+	+
Propanol	−	−	+	+	+	+
Formate	−	−	−	−	−	+
Succinate	−	−	−	−	+	−
Fumarate	−	−	−	−	+	+
Glucose	−	−	−	−	ND	+
Fructose	−	−	−	−	ND	+
H_2 + CO_2 Acetate	−	−	+	+	+	+
Utilization of electron acceptors						
Sulfate	+	+	+	+	+	+
Thiosulfate	+	+	+	+	+	+
Sulfite	−	+	+	+	+	−
Nitrate	−	−	+	−	−	ND
Utilization of substrates in the absence of electron acceptors:						
Pyruvate	+	+	+	+	+	+
Lactate	−	+	+	+	+	+
Glucose	−	ND	ND	ND	ND	+
DNA G+C content (mol%)	59.8	51.7	58.9	59	50.6	48.6
Genome size (Mbp)	2.8	ND	3.8	3.9	ND	5.8
Genome status	complete	ND	complete	draft	ND	draft
Isolation Reference	This study	El Houari, 2017	Widdel, 1982	Samain, 1984	Lien, 1998	Suzuki, 2007

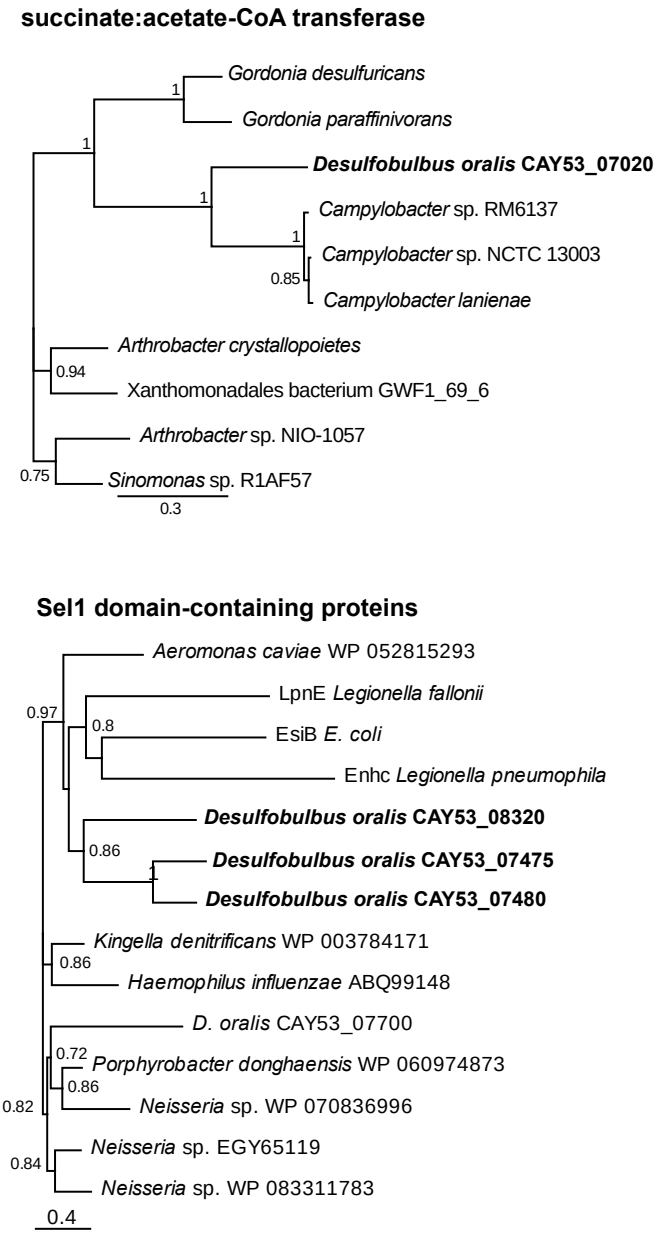


Figure A2.1. Phylogenetic analysis (FastTree) of selected horizontally-acquired genes in *D. oralis*. Closest sequence homologues from other bacteria (based on Blast) were used. For most genes proposed to have been horizontally acquired there were no close homologues in other *Desulfobulbus* genomes. The numbers indicate node support.

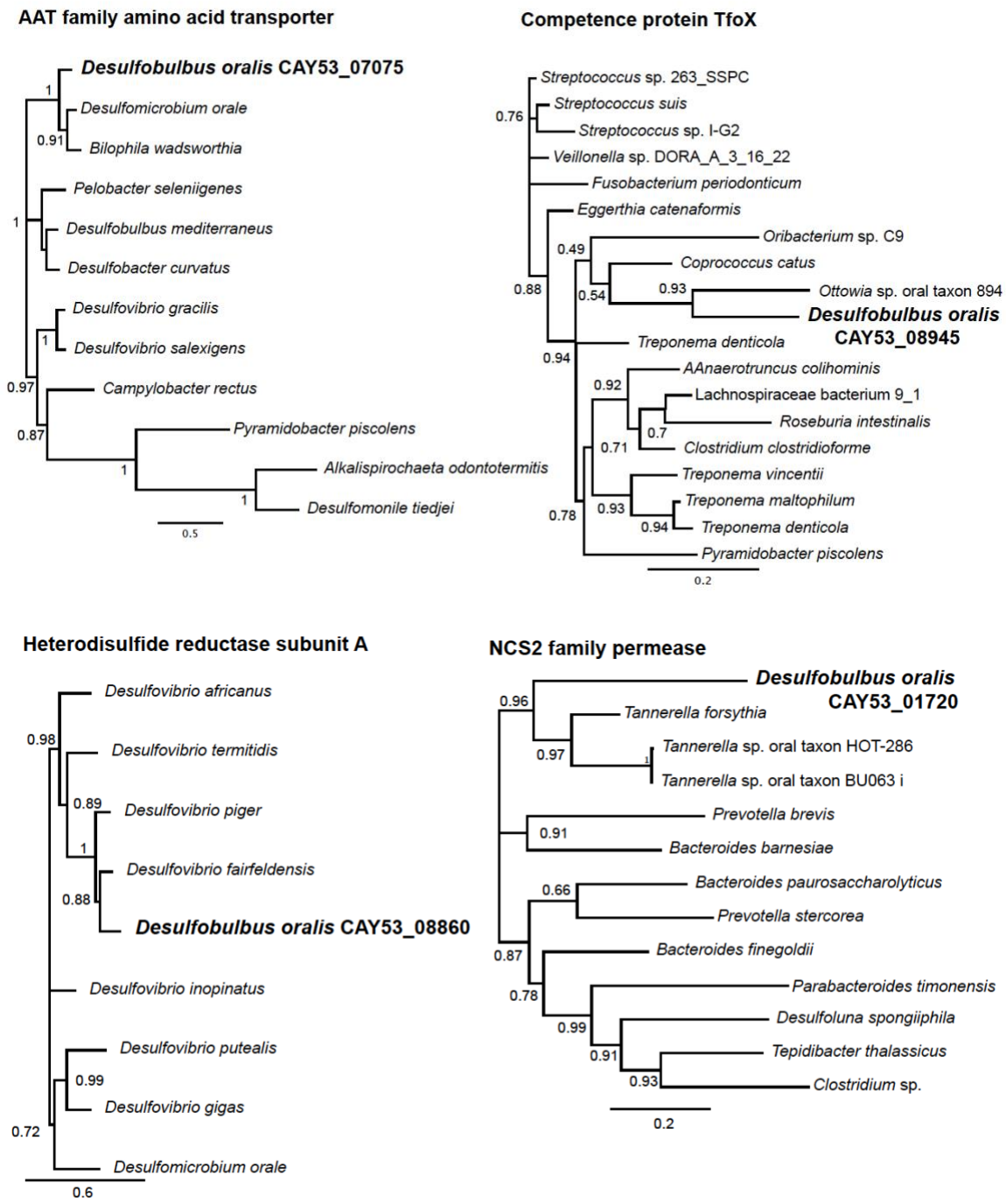


Figure A2.1. (continued)

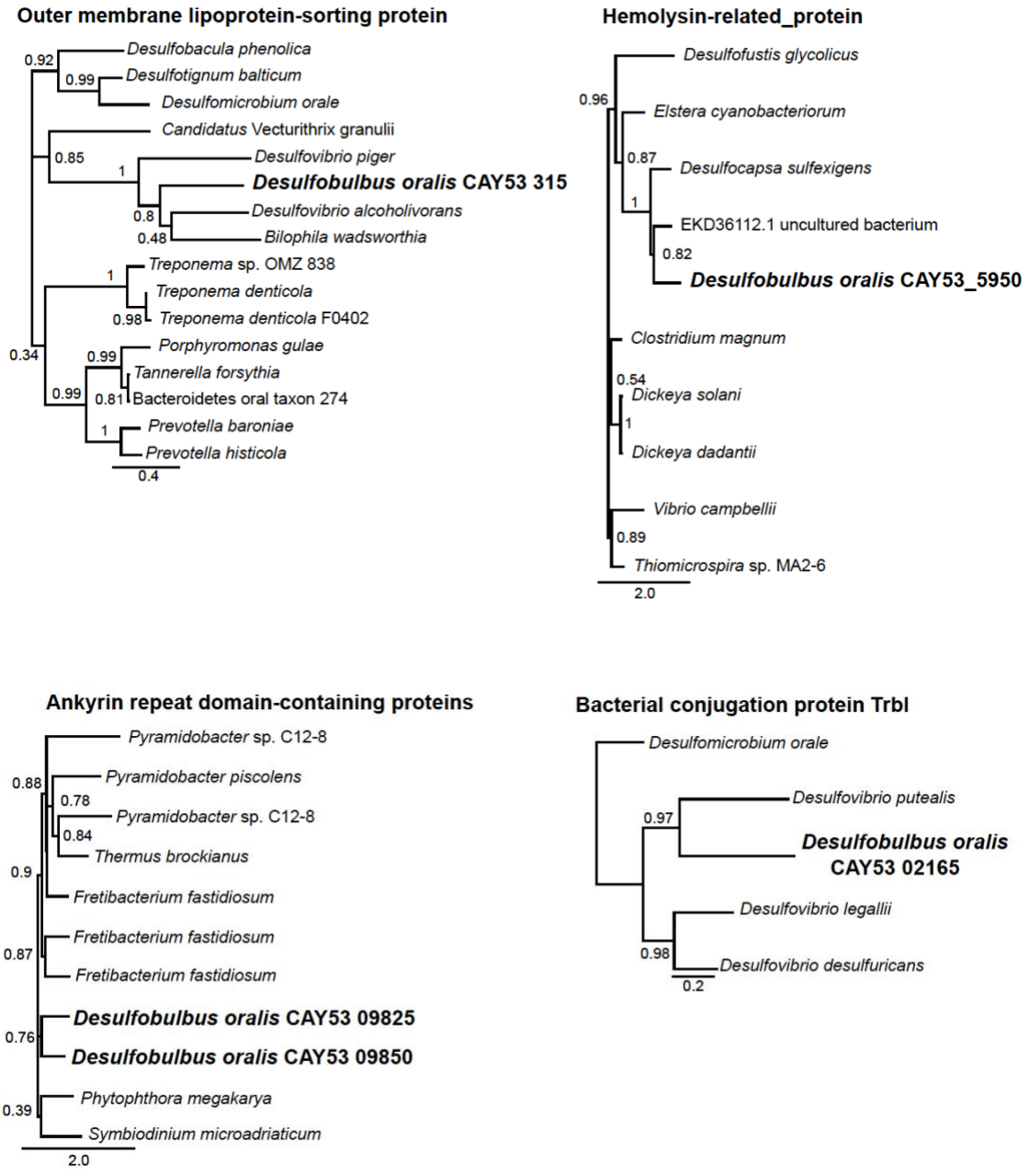


Figure A2.1. (continued)

Figure A2.2. Cytokine stimulation in oral epithelial keratinocytes by *Desulfobulbus oralis* supernatant. Keratinocytes were incubated with *D. oralis* supernatant for 48h and cytokine stimulation was measured using Biolegend LEGENDplex multiplex beads. Each bar represents ± standard deviation of the mean of triplicate measurements of mean technical duplicates. Significantly different from *D. oralis* medium-only control under fermentative conditions (lacking sulfate): * $p < 0.05$, ** $p < 0.01$. IL, interleukin; IFN, interferon.

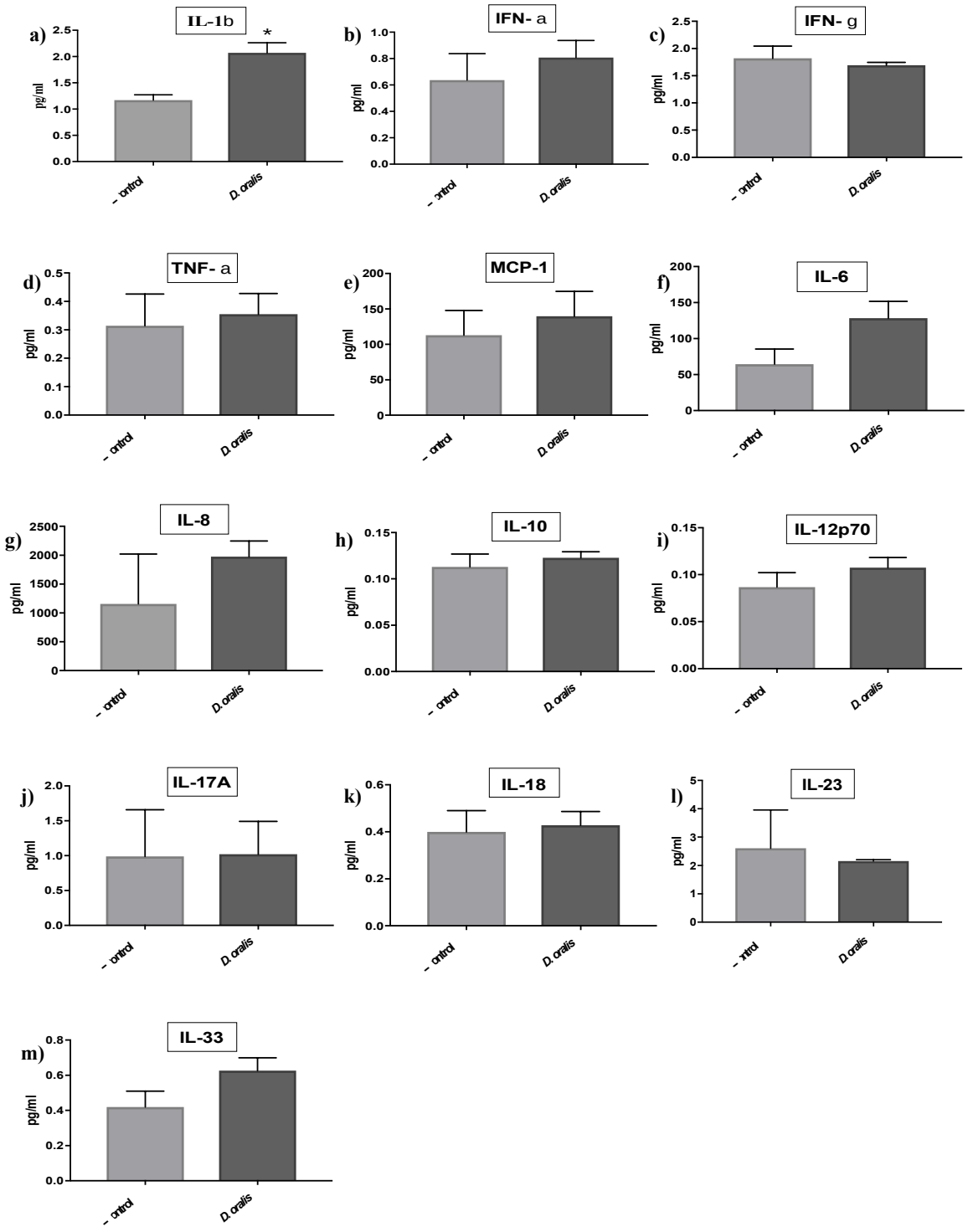
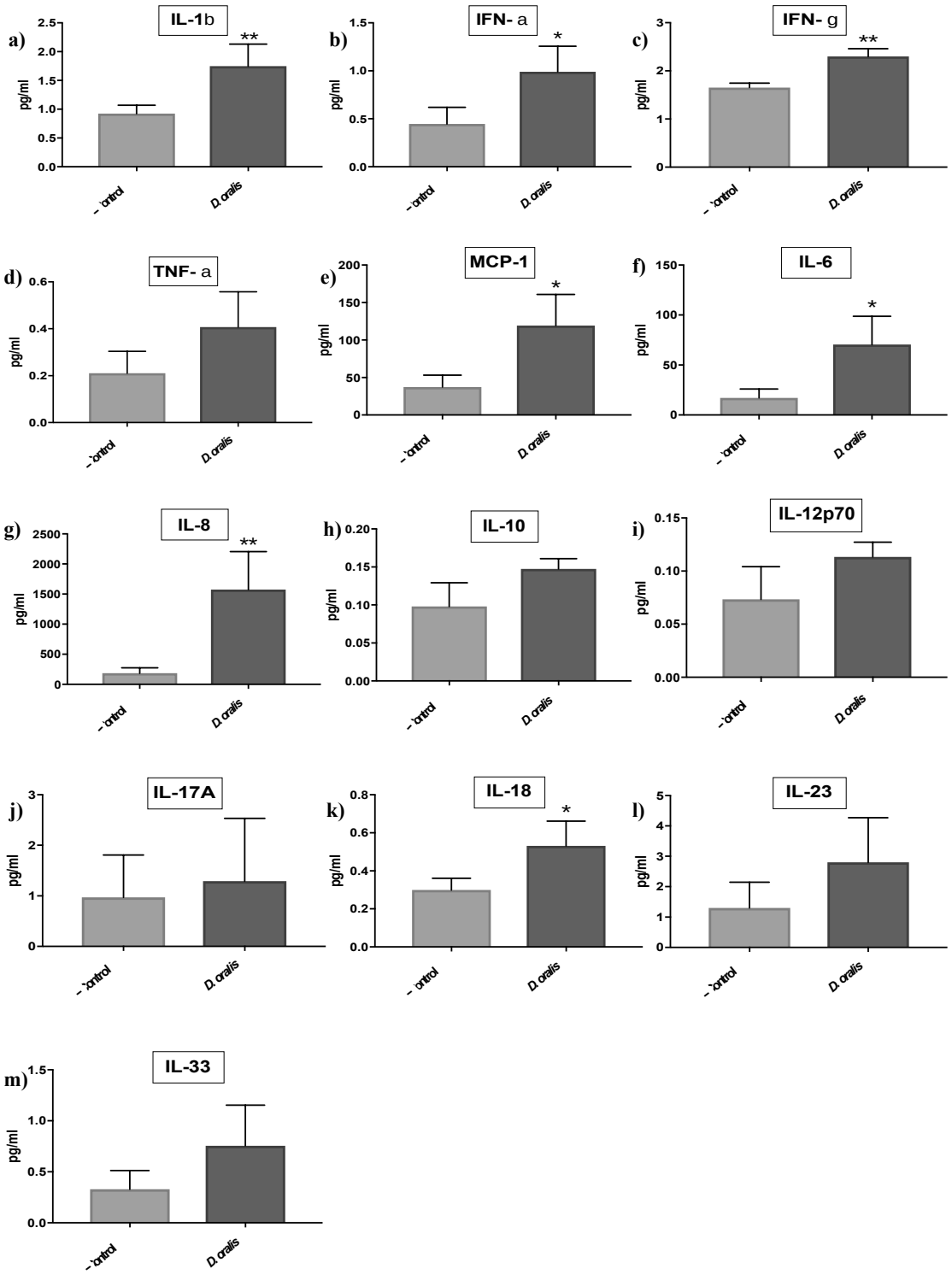

a)
IL-1b
*
pg/ml
2.5
2.0
1.5
1.0
0.5
0.0
- ontrol
D. oralis
b)
IFN- a
pg/ml
1.0
0.8
0.6
0.4
0.2
0.0
- ontrol
D. oralis
c)
IFN- g
pg/ml
2.5
2.0
1.5
1.0
0.5
0.0
- ontrol
D. oralis
d)
TNF- a
pg/ml
0.5
0.4
0.3
0.2
0.1
0.0
- ontrol
D. oralis
e)
MCP-1
pg/ml
200
150
100
50
0
- ontrol
D. oralis
f)
IL-6
pg/ml
200
150
100
50
0
- ontrol
D. oralis
g)
IL-8
pg/ml
2500
2000
1500
1000
500
0
- ontrol
D. oralis
h)
IL-10
pg/ml
0.15
0.10
0.05
0.00
- ontrol
D. oralis
i)
IL-12p70
pg/ml
0.15
0.10
0.05
0.00
- ontrol
D. oralis
j)
IL-17A
pg/ml
2.0
1.5
1.0
0.5
0.0
- ontrol
D. oralis
k)
IL-18
pg/ml
0.6
0.4
0.2
0.0
- ontrol
D. oralis
l)
IL-23
pg/ml
5
4
3
2
1
0
- ntrol
D. oralis
m)
IL-33
pg/ml
0.8
0.6
0.4
0.2
0.0
- ontrol
D. oralis

Figure A2.3. Cytokine stimulation in oral epithelial keratinocytes by *Desulfobulbus oralis* cells. Keratinocytes were incubated with *D. oralis* supernatant for 48h and cytokine stimulation was measured using Biolegend LEGENDplex multiplex beads. Each bar represents ± standard deviation of the mean of triplicate measurements of mean technical duplicates. Significantly different from keratinocyte medium-only control: * $p < 0.05$, ** $p < 0.01$. IL, interleukin; IFN, interferon.


a)
IL-1b
pg/ml
**
b)
IFN- a
pg/ml
*
c)
IFN- g
pg/ml
**
d)
TNF- a
pg/ml
e)
MCP-1
pg/ml
*
f)
IL-6
pg/ml
*
g)
IL-8
pg/ml
**
h)
IL-10
pg/ml
i)
IL-12p70
pg/ml
j)
IL-17A
pg/ml
k)
IL-18
pg/ml
*
l)
IL-23
pg/ml
m)
IL-33
pg/ml
D.oralis

CHAPTER THREE - A step towards elucidating the growth promoting characteristics of *Fusobacterium nucleatum* for cultivation of oral bacteria

Publication Note

The following chapter will be prepared for submission to an appropriate peer-reviewed journal once completed. However it will not be completed by the submission of this dissertation. Other co-authors will include P. Chirania, N. L. Engle, R. J. Giannone, Z. Yang, R. L. Hettich, T. J. Tschaplinski, and M. Podar.

Author's Contributions:

K.L.C. and M.P. conceived the study. K.L.C. was awarded the UTK Penley Fellowship to design and implement this project.

K.L.C. performed laboratory microbial cultivation for *D. oralis* and *F. nucleatum*.

K.L.C. performed the ethyl-acetate and methanol extractions and tests.

P.C., N.L.E., R.J.G., and Z.Y. conducted the mass spectrometry data acquisition and analysis with input from R.L.H., and T.J.T.

K.L.C. wrote this chapter with input from M.P.

Abstract

Periodontitis is the leading cause of tooth loss worldwide. Moreover, it has a complex etiology originating from within the human oral microbiome. Some of the periodontal pathobionts have been identified but many disease-associated lineages remain uncultured, in part due to not yet understanding what these microorganisms may require for growth in the laboratory. Cultivation is important because it provides the opportunity to test specific hypotheses that can tease apart correlation vs. causation in disease studies. Some oral microorganism have been shown to be dependent on *Fusobacterium nucleatum* for growth (39-41, 121). However, it remains unknown what molecules produced by *F. nucleatum* stimulate the growth of other oral species. Using the culture system between the recently isolated *Desulfobulbus oralis* and *F. nucleatum*, we began to investigate what results in the strict requirement of *D. oralis* for *F. nucleatum* by-products for growth in pure culture to occur. Understanding the relationship between *D. oralis* and *F. nucleatum* and what *D. oralis* may require from *F. nucleatum* for growth could allow for the cultivation of more disease-associated bacterial lineages. Using

metabolomics, we began to elucidate the metabolic by-products of *F. nucleatum* growth as a step towards uncovering the growth requirements of *D. oralis*.

Background

Within the human oral cavity, great strides have been made in culturing the diverse array of microorganisms present there. The human oral cavity is home to over 770 species-level taxa (14) and has evolved to exist in complex biofilms/communities that have resulted in genome reduction and strict interspecies interactions (50). As a result, many bacteria remain recalcitrant to cultivation until such dependencies can be identified. We recently cultured the first oral *Desulfobulbus*, *D. oralis*, and showed a dependency of *D. oralis* on *F. nucleatum* for growth in pure culture to be maintained (186). Although this interaction has been seen before (39-41), it still remains unknown what is produced by *F. nucleatum* that stimulates the growth of other oral species. Understanding the relationship between *D. oralis* and *F. nucleatum* and what *D. oralis* may require from *F. nucleatum* for growth could allow for the cultivation of more bacteria and disease-associated bacterial lineages. Additionally, understanding how such species interactions have developed can shed light on how host microbiomes have evolved.

We used the *D. oralis-F. nucleatum* culture system to begin to investigate the metabolic by-products of *F. nucleatum* growth to try and determine the basis of this dependency. We began by optimizing a chemically defined medium (CDM) that supports the growth of *F. nucleatum* and determined that this medium alone could not support growth of *D. oralis*. We then sought to identify the metabolic by-products of *F. nucleatum* and further fractionated the resulting supernatant for further analysis via mass spectrometry. We tested the effect of these individual fractions on *D. oralis* growth and began to build hypotheses about what may result in the inability of *D. oralis* to grow in the absence of *F. nucleatum*.

Results and Discussion

Optimization of a chemically defined medium for growth of *Fusobacterium nucleatum*. *Desulfobulbus oralis* was originally isolated along with a co-isolated strain of *Fusobacterium nucleatum animalis* (186). A cell-free-supernatant (CFS) following *F.*

nucleatum growth also supports growth of *D. oralis* in lab media. Further characterization identified that *F. nucleatum nucleatum* ATCC® 25586 also supported growth of *D. oralis* and was therefore used for the purpose of these tests as it is a well characterized, whole-genome-sequenced reference strain of *F. nucleatum*. Standard growth conditions for *F. nucleatum* involve rich media (e.g. ATCC® 1490, modified chopped meat medium). However, the very high chemical complexity of such media makes it difficult to identify the compounds *F. nucleatum* secretes as a result of its metabolism. Therefore, we selected a chemically defined medium (CDM), previously reported for the growth of oral bacteria (187, 188). *F. nucleatum* grows in that medium under continuous culture conditions (189, 190). To improve the growth in batch cultures we made some modifications: decreasing the levels of reducing agents and buffers and selecting glucose instead of amino acids as the primary energy source. The final CDM included amino acids, nucleotides, and vitamins **(Table 3.1)** and supported *F. nucleatum* growth similar to rich media.

***Fusobacterium* grown in chemically defined medium supports the growth of *Desulfobulbus oralis*.** Although *F. nucleatum* in CDM paralleled that in complex media, it was unknown if in the CDM *F. nucleatum* produced the metabolite(s) required by *D. oralis*. We tested samples from a batch culture of *F. nucleatum* over 70 hours of growth, spanning the various growth stages (lag, early and late exponential and stationary phases) **(Figure 3.1)**. Cell-free-supernatants (CFS) prepared (as detailed in (186)) from those samples were supplemented (10% v/v) into *D. oralis* cultures and growth was measured by OD_{600} **(Figure 3.2)**. *D. oralis* growth was induced by *F. nucleatum* CFS from early culture stages (6 hours) and peaked around 27 hours when exponential phase was reached. The CDM itself did not support the growth of *D. oralis* suggesting that *F. nucleatum* secretes a metabolite distinct from the standard amino acids, nucleotides or vitamins present in the medium.

Metabolomics analysis of *Fusobacterium* cell-free-supernatant by GC-MS. Passing all CFS through a 3kD filter did not remove its growth promoting characteristic, suggesting that compound(s) required by *D. oralis* is not a macromolecule. To begin characterizing the small molecule metabolites produced by *F. nucleatum* we analyzed CFS from mid-log cultures (OD_{600} = 0.46) by gas chromatography mass spectrometry (GC-MS). Using CDM as a control, we compared the chromatogram profiles aiming to

Table 3.1. Composition of the chemically defined medium (CDM).

Components	**Concentration**
Glucose	15mM
NaCl	200μM
CaCl-2H_20	150μM
$MnSO_4$	60μM
$MgSO_4$	1.66mM
NaCitrate	1mM
$FeSO_4$	66μM
Na_2MoO_4	700nM
K_2HPO_4	5mM
Na_2HPO_4	5mM
K_2CO_3	8mM
$(NH_4)_2SO_4$	4.5mM

Vitamins	**Concentration**
Folic acid	250nM
Pyridoxine hydrochloride	2.5μM
Riboflavin	650nM
Biotin	500nM
Thiamine	959nM
Nicotinic acid	2μM
Calcium Pantothenate	500μM
Vitamin B12	1nM
p-Aminobenzoid acid	4μM
Thioctic acid	1μM

Amino Acids	**Concentration**
Glutamic acid	2mM
Histidine	2mM
Lysine	2mM
Leucine	8.75mM
Alanine	2.55mM
Arginine	1.31mM
Asparagine	1.72mM
Aspartic Acid	1.72mM
Cysteine	1.88mM
Glutamine	1.56mM
Glycine	3mM
Myo-inositol	1.26mM
Isoleucine	1.74mM
Methionine	1.53mM
Phenylalanine	1.38mM
Proline	2mM
Serine	2.17mM
Threonine	1.9mM
Tryptophan	1.1mM
Tyrosine	1.26mM
Valine	1.95mM

Nucleotides	**Concentration**
Guanine	175μM
Uracil	275μM
Adenine	400μM

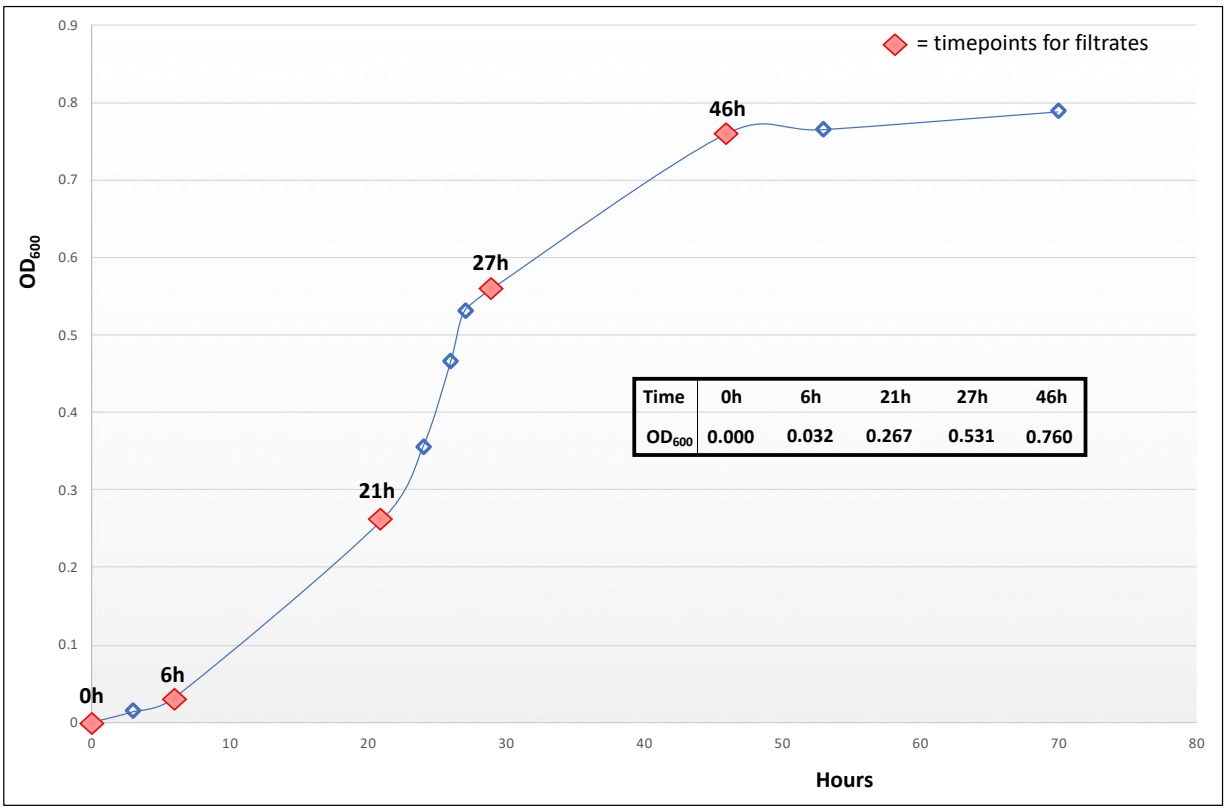


Figure 3.1. Growth of *Fusobacterium nucleatum* on chemically defined medium (CDM). Red diamonds indicate points of growth where cell-free-supernatant (CFS) was collected and tested in Figure 3.2. Inset table shows the OD_{600} of the culture at the time of collection.

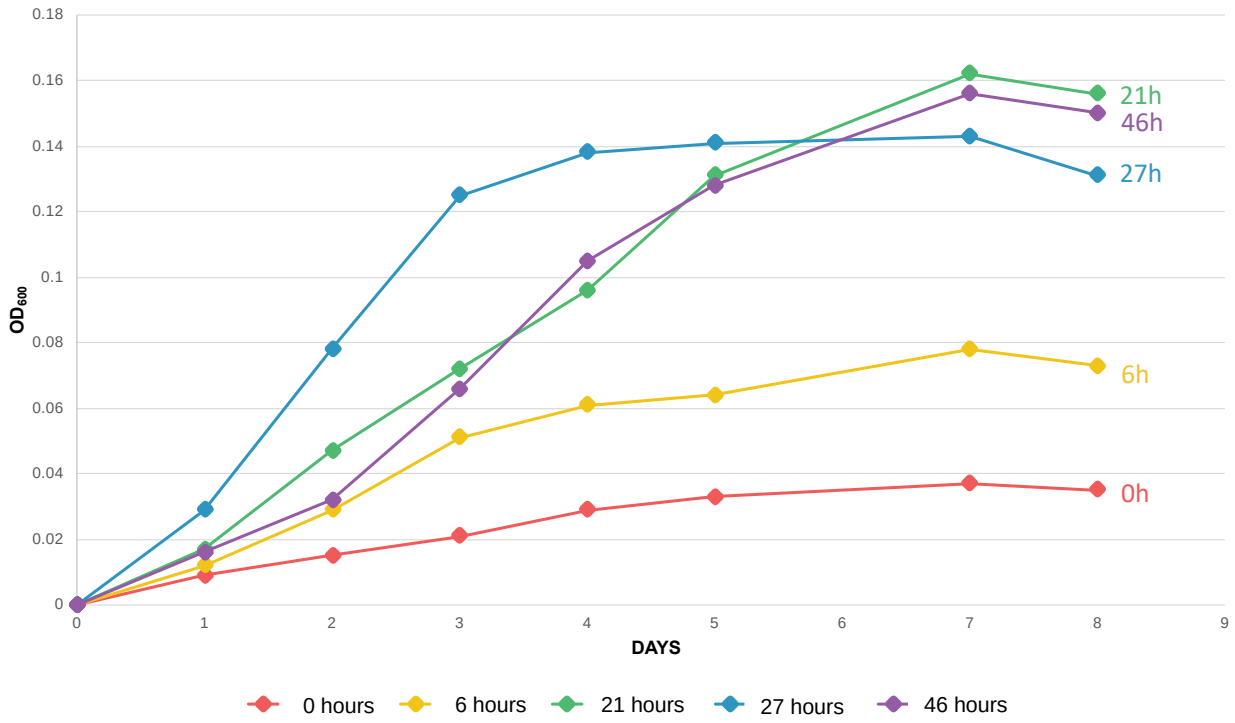


Figure 3.2. Growth of *Desulfobulbus oralis* after supplementation with CFS collected at times 0, 6, 21, 27, and 46 hours of *F. nucleatum* growth. Growth curve is representative of duplicate tests.

identify peaks (metabolites) unique to the CFS. The results from GC-MS detected the presence of cadaverine, cystine, 2-hydoxybutyric acid, and 2-hydroxyglutaric acid **(Figure 3.3)** as well as three unknowns not present in the CDM. Based on these findings, we tested pure 3-indolebutyric acid, gamma-aminobutyric acid (GABA), cystine, 2-hydoxybutyric acid, 2-hydroxyglutaric acid, 2-ketoglutaric acid, and 2-ketobutyric acid for growth support of *D. oralis*. A mixture, as well as individual compounds, were added to *D. oralis* cultures at varying concentrations to determine whether they could stimulate growth. None of the amendments led to *D. oralis* growth.

Chemical fractionation of *F. nucleatum* CFS. To reduce the complexity of the CFS we applied organic phase extraction by using ethyl-acetate and methanol solvent extractions. A total of 14 fractions were collected and tested as to whether or not they supported *D. oralis* growth **(Figure 3.4)**. The ethyl-acetate extractions showed the biggest degree of variability between the samples with fraction 2 (EtAc2) stimulating growth to native growth densities **(Figure 3.4a)**. Fractions 3 and 4 (EtAc3 and EtAc4) showed the most similar growth stimulation to that of the positive control while EtAc1 was most comparable to the negative control sample suggesting growth stimulation did not occur in that fraction. The alkaline and acidic methanol extractions showed less variability when compared to each other (**Figure 3.4c,d)** with all fractions supporting growth to varying degrees. As we wanted to target a distinct fraction that supported growth relative to the others, we focused on ethyl-acetate fraction 2 (EtAc2). We analyzed fraction EtAc2 via GC-MS and identified a single pronounced peak for 2-hydroxybutyric acid.

Ongoing experiments are testing pure 2-hydroxybutyric acid in combinations with amino acids and vitamins to determine whether this compound is capable of stimulating *D. oralis* growth in the absence of *F. nucleatum*. Additionally, as GC-MS may miss some compounds, we are analyzing EtAc2 via liquid chromatography mass spectrometry (LC-MS) and nuclear magnetic resonance (NMR).

Conclusions

The recent cultivation of *Desulfobulbus oralis* provides a novel platform with which to further understand how such organisms have evolved and became dependent on other microorganisms for growth. Understanding how microorganisms have evolved with each

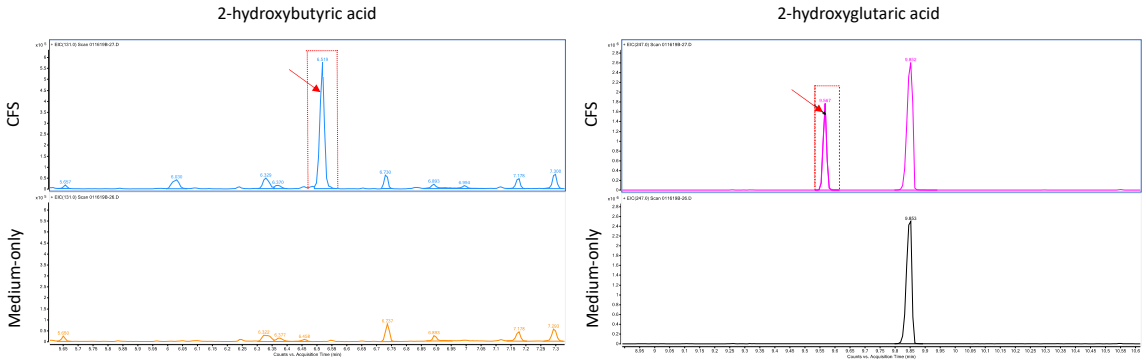


Figure 3.3. Example extracted ion chromatograms from GC-MS for the presence of 2-hydroxybutyric acid (left) and 2-hydroxyglutaric acid (right) in CFS after the growth of *F. nucleatum*.

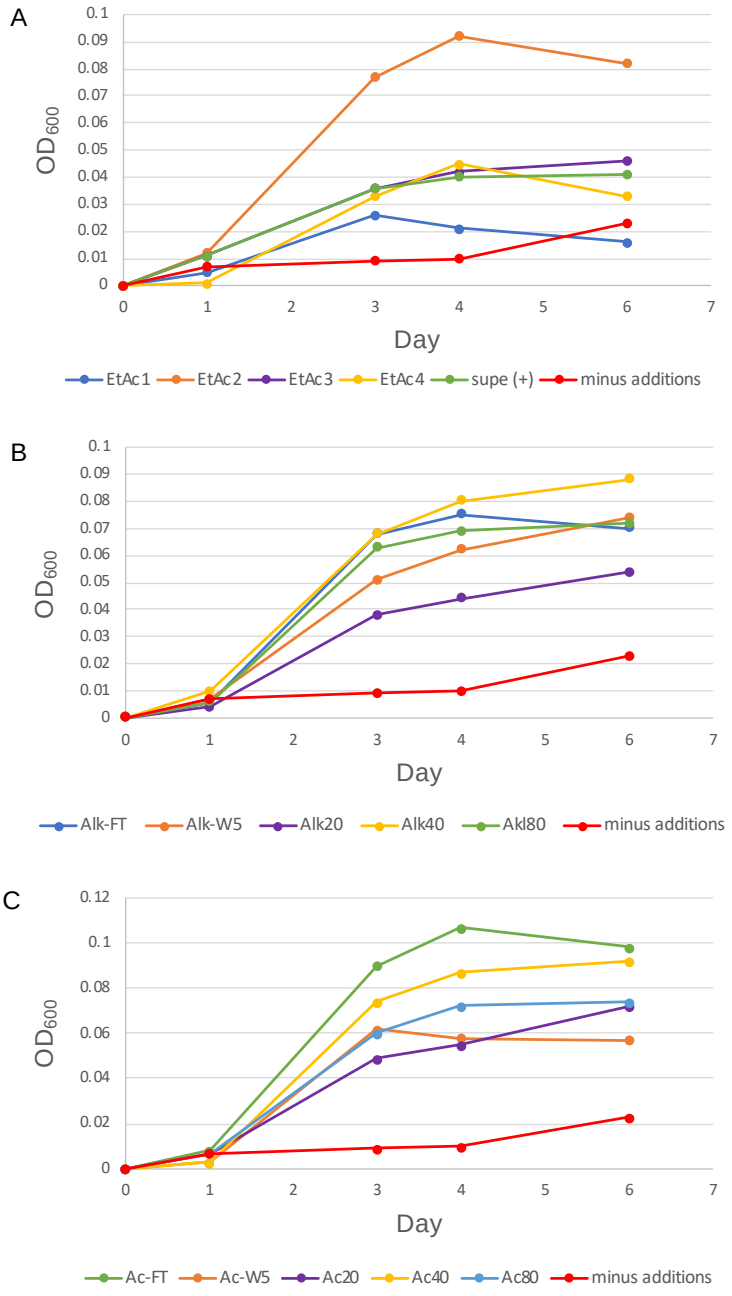


Figure 3.4. Growth of *D. oralis* on the various CFS extractions. **a.** Ethyl-acetate (EtAc) extractions 1-4 (see methods) as compared to a CFS dehydrated positive control (supe +) and a no additions negative control (minus additions). **b.** Alkaline methanol (Alk) extractions for flow-through (FT), wash 5% (W5), 20%, 40%, and 80% methanol as compared to a no additions negative control. **c.** Acidic methanol (Ac) extractions for flow-through (FT), wash 5% (W5), 20%, 40%, and 80% methanol as compared to a no additions negative control.

other, and with the human host, will provide insights into what results in microbial "unculturability". As periodontal disease is a result of a community shift favoring specific microorganisms, understanding these microbe-microbe interactions can shed light on what may trigger a dysbiotic environment. Here we made strides in understanding the metabolic signature of *Fusobacterium nucleatum* and what may result in the stimulation of other oral bacteria.

Materials and Methods

Microbial Cultivation. Pure cultures of *D. oralis* were grown in a modified Lactate-Baar medium containing (L^{-1} distilled water): 2g $MgSO_4.7H_2O$, 1g $(NH_4)_2SO_4$, 10mM MOPS, 10mM HEPES, 10mM sodium lactate, 0.1% yeast extract, 2.5mM cysteine, and 0.5 mg resazurin 0.5%. The pH of the media was adjusted to 7-7.1 with 8M NaOH, degassed with N_2 (100%) and autoclaved for 30 minutes at 121°C in balch tubes fitted with butyl rubber stoppers. After cooling, sterile K_2HPO_4 (0.5M, pH 7) was added to 3mM final. *F. nucleatum* was grown in either BHI autoclaved or CDM **(Table 3.1)** (amino acids and vitamins were prepared using synthetic amino acids supplements (Y1501, Sigma) and a vitamin mix (ATCC® MD-VS™)) pH 7.3. filter-sterilized with 0.22µm Nalgene™ Rapid-Flow™ sterile filters, both under 85%N_2, 10%CO_2, 5%H_2. *F. nucleatum* cell-free-spent supernatant (CFS) was prepared by filtering cultures through two 0.22µm Nalgene™ Rapid-Flow™ sterile filters and degassing with N_2 (100%) and 1mM cysteine was added to keep it reduced in balch tubes fitted with butyl rubber stoppers and added to *D. oralis* medium 10% v/v. Culture stocks were preserved by freezing at -80°C after adding DMSO to 10% (v/v). All OD_{600} results were measured with either an Eppendorf BioPhotometer or a Milton Roy Spectronic 21D fit for balch tubes.

Physiological analyses. Preliminary metabolite tests were performed with L-alpha-hydroxyglutaric acid disodium salt, 2-hydroxybutyric acid sodium salt (97%), L-cystine dihydrochloric acid, indole-3-butyric acid, 2 ketobutyric acid, 2 ketoglutaric acid, and gamma-aminobutyric acid (Sigma-Aldrich). Hydroxybutyric acid, hydroxyglutaric acid, ketobutyric acid, and ketoglutaric acid were prepared at 100mM in water. Indole-3-butyric acid was prepared at 100mM in 100% ethanol. Cystine (100mM) was prepared in

1M HCl. Additions were filter-sterilized, degassed with N_2, and added to *D. oralis* medium at 1-2mM, 0.1-0.2mM, and 0.001-0.002mM and pH was controlled by MOPs.

Ethyl-acetate extractions were prepared by adding *F. nucleatum* CFS (neutral pH) to ethyl-acetate (EtAc) and mixing vigorously. Samples were spun at 2000 x g for 2 minutes to facilitate separation. Top fraction was recovered (**EtAc1)**. Remaining aqueous solution was acidified with HCl to pH 4.5 and EtAc was added to this and mixing and spinning was repeated. Top fraction was recovered **(EtAc2)** and bottom phase **(EtAc3)**. Remaining aqueous phase was acidified to pH 2.3 and further extracted with EtAc by mixing and spinning as above (**EtAc4**). Solvent was removed from samples using a speed vac and either directly prepared for mass spectrometry or resuspended in CDM and brought to neutral pH, filter-sterilized, degassed with N_2, and added to *D. oralis* medium at 10% v/v.

Methanol extractions were prepared by using a Sep-Pak C18 plus short cartridge (360mg Sorbent per Cartridge) product #WAT02515. *F. nucleatum* supernatant was split into two equal portions, one was acidified with HCl to pH 2.2 and the other was made alkaline with KOH to pH 10. Each was then processed for fractions. Solutions of methanol were prepared: 5%, 20%, 40%, 80% in molecular grade water solutions. Using 10ml syringes, the columns were first slowly washed with 100% methanol to remove impurities and to wet the column. Then the columns were equilibrated with a 5% methanol solution. Sample was then slowly passed through the filter and collected (**Ac or Alk flow-through or FT**). Then, another 5% methanol wash was passed through the filter and collected (**Ac or Alk wash 5% or W5**). Then the different percentages of methanol were used to elute (20->40->80 **Ac or Alk 20/40/80**"). Solvent was removed from samples using a speed vac and either directly prepared for mass spectrometry or resuspended in CDM and brought to neutral pH, filter-sterilized, degassed with N_2, and added to *D. oralis* medium at 10% v/v.

Cellular and molecular analyses. To monitor cultures, 16S rRNA gene sequencing was used throughout to confirm purity of the cultures. DNA was extracted using a Zymogen Research fungal/bacterial DNA MiniPrep kit and PCR-amplification was carried out using the universal primer set 27F (5'-AGAGTTTGATYMTGGCTCAG-3') and 1492R (5'-TACGGYTACCTTGTTACGACTT-3'). Amplification with EconoTaq®

(Lucigen©) used the following thermal profile: 95°C for 3 minutes, followed by 32 cycles of 95°C for 30 seconds, 56°C for 45 seconds, 72°C for 90 seconds, and a final extension of 72°C for 5 minutes. Single band products were directly sequenced at Eurofins Genomics (Louisville, KY, USA) using Sanger chemistry.

GC-MS analyses. For GC-MS, 15 µL of sorbitol (1 mg/mL) was added as an internal standard to a 50 µL aliquot of sample and then dried under a stream of nitrogen. Following evaporation, trimethylsilyl derivatives were generated by dissolving the dried sample residue in 0.5 mL of silylation–grade acetonitrile followed by addition of 0.5 mL of N-methyl-N-trimethylsilyltrifluoroacetamide with 1% trimethylchlorosilane (Thermo Scientific, Bellefonte, PA) and heating for 1 hour. A 1 µL aliquot of the trimethylsilyl derivatives solution was injected into an Agilent Technologies Inc. (Santa Clara, CA) 5975C inert XL gas chromatograph-mass spectrometer (GC-MS) fitted with an Rtx-5MS with Integra-guard (5% diphenyl/95% dimethyl polysiloxane) 30 m x 250 µm x 0.25 µm film thickness capillary column. The quadrupole GC-MS was operated in the electron ionization (70 eV) mode with 6 full-spectrum (50–650 Da) scans per second; gas (helium) flow was set at 1 mL/min with the injection port configured in the splitless mode. The injection port, MS source, and MS quad temperatures were set to 250 °C, 230 °C, and 150 °C, respectively. The initial oven temperature was held at 50 °C for 2 min and was programmed to increase at 20 °C per min to 325 °C, to hold for 11 min, and then to return to the initial 50 °C hold. Peaks were assigned using a custom mass spectral database, the Wiley Registry 8th Edition mass spectral database and the NIST 05 mass spectral database.

CHAPTER FOUR - Targeted isolation and cultivation of microbial "dark matter" through reverse genomics

Publication Note

The following chapter is currently in review and represents a revised and re-submitted draft. Other co-authors include J.H. Campbell, M. Balachandran, A.G. Campbell, S.J. Cooper, A. Griffen, M. Heaton, S. Joshi, D. Klingeman, E. Leys, Z. Yang, J.M. Parks, and M. Podar.

Author's contributions.

M.P. conceived the study, designed, and analyzed data.

M.P., M.H., A.G. and E.L. recruited human subjects and collected oral samples.

J.H.C. and M.P. designed the antibodies and performed oral community analysis.

S.C. and J.P. performed protein sequence analyses and structure modelling.

A.G.C. and S.J. aided in single-cell genome collection.

Z.Y. and D.K. performed the MDAs and in-house sequencing.

K.L.C. assembled, annotated and analyzed the single-cell genomes.

K.L.C. performed the TM7 cultivation with M.B.

K.L.C. and J.H.C. contributed equally to this work.

K.L.C. and M.P. wrote the manuscript.

Abstract

The vast majority of microorganisms, across all taxonomic levels, have remained uncultured. Single cell genomes and metagenomes are expanding the known diversity of Bacteria and Archaea and have facilitated inferences of potential physiological and environmental roles, but most are not experimentally validated. Ideally, sequence data should enable targeted microbial isolation, cultivation, and characterization. Here we used genome-informed antibody engineering to selectively isolate single cells, cultivate previously "unculturable" bacteria, and identify novel symbioses. By targeting the candidate phylum TM7 ("Saccharibacteria") from the human oral microbiota, we sequenced genomes representing 8 distinct lineages. Viable cell isolation enabled cultivation of three different species-level lineages of oral TM7 as epibionts on diverse *Actinobacteria*, demonstrating that this approach can lead to discovery and characterization of novel interspecies partnerships. As the concept could be applied to

any group of microorganisms, from any environment, it unlocks the potential to directly isolate, culture, and characterize representatives of the yet-uncultured branches of the microbial tree of life.

Introduction

Our understanding of the microbial tree of life has undergone continuous reshaping and refinement ever since the introduction of molecular, culture-independent methods to identify and classify organisms. While pre-genomic, ribosomal RNA gene sequencing resulted in the discovery of nearly two dozen candidate Bacteria and Archaea phyla (191), more recent single-cell genomic and metagenomic assemblies (SAGs and MAGs) have led to a deluge of proposed novel lineages (27, 192-194). Notably, however, SAGs and MAGs of specific, target taxa are generally retrieved by chance, depending on their abundance in the community and currently there are no effective ways to focus on specific, individual taxa. Furthermore, although microbial cultivation has been successful for representatives of a handful of phyla initially discovered by rRNA sequencing, the vast majority of candidate microbial taxa remain uncultured. Genome-based metabolic reconstructions can provide important insights into the lifestyle and characteristics of various uncultured Bacteria and Archaea but, with few exceptions, have not led to targeted isolation of novel microbes (164, 195). Growth requirements of many organisms may be too complex, or too subtle, for genomic-based predictions and may depend on specific interspecies interactions (e.g. Nanoarchaeota (164, 196, 197), TM7/"Saccharibacteria" (45)). High-throughput cultivation approaches, multispecies consortia, and addition of various growth factors have been used successfully to propagate and isolate "unculturable" organisms from various environments including soils, water, sediments, and the human microbiome (41, 198-202). However, being able to directly link the genotype to actual cells while maintaining cellular viability would enable isolation of cells from target microbial taxa for selective genomic sequencing or cultivation and characterization of interspecies interactions. Here we introduce such an approach, using antibodies against predicted cell surface protein epitopes from uncultured bacteria, with human oral TM7/"Saccharibacteria" as a test case **(Figure 4.1a)**.

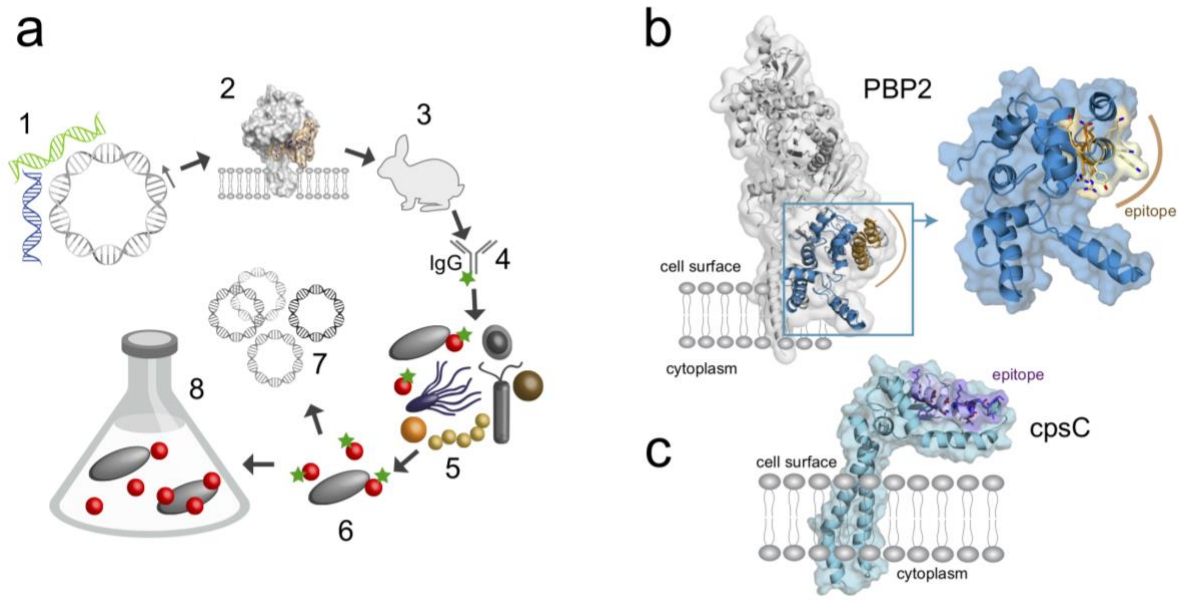


Figure 4.1. Overview of targeted microbial isolation through reverse genomics. **a.** Diagram showing the steps in reverse genomics-enabled microbiology: (1) Identification of membrane protein-encoding genes in SAGs and MAGs, (2) selection of predicted exposed epitopes, (3) antibody production, (4) purification and fluorescent labelling, (5) staining of target cells from microbiome samples (6) and isolation for genomic sequencing (7) or cultivation (8). **b.** X-ray structure of penicillin-binding protein from *E. coli*, with the modeled TM7 glycosyltransferase (GT) domain colored blue and the selected epitope region highlighted in yellow-brown. c. Structural model of TM7a capsular polysaccharide biosynthesis protein (cpsC), with the peptide epitope region in purple.

TM7 was one of the first candidate bacterial phyla proposed two decades ago based on rRNA gene sequences (76, 203) and was subsequently detected in a wide range of environments, including in association with plants and animals. Following metagenomics-based genomic assembly of near-complete genomes for several TM7 lineages from activated sludge, the group has been renamed “Saccharibacteria” (44) (204), although as a formal classification and valid description are still lacking. In humans approximately a dozen distinct species-level TM7 are recognized (205), mostly in the oral cavity, and their increased abundance has been correlated with the development of periodontitis (206) and inflammatory bowel disease (74). Related lineages have also been identified in dogs, cats and dolphins (127, 207, 208). The first genomic data for a human oral TM7 was generated by single-cell sequencing following micro-isolation from supragingival plaque (28). Recently, the first TM7 bacterium (TM7x) was cultured as an obligate epibiont on the human oral *Actinomyces odontolyticus*, based on serendipitous streptomycin resistance (45). Considering the large taxonomic diversity and ubiquitous environmental presence of these organisms, generally at low abundance (<1%), we selected TM7 for developing reverse genomics-enabled microbial isolation.

Results

Selection of TM7 bacterial cell surface antibody targets. We used the first single-cell genomic data from a human oral TM7 (TM7a)(28) to identify genes encoding predicted membrane-associated proteins for which function could be inferred, based on homologues from other bacteria. We avoided hypothetical proteins, because in the absence of gene expression information we considered those less likely to be present on the cell surface. It should be noted, however, that if experimental data (e.g. based on metaproteomics, metatranscriptomics) would be available, any categories of membrane protein could be a potential target. We then screened for proteins with extracellular domains that could be used as antigens for antibody development and which had homologues with solved three-dimensional structures. A candidate protein that satisfied these criteria was a predicted penicillin-binding protein (PBP2). In pathogenic bacteria, PBP2 has been shown to be an effective vaccine target, indicating that the protein is expressed and can be recognized by an antibody (209). Crystal structures for

Staphylococcus aureus (PDB entry 3DWK) and *Escherichia coli* (3VMA) protein homologs were used to model the TM7a PBP2 and identify potentially exposed epitopes. PBP2 is anchored in the membrane by a single α-helix and consists of two extracellular domains, an N-terminal glycosyltransferase (GT) domain and a C-terminal transpeptidase (TP) domain connected by a β-rich linker region (210) **(Figure 4.1b)**. A 19-amino acid sequence located in the GT domain was selected as antigen based on surface exposure, predicted immunogenicity and solubility, and was used for raising polyclonal antibodies **(Figure A4.1)** (all additional material occurs in the chapter appendix). As experimental structure data for membrane proteins is relatively scarce, computational prediction may also be used for antigen selection. Capsular polysaccharide biosynthesis protein (CpsC) was selected as second target but it lacked close homologues with solved structures. We generated a Rosetta 3D structure model using coevolution-derived residue-residue contact restraints (211), and a predicted antigenic peptide from an extracellular α-helical domain was used to raise antibodies.

While for this study we selected short peptides, entire protein domains could be used as antigens for generating polyclonal antibodies. Such antibodies would target multiple regions of the membrane protein and could increase the chance of binding, especially if some residues may be sterically hindered (e.g. due to glycosylation). However, depending on the degree of sequence conservation, antibodies raised against protein domains may be less specific and bind to bacteria other than the target group.

Antibody-enabled isolation and cultivation of oral TM7/Saccharibacteria symbiotic lineages. Human oral samples (saliva, subgingival fluid) from orally healthy donors and from patients with periodontitis were incubated with fluorescently labeled anti-PBP2 and anti-CpsC IgGs (anti-TM7). Stained samples were analyzed by flow cytometry, revealing a distinct population of fluorescent cells (0.1-5%, depending on the sample) **(Figure 4.2a, b)**. To determine if Saccharibacteria were labeled by the antibody, we sorted pools of 10-100 fluorescent cells followed by multiple displacement genomic amplification (MDA). We then assayed for the presence of TM7 bacteria using PCR with specific primers targeting the small subunit ribosomal RNA (SSU rRNA) gene followed by Sanger sequencing or by MiSeq amplicon sequencing of the V4 region of the SSU rRNA gene amplified with universal primers. Between 25-100% of the cell pools were positive

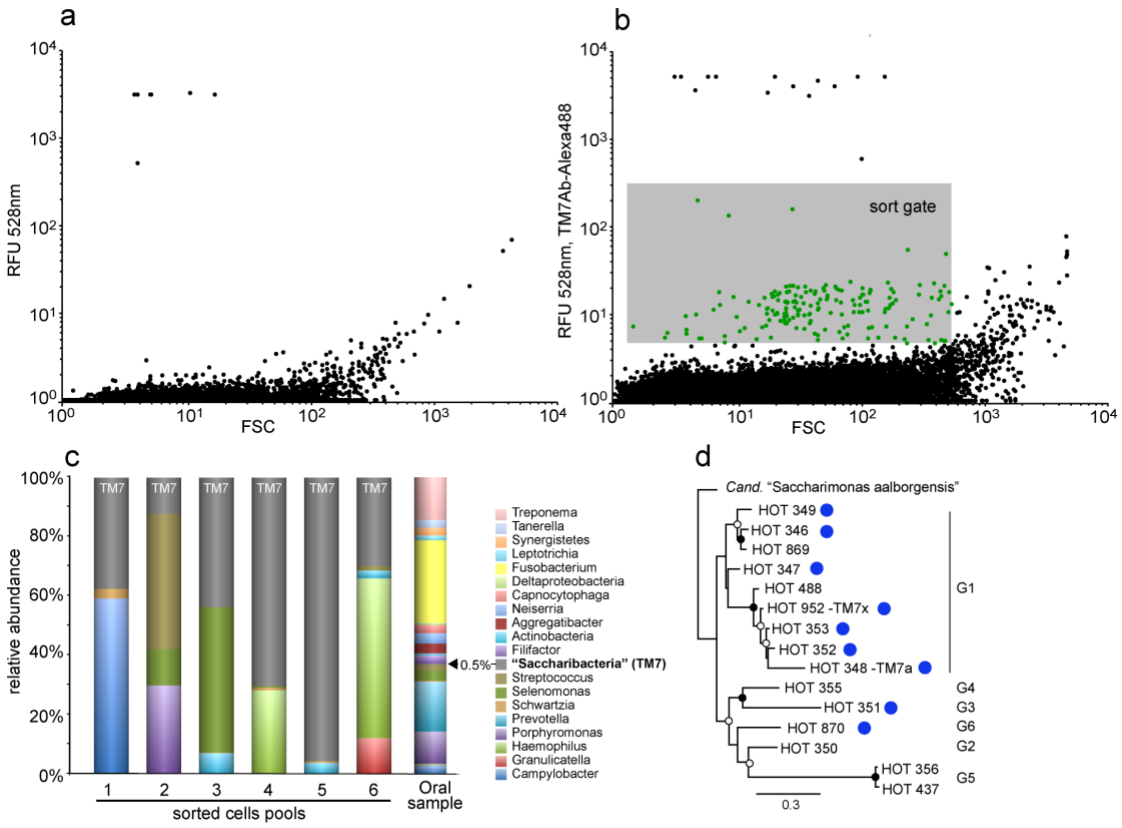


Figure 4.2. Isolation of TM7 cells by flow cytometry cell sorting. **a-b.** Flow cytometry plot (forward scatter "FSC" versus green fluorescence "RFU") of native **(a)** and TM7 antibody-labelled oral samples **(b)**. The sort gate area indicates selection characteristics for cell sorting. Highly fluorescent particles (RFU>103) are auto fluorescent precipitates/minerals. **c.** Microbial diversity in pools of 10 sorted cell pools, following MDA and 16S rRNA gene amplicon sequencing. Also shown for comparison is the diversity of the original subgingival samples (far right column). **d.** Maximum likelihood tree of human oral TM7 SSU rRNA sequences, indicating phylotypes retrieved by antibody labeling and cell sorting (blue dots) and TM7 groups (G1-G6). Circles at nodes indicate bootstrap values (not shown if <50, black>80). Scale bar = substitutions per site.

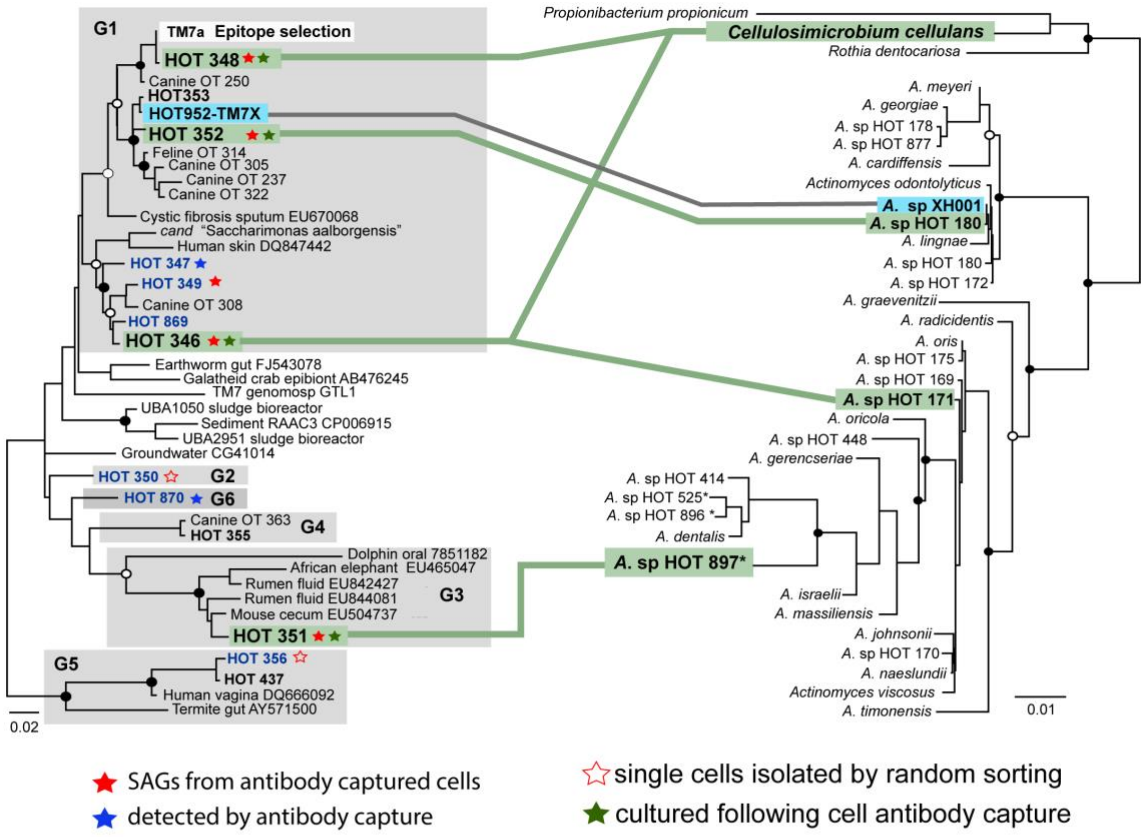
for TM7, depending on individual subjects/samples and gating parameters, with both anti-PBP2 and anti-CpsC successfully retrieving TM7 cells. To enhance the fluorescence signal through dual labelling and to cover as much as possible of sequence variation space, most experiments were subsequently performed with a mix of the two antibodies. As such, when we analyzed the pools for taxonomic representation, Saccharibacteria/TM7 comprised up to 95% of the sequences and represented phylotypes from groups 1, 3 and 6 **(Figure 4.2c, d and Figure A4.2)**. Other bacteria were primarily members of *Fusobacteria, Actinobacteria, Bacteroidetes, Haemophilus* and *Streptococcus*. The identification of species other than the target TM7 bacteria may indicate cross-reactivity of the antibodies, free genomic DNA, and interspecies cellular adhesion (specific or non-specific) which is extensive in the microbiota (212).

We demonstrated that the anti-TM7 peptide antibodies can label target cells in biological samples and enable their isolation by flow cytometry. We tested if cell isolation could then be viable and propagate in cultures. Importantly, such cultures would enable identification of potential physiological interactions between various TM7 lineages and other bacteria in the oral microbiota. Using freshly collected oral samples, we labeled the bacteria with the anti-TM7 antibodies and sorted individual fluorescent particles (cells) on culture media. Following incubation under anaerobic or hypoxic conditions, the cultures were tested for the presence of TM7 bacteria using rRNA gene amplicon sequencing.

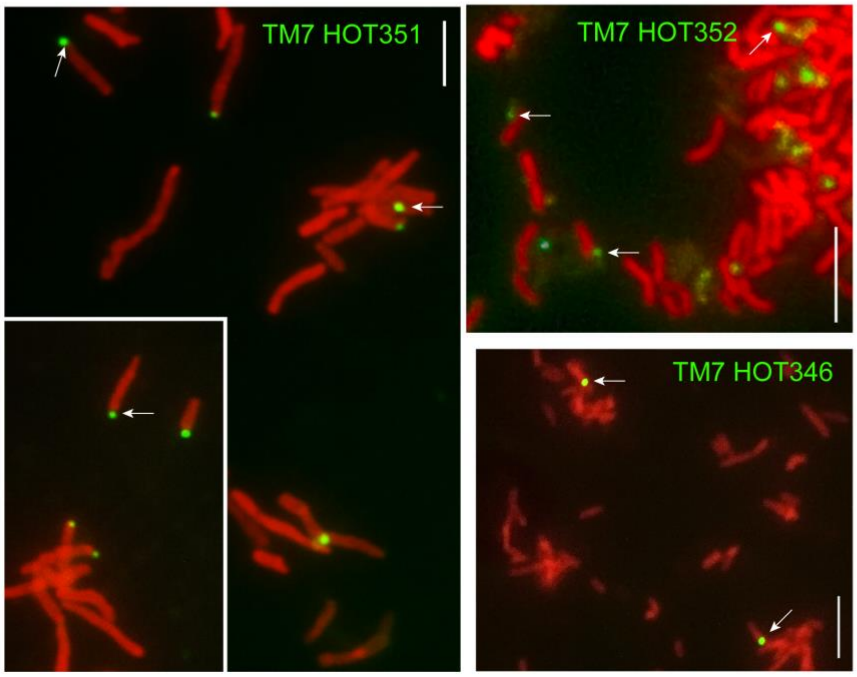
On solid media, we identified colonies containing five distinct TM7 phylotypes (species to family/order taxonomic range), HOTs 346, 348, 352, 952/TM7x (all group 1) and HOT 351 (group 3) **(Figure 4.3, Figure A4.3)**. Based on amplicon sequencing, most TM7 positive colonies also contained species of *Actinobacteria* and some also contained *Streptococcus*. This finding suggested that *Actinobacteria* may be host organisms, as with TM7x, and that *Streptococcus* cells, known to be highly adherent (213), are co-isolates. Surprisingly, TM7 HOTs 346 and 348 (distinct species, based on 93% rRNA identity), were identified in the same colony with *Cellulosimicrobium cellulans*, an actinobacterium present in healthy saliva but occasionally associated with bacteremia and other diseases (214). As the fluorescent particle could not have contained multiple *Cellulosimicrobium* cells that independently also harbored different TM7, multiple epibiont TM7 species may simultaneously parasitize the same host cell. A second round of

Figure 4.3. Diversity of cultivated and uncultivated TM7 bacteria. **a.** Maximum-likelihood phylogenies of TM7 bacteria and of selected human oral *Actinobacteria*, based on small subunit rRNA genes. TM7 bacteria include human oral phylotypes (HOTs) as well as related host-associated lineages from open environments, identified by sequence accession numbers and associated with six operational groupings (G1-6). HOTs detected using antibodies, isolated as single cells, or identified in cultures are indicated by colored stars. Circles at nodes indicate bootstrap support (>80% filled, 50-80% open, only shown for major groupings). Lines connecting TM7 and *Actinobacteria* species indicate epibiont-host systems identified previously (blue highlight) and in this study (green). **b.** Epifluorescence microscopy images of novel TM7-host associations identified in this study using immunolabeling (TM7 HOT351-*Actinomyces* HOT897 and TM7 HOT346-*Cellulosimicrobium cellulans*) or FISH (TM7352-*Actinomyces odontolyticus* strain OR). Arrows point to some of the green fluorescent TM7 cocci. Scale bar is 5 μm).

a


G1
TM7a Epitope selection
HOT 348
Canine OT 250
HOT353
HOT952-TM7X
HOT 352
Feline OT 314
Canine OT 305
Canine OT 237
Canine OT 322
Cystic fibrosis sputum EU670068
cand "Saccharimonas aalborgensis"
Human skin DQ847442
HOT 347
HOT 349
Canine OT 308
HOT 869
HOT 346
Earthworm gut FJ543078
Galatheid crab epibiont AB476245
TM7 genomosp GTL1
UBA1050 sludge bioreactor
Sediment RAAC3 CP006915
UBA2951 sludge bioreactor
Groundwater CG41014
HOT 350 G2
HOT 870 G6
Canine OT 363 G4
HOT 355
Dolphin oral 7851182
African elephant EU465047
Rumen fluid EU842427
Rumen fluid EU844081 G3
Mouse cecum EU504737
HOT 351
G5
HOT 356
HOT 437
Human vagina DQ666092
Termite gut AY571500
0.02
Propionibacterium propionicum
Cellulosimicrobium cellulans
Rothia dentocariosa
A. meyeri
A. georgiae
A. sp HOT 178
A. sp HOT 877
A. cardiffensis
Actinomyces odontolyticus
A. sp XH001
A. sp HOT 180
A. lingnae
A. sp HOT 180
A. sp HOT 172
A. graevenitzii
A. radicidentis
A. oris
A. sp HOT 175
A. sp HOT 169
A. sp HOT 171
A. oricola
A. sp HOT 448
A. gerencseriae
A. sp HOT 414
A. sp HOT 525*
A. sp HOT 896 *
A. dentalis
A. sp HOT 897*
A. israelii
A. massiliensis
A. johnsonii
A. sp HOT 170
A. naeslundii
Actinomyces viscosus
A. timonensis
0.01
SAGs from antibody captured cells
detected by antibody capture
single cells isolated by random sorting
cultured following cell antibody capture

b


TM7 HOT351
TM7 HOT352
TM7 HOT346

labeling and cell sorting resulted in a pure co-culture of *C. cellulans*-TM7 HOT346. In a separate experiment using an oral sample from the same donor, a second TM7 HOT346 strain was co-isolated with *Actinomyces* sp. HOT171. This represents the first demonstration that in their natural environment (oral microbiota) a TM7 species can be an epibiont on bacteria that belong to distinct genera.

TM7 HOT351 is the first cultivated member of a novel family to order-level TM7/Saccharibacteria taxon (85% 16S RNA identity to TM7x) and was identified in pure co-culture with a previously uncultured *Actinomyces* species (*A.* sp. HOT897), most closely related to *A. dentalis* (92% identity). As characterization of all human oral bacteria has been a continuous effort over decades, the isolation of a novel species of *Actinobacteria* further highlights the importance of novel approaches in enabling co-cultivation of interacting organisms that were refractory to traditional methods.

For two closely related phylotypes, TM7 HOT352 and HOT952, which belong to the same species complex as TM7x, we identified the host as *Actinomyces odontolyticus* XH001, the known host of TM7x (45). In liquid microcultures we also detected TM7 HOT347 (Group 1) and HOT870 (Group 6) however those cultures could not be maintained, and we did not identify the native host of those Saccharibacteria lineages. Identification of multiple oral TM7-host pairs provides the first substantial view into the evolution of such interbacterial relationships within a candidate phylum. Interestingly, there is no apparent correlation in phylogenetic distance between the TM7/Saccharibacteria species and their *Actinobacteria* hosts, which would indicate co-speciation **(Figure 4.3a)**. This suggests that the parasite-host interaction evolved separately, likely after oral colonization, and were not necessarily linked to speciation events. Host specificity appears stringent for TM7x, as closely related Actinomyces cannot substitute the strain that it was co-isolated with (45). We have further expanded on our understanding of TM7 physiology by identifying cells from distinct TM7 lineages that can simultaneously “infect” the same host cell and that the same TM7 species can “infect” distinct hosts *in vivo*. That indicates some flexibility in the establishment of the ectosymbiosis. The molecular mechanisms of interaction and the host specificity determinants for the different TM7 species and their hosts remains unknown. As other members of the “Candidate Phyla Radiation” are hypothesized to have an obligate

parasitic/symbiotic lifestyle (215), isolation of additional TM7 and related symbiotic systems will be important for better understanding of the evolution of these microbial lineages and characterizing the molecular strategies that ectobionts use to attach to host cells and obtain metabolic building blocks.

All TM7 are very small epibiont cocci (<0.5 μm in diameter) on the surface of host cells, observed using both fluorescence in situ hybridization (FISH) and staining with the TM7 antibody **(Figure 4.3b)**. TM7 HOT351 cells appear primarily associated with the pole of the bacillus-shaped Actinomyces host cells, while for the others, there does not appear to be a localization preference (46). As with TM7x-*A. odontolyticus* XH001, we found that growth conditions affect the relative abundance of the different TM7 in the co-cultures, although each appears to have its own characteristics. Whereas TM7x abundance increased in the presence of oxygen, TM7 HOT351 and 346 were inhibited by micro-oxic environment as their hosts were also inhibited. The physiology of each TM7 lineage is therefore linked to that of their hosts and the subject of ongoing studies for all of the symbiotic systems we isolated.

Antibody-enabled single cell genomics. Because not every organism may be readily culturable, we also tested the targeted isolation approach to obtain single cell genomic information for the various oral TM7/Saccharibacteria lineages. Targeted collection of specific genomes could further provide hypotheses to begin to isolate other uncultured lineages of bacteria. Flow cytometry enables analysis of over 5,000 cells per second, therefore taxa present at extremely low abundance could be targeted for single cell isolation and genomic sequencing. Labeled samples from healthy individuals and from individuals with periodontitis were therefore used for single cell sorting and MDA. We identified over two dozen TM7 single-cell amplified genomes (SAGs), which were assigned to six human oral taxa, HOT's 346, 348, 349, 351, 352 and 353/952. SAGs from two other TM7 HOTs (350 and 356), identified through un-targeted single-cell isolation (32, 34), were also included in the analysis. Unlike genomes derived from metagenomes (MAGs), that most often lack rRNA sequences and are difficult to match with lineages previously known based on 16S studies, SAGs enable such comparisons. This is important for organisms that span a wide taxonomic space (such as the oral Saccharibacteria), ranging from species/subspecies to genera and families/orders. In

total, we sequenced and assembled 23 of the SAGs, representing four of the six groups of TM7/Saccharibacteria in the human oral microbiome (Groups 1, 2, 3 and 5) **(Table S4.1)**. The assembled SAGs ranged from 0.25 to 1 Mbp (median 0.62 Mbp), reflecting different levels of genome completion **(Figure 4.4a, Figure A4.4, and Table S4.1)**. In comparison, the closed genome of TM7x is 0.7 Mbp (45), while MAGs of TM7/Saccharibacteria from open environments range between 0.8-1.2 Mbp (44, 47, 193). Based on conserved single copy core genes **(Table S4.1)**, we estimated the level of completion of the SAGs to be between 40 and more than 90%. The oral TM7 genomes share a core of 279 shared orthologous gene families, mostly related to genetic information processing, transport and secretion functions. The small genome size, especially for oral TM7, is a result of gene (and metabolic capability) loss associated with their parasitic lifestyle, which likely extends across all oral phylotypes. However, it is not known if TM7 from open environments are dependent on physical association with other bacteria. It has been proposed, based on the interspersed oral and environmental TM7 lineages, that multiple independent acquisitions of oral microbiota were followed by independent genome reductions (and possible bacterial host associations) among the TM7 subtypes (216). As humans share close relatives of oral TM7 with other mammals **(Figure 4.3a and Figure A4.5)**, oral colonization by these bacteria likely occurred early in the evolution of mammals, if not before. A two-fold increase in the abundance of genes related to defense mechanisms (CRISPR-Cas, restriction-modification systems and putative virulence factors) and a reduction in genes associated with cellular motility and redox processes (**Figure A4.6 and Table S4.1**) are also potentially linked to adaptation to a epibiotic lifestyle (45, 208, 216).

As with other related lineages from the “Candidate Phyla Radiation” (CPR)(27, 215), genome reduction in TM7/Saccharibacteria has led to the loss of numerous functions, including biosynthesis of essential amino acids and vitamins. All TM7 appear to have an obligately fermentative metabolism with no respiratory functions, as electron transport chains are lacking. None of the TM7 can synthesize sugars by gluconeogenesis, as the key enzyme fructose-1,6-bisphosphatase is absent. Glucose and other monosaccharides may be fermented via the pentose phosphate and heterolactic fermentation pathways. We identified, however significant differences between the

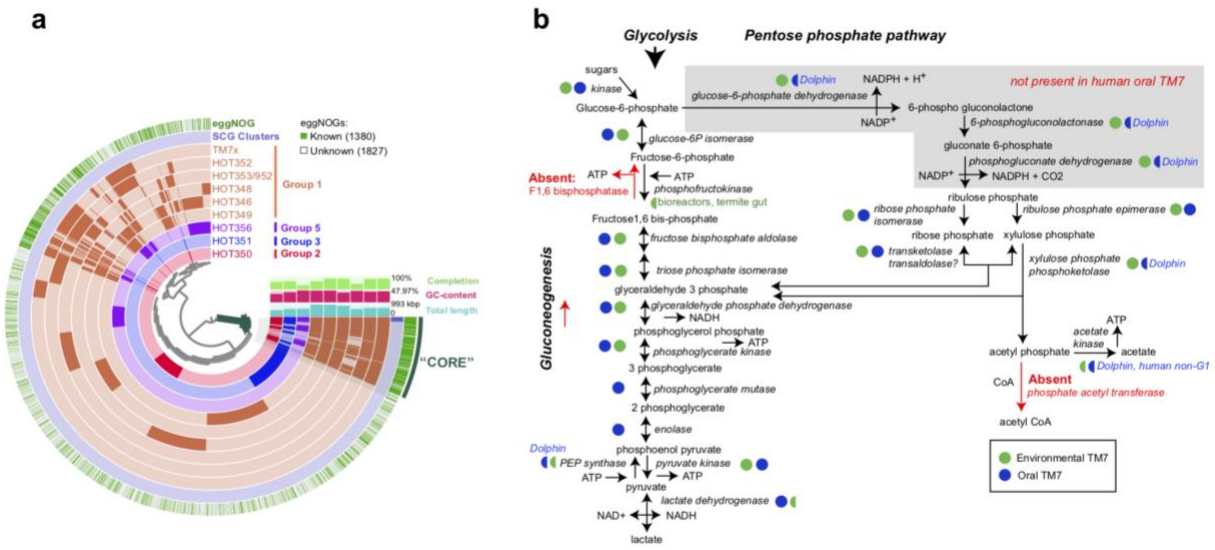


Figure 4.4. Comparison of oral TM7 at genomic and metabolic levels. **a.** Alignment of oral TM7 genomes using Anvi'o (217, 218), showing basic genome characteristics, distribution of core genes and functional gene categories across the various phylotypes. **b.** Reconstruction of central metabolic pathways for TM7 bacteria, based on completed genomes and draft SAG/MAG assemblies. Blue or green circles indicate broad presence in environmental and animal-associated TM7. Semicircles indicate occasional presence, associated with specific environments. Absent reactions are in red. The absence of genes encoding the three essential pentose phosphate pathway enzymes in human oral TM7 is highlighted in gray.

various types of TM7 based on environmental origin and phylogenetic grouping **(Figure 4.4b)**. TM7 genomes from open environments encode a nearly complete pentose phosphate pathway coupled with heterolactic fermentation, which could generate reduced NADPH and ATP via substrate-level phosphorylation by acetate kinase and pyruvate kinase. However, genes encoding two of the enzymes for the terminal part of the glycolytic cycle, phosphoglycerate mutase and enolase could not be identified in any of the environmental TM7 genomes, although they are present in all oral TM7. On the other hand, all human TM7 appear to lack important enzymes for the initial NADPH-generating reactions of the pentose phosphate pathway and also a key enzyme for heterolactic fermentation, xylulose phosphate phosphoketolase. Therefore, oral TM7 may use hexose-pentose interconversions and primarily operate glycolysis reactions for ATP synthesis. Evidence of several other differential gene (and metabolic potential) loss between various lineages of TM7 (acetate kinase, lactate dehydrogenase phosphofructokinase and phosphoenol pyruvate synthase), suggests that central metabolic pathways have been in continuous reshaping and reflect adaptation to specific open environments and oral niches **(Figure 4.4b)**.

Because TM7 bacteria are auxotrophs for amino acids, nucleotides and vitamins, they require efficient uptake mechanisms. Based on comparative genomic analysis of host-associated and open environment TM7 genomes we identified only a handful of recognizable transporters. Paucity of canonical transport mechanisms has been noted in other obligate symbionts (e.g. Nanoarchaeota, insect intracellular endosymbiotic bacteria), where it has been proposed that direct cytoplasmic bridges (219) or specialized proteins that change membrane permeability (220) may enable intercellular molecular fluxes. In TM7, as in Nanoarchaeota, such transport processes must occur over an area only few nanometers across (164). As such ectobionts may be able to detach from their host cells and survive free for a limited time, they may have mechanisms to maintain some energetic independence (ATP and reducing equivalents) by acquiring specific central metabolic substrates directly from the environment. We hypothesized, based on the glycolytic pathway predicted in all human oral TM7 (**Figure 4.4b**), that supplementation of the culture medium with glycolysis substrates might stimulate growth if they could be utilized directly. As such, we tested media supplemented with 2-5mM

additions of glucose, ribose, phosphoenolpyruvate, pyruvate or lactate, singly or in combination. However, we did not detect any effect relative to any of the TM7 strains. While we cannot conclude TM7 detached from their host cells could not uptake glycolytic substrates *in vivo*, in laboratory cultures they appear to exclusively rely on metabolites provided through cell contact with *Actinobacteria*. In several of the oral TM7 genomes we identified the gene encoding a sugar efflux transporter for intercellular exchange (PF03083), which, together with proteins containing cell wall binding repeat domains (PF01473) might be involved in that process. Ultrastructural studies ultimately will be required to better understand how the oral TM7 ectobionts interact with their hosts.

Reverse genomics for other "unculturable" microbial lineages. While the success of using two effective antibodies against TM7/Saccharibacteria is unlikely a result of any special features of its proteome, we further sought to test and expand this approach to a microbial lineage that currently has no cultured representatives. We selected candidate bacterial phylum SR1/"Absconditabacteria", organisms present at low abundance in a variety of environments including in the human oral cavity. SR1 bacteria have highly reduced genomes, indicating likely dependence on interacting organism(s), and have recoded the opal stop codon (UGA) for encoding glycine (32). Based on single cell genomic data of human oral SR1 HOT345, we applied the same search criteria for conserved membrane proteins with extracellular domains. We selected a predicted surface exposed transglycosylase, also annotated as surface antigen involved in cell wall biogenesis (Genbank accession RAL55607). The protein has one predicted transmembrane helix and a ~300 amino acids globular extracellular domain. For this protein target we selected that entire exposed domain for which gene synthesis, protein expression and purification followed by antibody production were performed commercially. The resulting antibody has been effective in capturing live SR1 cells from human oral samples. Culture enrichments as well as analyses on the binding specificity of that antibody on samples from various environments are under current investigation.

Discussion

Genome sequences have greatly expanded our understanding of microbial diversity and evolution. However, the absence of cultured representatives of many

lineages still hinders our ability to study the roles numerous lineages of Bacteria and Archaea play in their environments and how they interact with others. Here we have leveraged single cell genomic data to selectively isolate, using engineered antibodies, cells of diverse species of TM7/Saccharibacteria for sequencing and cultivation along with their interacting *Actinobacteria* hosts. As many bacterial and archaeal lineages have no cultured representatives and are inferred based on their genomic sequence to depend on other members of their communities, cultivation and partnership characterization are essential. The causes of microbial "uncultivability" are numerous, and includes requirements for specific nutrients and growth factors produced by other species, strict interspecies interactions, slow growth, competition/inhibition, and dormancy (221, 222). Traditional microbial cultivation approaches are stochastic or rely on selection for specific physiological traits. Genomic inferences have occasionally been instrumental in enrichment and isolation of several organisms but depend on identifying such traits, which rarely allow for selection in complex enrichments and therefore are not universally applicable. Physical retrieval of single live cells of target microbial taxa from complex communities enables a high throughput search for optimal cultivation parameters and, conceptually, could be applicable to any organism for which a suitable surface target can be identified. While not every organism may be immediately culturable, single cell genomics data from such target cells would further enable population studies, specific physiological inferences and may ultimately lead to successful cultivation. Applying and further developing this strategy to other yet-uncultured taxa could begin to fill the many cultivation gaps in the tree of life.

Methods

Ethics Statement. Human subjects recruitment and sampling protocols were approved by the Ohio State University Institutional Review Board and by the Oak Ridge Site-Wide Institutional Review Board. Written informed consent was obtained from all participants. The use of commercial custom antibodies was approved by the Oak Ridge National Laboratory Animal Care and Use Committee.

Sample Collection and Processing. Saliva samples were self-collected by healthy donors by passive drool, using a SalivaBio collection aid (Salimetrics LLC,

Carlsbad, CA, USA), in sterile cryovials. 1-2 mL saliva were collected at least 2 hours after eating or drinking, kept at room temperature and processed immediately or within 2 hours from collection. Recurring samples were collected from the same donors as needed for individual experiments. Gingival crevicular fluid from patients with periodontitis was collected from affected teeth using sterile endodontic paper points that were placed immediately in sterile vials with 1 mL reduced dental transport media (Anaerobe Systems, Morgan Hill, CA, USA), transported to the lab and further processed within 24 hours. Patients with periodontitis were sampled once.

Oral microbiota samples were diluted 1:5 in dental transport media (L^{-1} distilled water): Na thioglycollate 1.0 g, Na_2HPO_4 1.15 g, NaCl 3.0 g, KCl 0.2 g, KH_2PO_4 0.2g, $MgSO_4$ x $7H_2O$ 0.1 g, L-cysteine HCl 0.5g, pH 7.3), mixed by vortexing for 1 minute then centrifuged at 1000 x g for 2 minutes to remove large particles. The supernatant was passed through 10-μm CellTrics filters (Sysmex, USA) and the cells were collected by centrifugation (12,000 x g for 15 minutes), followed by resuspension in 1 ml PBS. An aliquot was used for total bacterial DNA extraction using proteinase K digestion and phenol-chloroform extraction (164). For fixation, an aliquot was fixed by addition of an equal volume of ice cold, 100% ethanol and stored at -20^{o}C overnight. For live cell sorting, the processed microbiota samples were used immediately.

TM7a genome mining, epitope selection and antibody generation. For selection of candidate proteins to serve as immunogens, we used the single cell genomic data generated by Marcy et al. (28), the only oral TM7 genomes available at the time we initiated this work. We used the TM7a (6806 protein genes) and TM7c datasets (590 protein genes), based on the annotations available in IMG (https://img.jgi.doe.gov)(223), with the expectation that some of the genes may represent organisms other than TM7 (44). We first selected genes encoding proteins associated with membrane processes, based on COG categories (transport functions, cell wall/membrane/envelope biogenesis, extracellular structures), using the IMG assignments. The corresponding proteins were analyzed for the presence of transmembrane anchoring helical domains using TMHMM v.2.0 (224). From the list of proteins with such domains, we eliminated the ones annotated as hypothetical as well as those shorter than 100 amino acids, as we aimed for final selection of relatively large proteins with known function. The proteins that passed those

criteria (117 combined between TM7a and TM7c) were analyzed by blastp against the proteomes of all oral bacteria downloaded from the Human Oral Microbiome Database (http://www.homd.org/). Forty eight of the proteins had high sequence similarity (>90%) to proteins from species of *Leptotrichia* (*Fusobacteria*), indicating they are likely contaminants, as previously reported (44). We then individually analyzed the remaining sequences for number of transmembrane domains, number and size of predicted extracellular domains, availability of three-dimensional structures for homologues from other bacteria, and experimental data related to abundance or prior use as immunogens. The top pick was penicillin-binding protein 2 (PBP2), which belongs to a class of proteins with important roles in bacterial peptidoglycan synthesis (225). Both high resolution crystal structure data (210) and multiple reports that anti-PBP2 antibodies can bind to the cell surface of pathogenic bacteria (226, 227) were available. The closest homologues to the TM7a PBP2 present in genomes of human oral bacteria had 30-40% pairwise sequence identity (species of *Selenomonas*, *Leptotrichia*, *Fusobacterium*, *Streptococcus*). To test the applicability of the approach with a protein for which 3D structure or antigenic data were not available, we also selected CpsC as a target, as cellular studies indicated a cell surface-exposed domain (228, 229). CpsC belongs to the polysaccharide co-polymerases (PCPs) superfamily and is involved in capsular polysaccharide biosynthesis (230). We found no homologues in the oral bacterial genomes with a pairwise identity to TM7a CpsC over 30%.

For selecting peptides to serve as antigens, we analyzed the two protein sequences for antigenic regions and peptide hydrophilicity using the online servers ABCpred (231), BepiPred (232), IEBD Analysis Resource (http://tools.immuneepitope.org/bcell) and Bio-Synthesis peptide design tools (https://www.biosyn.com), using a 16-18 amino acid-peptide threshold. We used Bio-Synthesis (Lewisville, TX, USA) for peptide synthesis and antibody production. For PBP2, we identified 7 peptides that could serve as antigens and for CpsC we identified four peptides (**Figure A4.1**). We next analyzed the location of the 7 peptides from TM7a on the 3D structure of the protein based on the homologous regions of the *E. coli* PBPB structure (PDB 3vma) using PyMOL (https://pymol.org/2/). Further details on structural modeling and visualization are presented in the comparative structural analysis section.

The selected antigen peptide (SGNIKYAKQRQKTVLTRMV) was one of the three peptides with the most exposed residues. In the case of CpsC, because no experimental 3D structure was available, we selected an antigenic peptide (ESAIQEFKEQSKSLYGNS) that is in a helical region of the predicted extracellular domain (**Figure A4.1**).

The two selected peptides were commercially synthesized by Bio-Synthesis (Lewisville, TX, USA), with cysteine added at N (CpsC) or C terminus (PBP2), purified and injected in rabbits. For antisera production, Bio-Synthesis used a 70-day protocol, with 5 booster doses and 4 test bleeds followed by ELISA. We received serum batches confirmed by ELISA to be reactive against the antigenic peptides. For IgG purification we used HiTrap Protein A HP antibody purification columns (GE Bio-Sciences (Pittsburgh, PA, USA) according to manufacturer's protocol. For fluorescent IgG labelling of 100 μg IgG aliquots we used the Alexa Fluor™ 488 Antibody Labeling Kit from ThermoFisher (A20181) according to manufacturer's instructions.

Immunofluorescent labeling of oral microbiota samples and flow cytometry. Aliquots of oral microbiota samples, processed as described above and suspended in PBS were blocked for 30 minutes with 5% goat serum and a mixture of IgGs (1 μg/mL each rabbit anti-*Ignicoccus* IgG (233), anti-*Clostridium* IgG (234) and human IgG, I4506, Sigma, St. Louis MO, USA) for blocking potential non-specific antibody-binding sites on some oral bacteria. We also used 5% rabbit pre-bleed serum instead of the goat serum, with no observed differences. Following blocking, fluorescently labelled anti-PBP2 or anti-CpsC IgG were added to the samples at 0.1-1 μg/mL. It was important to centrifuge the antibody solution (15,000 x g for 5 minutes) and carefully recover supernatant, prior to addition to the microbiota samples in order to reduce the background of fluorescent particles. Antibody labelling was conducted for 1 hour at room temperature, followed by either centrifugation (12,000 x g for 15 minutes) and resuspension in 2 mL PBS supplemented with 1% serum or directly diluting in 2 ml PBS with 1% serum (mainly for live cell sorting for cultivation experiments). After 30 minutes, the samples were analyzed by flow cytometry. Negative controls included unstained oral sample (to check for auto-fluorescence), and samples processed in parallel with the oral sample but without microbiota (to check for fluorescent antibody precipitates).

For flow cytometry and cell sorting, we used a Cytopeia/BD Influx Model 208S (Cytopeia, Seattle, now BD, San Jose, CA), in a biological safety cabinet (BSC) designed specifically for the Influx cell sorter (Baker Co). Preparation of the flow cytometer for single cell sorting for genomic amplification and sequencing was as we and other have previously described (33, 235, 236). For sorting, we used a 70-μm nozzle and the 488-nm laser for forward-side scatter (FSC-SSC) analyses as well as fluorescence detection and cell sorting trigger. Gating parameters were selected based on FSC-SSC and fluoresce levels for each experiment and sample. A very low level of highly autofluorescent particles (<0.01%, RFU>10^3) were observed in all samples in the absence of staining with the antibody and gating parameters were set to avoid sorting such particles. For genomic and community diversity analyses single or pools of fluorescent particles were deposited in 3 μL of DNA-free, Tris–HCl (10mM pH 8.0), 1mM EDTA (TE) in individual wells of 96 well plates. We observed sufficient fluorescence signal with primary labelled antibody preparations. However, if the signal would be low, using a secondary labelled antibody (e.g. goat anti-rabbit IgG) would provide signal amplification.

Isolation and cultivation of oral TM7 bacteria. For cultivation, we used saliva samples collected recurrently from three healthy donors and processed them as detailed above. Antibody staining was performed as already described. For the initial cultivation tests, in 200 μL liquid media in 96-well plates, we used as basal media commercial or home-made brain heart infusion (BHI, Difco), Oral Treponeme Enrichment Broth (OTEB, Anaerobe Systems, Morgan Hill, CA, USA), MTGE Broth (Anaerobe Systems), Tryptic Soy broth (TSB, Difco) supplemented with various additional factors (ATCC vitamins, trace minerals, clarified filtered saliva, pig gastric mucin, sugars, amino acids and nucleobases, N-acetyl muramic acid, N-acetyl glucosamine, pyruvate). Following flow sorting, performed as already described, the plates were incubated in an anaerobic chamber (COY, Grass Lake, MI, USA) at 37°C, under 85%N_2, 10%CO_2, 5%H_2 or in anaerobic jars to which we added air via a syringe port to reach 2% O_2. To test for TM7 growth, after 48-72 hours we retrieved 100 μl culture using multichannel pipettes and vacuum-filtered through 0.2 μm filter 96-well plates (Millipore, Burlington MA, USA). The cells were washed with 200 μL PBS. Then, 20 μL TE were added to the filter. The plate was placed on an orbital shaker at 100 rpm for 5 minutes at room temperature. 10 μL of

the cell suspension was then recovered in 96 well PCR plates, to which we added 10μL lysis solution (0.13 M KOH, 3.3 mM EDTA pH 8.0). The plate was covered with adhesive foil and heated in a PCR machine at 95°C for 5 min. The plate was immediately cooled on ice after which, 10 μL neutralization buffer (0.13M HCl, 0.42M Tris pH 7.0, 0.18M Tris 8.0) was added. We then used 1.5-2 μL of the lysate in PCR reactions with universal and TM7 primers, as detailed earlier. The presence of TM7 in PCR positive cultures were further confirmed by Sanger sequencing and the cultures were transferred into fresh medium under the original growth conditions. We initially obtained TM7 positive cultures using BHI, OTEB, MTGE and TSB, primarily under microaerobic conditions. In liquid microcultures, in addition to TM7 HOT352 (Group 1), we also identified TM7 HOT347 (Group 1) and HOT870 (Group 6). However, none of positive microcultures were pure and contained various other bacteria (*Streptococcus*, *Fusobacterium, Campylobacter, Actinomyces, Propionibacterium* and *Veillonella*). While TM7 survived in those liquid cultures for several passages, we have not identified the hosts for HOT347 and HOT870. Because liquid cultures in 96 well plates were difficult to monitor and to prevent overgrowth, we used cell deposition and cultivation on solid media for final isolation experiments.

For cultivation on solid medium we used DMM-agar (237) supplemented with 2% glucose, 0.1g/L casamino acids and 10% sheep blood as well as BHI-agar supplemented with 10% sheep blood. Single cells were deposited by flow sorting in a 10x10 array on each plate, followed by incubation at 37ºC under anaerobic or microaerobic conditions. Visible colonies formed after 3-5 days of incubation (**Figure A4.3**) were sampled with sterile tips and analyzed by direct PCR using primers 27FM-510R (TM7 specific) and 27FM-1492R (universal bacteria). Secondary screening for TM7 bacteria was also performed using primers 910F (5'CATAAAGGAATTGACGGGGAC 3') and 1177R (5'-GACATCATCCCCTCCTTCC-3') under the conditions 95°C for 2 minutes, 32-34 cycles of 95°C for 30 seconds, 61°C for 30 seconds, 72°C for 1 minute, followed by 72°C for 5 minutes. Single band products were directly sequenced at Eurofins Genomics (Louisville, KY, USA) using Sanger chemistry. For colonies that contained other bacteria (*Streptococcus*, *Veilonella*), the original TM7-positive colony was repeatedly streaked on fresh plates until colonies containing only TM7 and a host actinobacterium were obtained.

For TM7 HOT346, a pure co-culture with *Cellulosimicrobium cellulans* was obtained following a second round of antibody staining and sorting (**Figure A4.3**). TM7 enrichments and co-cultures with hosts could be maintained as freezer stocks in media containing 10% DMSO at -80°C. For further physiological characterization of TM7 metabolism potential we supplemented solid (with 10% blood) and liquid (without blood) BHI media with 2-5mM of ribose, lactate, pyruvate, or phosphoenolpyruvate at 37°C or 39°C in duplicate and screened for TM7 relative abundance after 2 and 5 days of incubation and a subsequent passage in the same media, as described above. Compared to a BHI only control, no significant differences were observed.

For microscopic characterization of co-cultures, cells were fixed in 50% ethanol-PBS overnight at -20°C. For fluorescence in situ hybridization we used cell immobilization and dehydration on gelatin-coated slides as described (238), with 30% formamide in the hybridization buffer. The probes we used were EUB338 (5'GCTGCCTCCCGTAGGAGT3') (universal bacteria), labeled with Alexa536 or Alexa647 and TM7-567 (5'CCTACGCAACTCTTTACGCC3') labeled with Alexa488. In some experiments we also labelled cells with CellBrite Fix 640 Membrane Stain (Biotium) prior to fixation. Following hybridization and washing, the slides were covered with EverBrite Mounting Medium (Biotium) and analyzed by epifluorescence microscopy using a Zeiss AxioImager microscope. Color photographs were taken using a Canon 7D digital SLR and were merged in Adobe Photoshop. For immunofluorescence microscopy, in addition to the anti-PBP2-CspC antibody mix (labelled with Alexa 488), we also used a broad anti-*Actinomyces* rat IgG, generated by immunizing rats with a mix of *Actinomyces* species (*A. odontoliticus, A. israelii, A. naeslundii, A. gerencseriae*) and labeled with Alexa 647. The slides were first blocked with the same cocktail used in antibody staining for flow cytometry and stained with a mix of anti-PBP2-CspC and anti-*Actinomyces* antibodies for 60 minutes. After washing with PBS five times, the slides were mounted and visualized as above.

Genomic amplification, taxonomic analysis and single cell genome sequencing. Sorted single cells or cell pools were stored frozen at -80ºC until genomic DNA amplification. For amplification by MDA we used a protocol that we and others have described in detail previously (33, 235, 236). Single cells amplified genomes (SAGs) were

screened for the presence of TM7 by PCR using the primers 27FM (5'AGAGTTTGATYMTGGCTCAG-3') and 510R (5'-CTCTTTACGCCCAGTCAC-3') by denaturation at 94°C for 3 minutes followed by 32 cycles of amplification (95°C for 30 seconds, 62°C for 30 seconds, 72°C for 30 seconds) and a final extension at 72°C for 5 minutes. PCR products were sequenced on an ABI3730 DNA Analyzer (Applied Biosystems), at the University of Tennessee Sequencing Facility or at Eurofins Genomics (Louisville, KY, USA), using the same primers. Resulting chromatograms were manually edited for low quality regions using Geneious (https://www.geneious.com/) and sequences were used for phylogenetic identifications using the HOMD SSU rRNA database (http://www.homd.org/). Samples for which the chromatograms were not homogeneous and that denoted potential heterogeneous template DNA (more than one sorted cell) were not further used. While a mixed chromatogram may indicate the presence of a symbiont of TM7, because they could not be unambiguously distinguished from either DNA contamination or non-specifically adherent cells co-sorted with TM7, so we did not pursue characterization of mixed SAGs.

To characterize the microbial diversity in cell populations sorted following antibody staining (10-100 cells), genomic amplification by MDA was conducted the same way as for single cells. The amplification product was first purified using a Zymo genomic DNA cleanup kit (Zymo Research, Irvine CA, USA). To determine the microbial composition of the original oral microbiota samples we used gDNA purified from the saliva and the subgingival samples. Microbial diversity was determined by amplification of the hypervariable V4 region of SSU rRNA gene followed by MiSeq amplicon sequencing. We followed the protocol described by Lundberg et al. (239), with a modification in the universal primers. Specifically, because the canonical 515F primer sequence (5'GTGCCAGCMGCCGCGGTAA) does not recognize TM7 rRNA, we supplemented the PCR reaction mix with 5% of a 515F-derivative that can bind TM7 rRNA (5'GTGCCAGCMGCCGCGGTCA). The reverse primer, 806R (5'GGACTACHVGGGTWTCTAAT) can recognize the full diversity of oral bacterial SSU rRNA. The purified amplicons were sequenced on the MiSeq platform (Illumina, San Diego CA, USA) on v2 or v3 sequencing chips using 2x250 nt reads. Read processing and microbial diversity analysis was performed using the software programs cutadapt

(240), usearch (241) and QIIME (242) as well as the HOMD SSU rRNA taxonomy database and following the commands and parameters listed in the “MySeq sequence analysis” supplementary file. To determine if the 254 nt V4 amplicon provides sufficient resolution to enable distinguishing the different human oral TM7 phylotypes, we determined the pairwise sequence identity between the 13 main types of oral TM7 OTUs and performed phylogenetic analysis including the TM7 OTUs identified based on sequencing the V4 amplicon (TM7HOTs_OTUs V4.fasta supplementary file). As shown in **Figure A4.2**, with the exception of 2 OTU pairs (HOT346-HOT869 and HOT352-HOT952), all other TM7 could be readily distinguished in MiSeq amplicons, with pairwise identity values bellow 98% and sufficient phylogenetic resolution.

Purified DNAs of TM7 SAGs selected for genomic sequencing were used to generate libraries and were sequenced on Illumina HiSeq (100 nt paired reads) at the Hudson Alpha Institute for Biotechnology (Huntsville, AL, USA) or on MiSeq (250 nt paired reads) at ORNL. Approximately 100 million reads were obtained for each SAG sequenced on HiSeq and 10 million for the MiSeq SAGs.

SAG assembly, annotation and comparative genomic analysis. For individual assembly, the reads for each SAG were trimmed using Trim Galore 0.4 (https://github.com/FelixKrueger/TrimGalore) and assembled using SPADES (v3.9.0)(243). Removal of non-TM7 contigs (human and other bacterial sequences) was performed by a combination of GC-content-based binning, and Emergent Self-Organizing Maps (ESOM) based on kmer frequency (244). Further contamination was removed using DNA and protein BLAST and single-copy gene analysis (245, 246). Lastly, further assembly in Geneious (v8-10)(179) consolidated some contigs and the final cleaned genomes were analyzed with Prokka (v1.11)(247) for gene prediction and annotation. Completion and contamination estimates were calculated using CheckM (v1.0.11)(246) using the lineage workflow and a manually adapted gene-list workflow containing 42 single copy genes conserved in CPR Bacteria (248) (**Table S4.1, Figure A4.4**). Pfam (v31)(249), eggNOG (v4.5.1)(128) and Interproscan (v5.27-66.0)(250) databases were used for protein assignment to families and orthologous groups. Anvi’o (v5.1) was used for core genome analysis and identification of conserved COGs following the pangenomics pipeline with mapping of eggNOGs (217, 218) (**Table S4.1** and **Figure 4.4**).

NCBI, IMG and Geneious were used for basic homology searches and for obtaining genomic data for further metabolic inference analyses. SSU rRNA and concatenated single-copy protein gene trees were generated in Geneious using RAxML, with estimation of all free parameters and 100 bootstraps **(Figure A4.5)**. The TM7 SAGs sequences and associated metadata were deposited in Genbank under Bioproject PRJNA472398. SSU rRNA sequence alignments that include reference strains as well as detected phylotypes, SAGs and isolates obtained in this study are provided as supplemental files "TM7_16S.fasta" and "Actinomyces_16S.fasta".

Protein structure modeling. The glycosyltransferase domain of PBP2 from TM7a (residues 95-278) and TM7x (residues 145-334) was modeled using a combination of homology modeling and coevolution-based residue-residue contact prediction. We used the GREMLIN web server (251-253) to identify structural templates, predict residue contacts based on coevolution analysis, and generate contact restraints for Rosetta structural modeling (211). Briefly, the server uses HHblits (254) and to generate a multiple sequence alignment (MSA), which is then filtered to include only sequences that cover at least 75% of the query sequence, have fewer than 75% gaps, and have a maximum mutual sequence identity of 90%. From this MSA, GREMLIN learns a global statistical model that accounts for both sequence conservation and coevolution. Next, HHsearch (255) was used to identify potential templates for homology modeling from the Protein Data Bank (PDB)(256). More than 9,300 sequences of PBP2 homologs were identified from the sequence database search. For the ~190 residue proteins this number corresponds to a sequence/length ratio of ~51, indicating the robustness of the predicted residue contacts.

For both TM7a and TM7x, the following five templates identified by HHsearch were used as templates for modeling PBP2: 3VMA (257), which is missing part of jaw subdomain, 3UDF (258) (also missing the jaw subdomain), 2OQO (259), 2OLV (210), and 3VMT(260) (**Figure A4.7**). Although multiple template structures are available, we included coevolution information to provide additional benefits compared to standard homology modeling. We used map_align (https://github.com/gjoni/map_align) to align the contact maps predicted from the coevolution analysis to selected templates identified by HHsearch. The initial threaded models generated by map_align were then used as input

for Rosetta hybrid modeling, and all models were given weights of 1.0. Fragment files were obtained with the Robetta server (261). We generated 10,000 models with both sigmoidal and bounded restraints during coarse-grained sampling and only sigmoidal restraints during full-atom refinement (253). Models were ranked on the basis of the sum of their Rosetta energy and the restraint score. The final models are shown in **Figure A4.7.** PyMOL (https://pymol.org/2/) and **pdb** format files of the models are also provided as supplementary data (PBP_structural_alignment.pse, tm7a_S_0876_ma1.1.pdb and tm7x_S_0959_ma1.5.pdb).

CpsC (residues 1-197) was modeled with Rosetta using coevolution-derived residue-residue contact restraints. More than 3,700 sequences obtained with the GREMLIN webserver were used to perform the coevolution analysis. A PyMOL file of the 3D model structure is provided as supplementary data (cpsc_epitope.pse).

We analyzed the SAGs for the presence of the genes encoding the two proteins used as epitopes for raising antibodies. For PBP2, we found the gene in SAGs representing all of the TM7 lineages isolated by antibody capture. For Cbp2, we found the gene in only three of the lineages. This gene is also missing in the complete genome of TM7x, so the Cbp2 protein is not present across all TM7. The level of sequence similarity at the epitopes level varied between the different TM7 as compared to the original TM7a sequence but there were multiple conserved amino acids across all lineages. To better understand the structural basis of epitope recognition in the different TM7 PBP2, we modeled the three-dimensional structure of the TM7a/HOT348 and TM7x/HOT952 proteins based on crystal structures available for several homologues (**Figure A4.7**). In addition to structural templates, we included coevolutionary information in the form of residue-residue contacts (252) predicted from a deep multiple sequence alignment to improve the structural models. Because we were primarily interested in identifying structural determinants of epitope recognition, we generated structural models only of the extracellular, N-terminal glycosyltransferase domain (residues 95-278 in the TM7a PBP2 sequence). The final models share a similar overall fold and structure between TM7a and TM7x PBP2 even though they are only 37% identical. The α-helix containing most of the epitope is located on the surface of the protein, with some of the conserved residues buried and others exposed or partially exposed (**Figure 4.1b**). A

similar epitope topology was observed in the Cbp2 protein model (**Figure 4.1c**). The exposed residues include multiple polar and charged amino acids, which may be responsible for antibody recognition and binding. It is important to note that although the sequence similarity in the epitope region is as low as 35-50 % across the TM7 lineages, IgG variants were able to bind their target on the cell surface. Although the specificity breadth cannot be predicted, especially when only a single target sequence is available (as was the case with TM7a), polyclonal antibodies can therefore recognize multiple related species and increase the diversity in targeted isolation of microbial cells. Broader specificity can also result in cross reactivity with more distant, non-target organisms, potentially explaining some of the additional taxa we observed in pooled cell populations.

Data Availability. Curated TM7 SAG sequences were deposited in GenBank under the Bioproject PRJNA472398. Aligned SSU rRNA sequences are provided as supplementary files. Raw sequence datafiles are available from the authors upon request.

Acknowledgements and Funding. We thank S. Allman, S. Kauffman. L. Sukharnikov and M. Robeson for technical assistance. Support for this work was provided by grants R56DE021567 and R01DE024463 from the National Institute of Dental and Craniofacial Research of the U.S. National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. J.M.P. was supported by the Laboratory Directed Research and Development program at Oak Ridge National Laboratory (ORNL), which is managed by UT-Battelle, LLC for the U.S. Department of Energy under Contract No. DE-AC05-00OR22725. S.J.C. was supported by a National Science Foundation Graduate Research Fellowship under Grant No. 2017219379. This work used resources of the Compute and Data Environment for Science (CADES) at ORNL.

Chapter Four Appendix

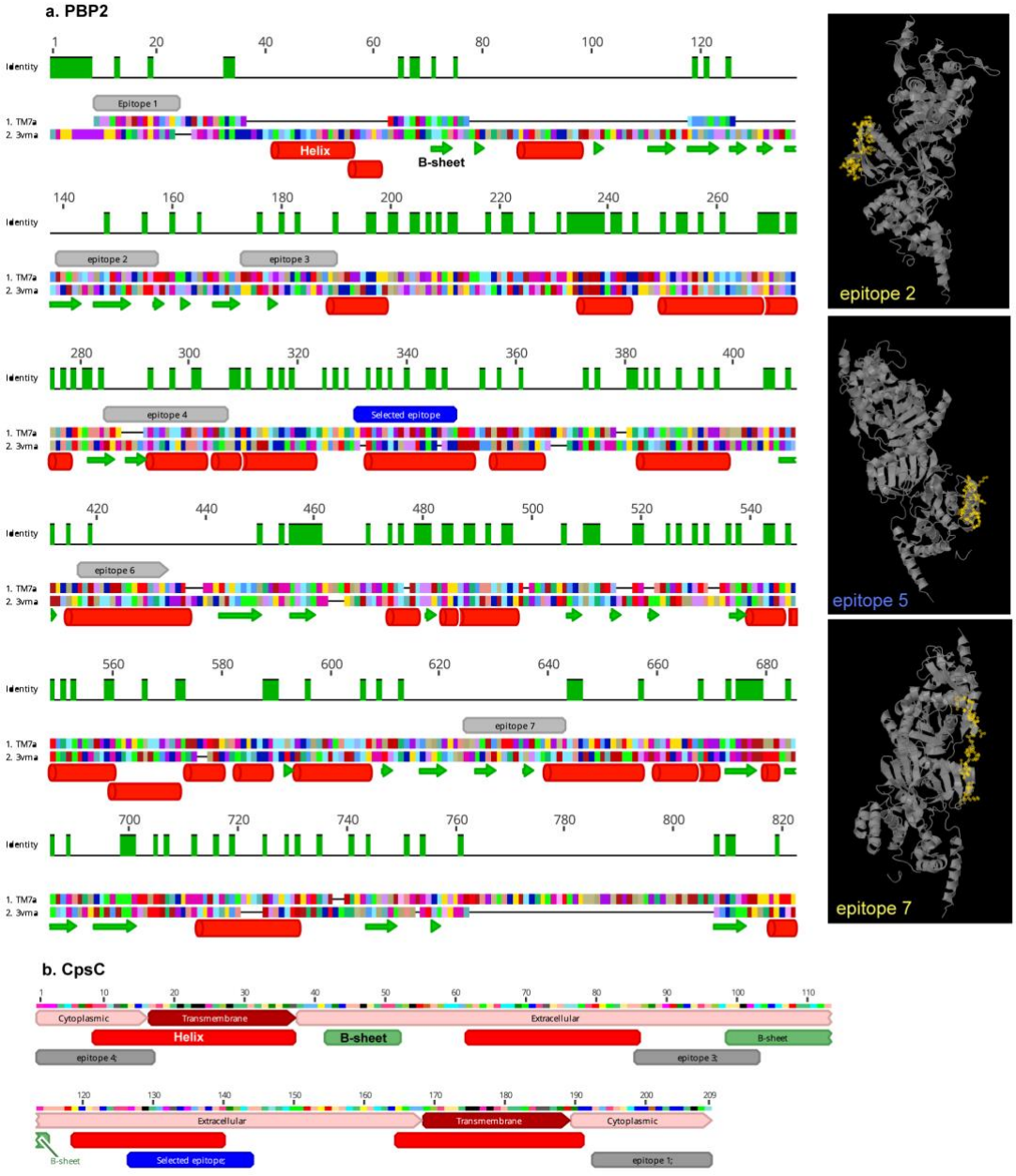


Figure A4.1. Antigenic peptide selection. **a.** Alignment of TM7a PBP2 with the *E. coli* homologue based on 3D structure (3vma). Alpha helices are in red, beta sheets are shown as green arrows. The location of the predicted antigenic peptides (epitope 1-7) are in gray, and the selected epitope for antibody production is in blue. Three-dimensional structure (3vma) on the right shows the location of epitopes 2, 5 (selected), and 7 in gold color. **b.** Predicted secondary structure and topology of TM7a CpsC. The four predicted antigenic peptides (epitope 1-4) are indicated, the selected one is in blue.

a

b

	HOT869 [G1]	HOT349 [G1]	HOT346 [G1]	HOT347 [G1]	HOT352 [G1]	HOT952 [G1]	HOT488 [G1]	HOT348 [G1]	HOT870 [G6]	HOT350 [G2]	HOT355 [G4]	HOT351 [G3]
HOT349 [G-1]	95.7											
HOT346 [G-1]	99.2	96.5										
HOT347 [G-1]	96.9	94.9	97.6									
HOT352 [G-1]	92.9	92.5	93.7	92.9								
HOT952 [G-1]	93.3	93.7	94.1	94.1	98.8							
HOT488 [G-1]	92.5	91.4	93.3	93.3	94.9	96.1						
HOT348 [G-1]	92.2	92.5	92.9	92.9	97.6	97.3	93.3					
HOT870 [G-6]	89.4	89	89.4	89.8	91.4	90.2	87.1	91.4				
HOT350 [G-2]	94.9	92.2	95.7	95.7	94.1	92.9	92.2	94.1	93.3			
HOT355 [G-4]	89.4	88.2	89.4	89	89.8	89.4	87.1	91	91	92.2		
HOT351 [G-3]	87.1	85.9	87.1	85.5	88.2	87.1	87.1	86.7	85.9	87.5	87.5	
HOT356 [G-5]	82.4	82	82.4	82.4	82	80.8	80	82	84.7	84.7	83.5	77.3

Figure A4.2. SSU rRNA V4 taxonomic resolution. **a.** Phylogenetic tree (Jukes Cantor-corrected distances) of human oral TM7 reference sequences (HOTs) and MiSeq OTUs. Most OTUs can be unambiguously assigned to known HOT types. **b.** Pairwise identity matrix for human oral TM7 HOTs (SSU rRNA V4 region).

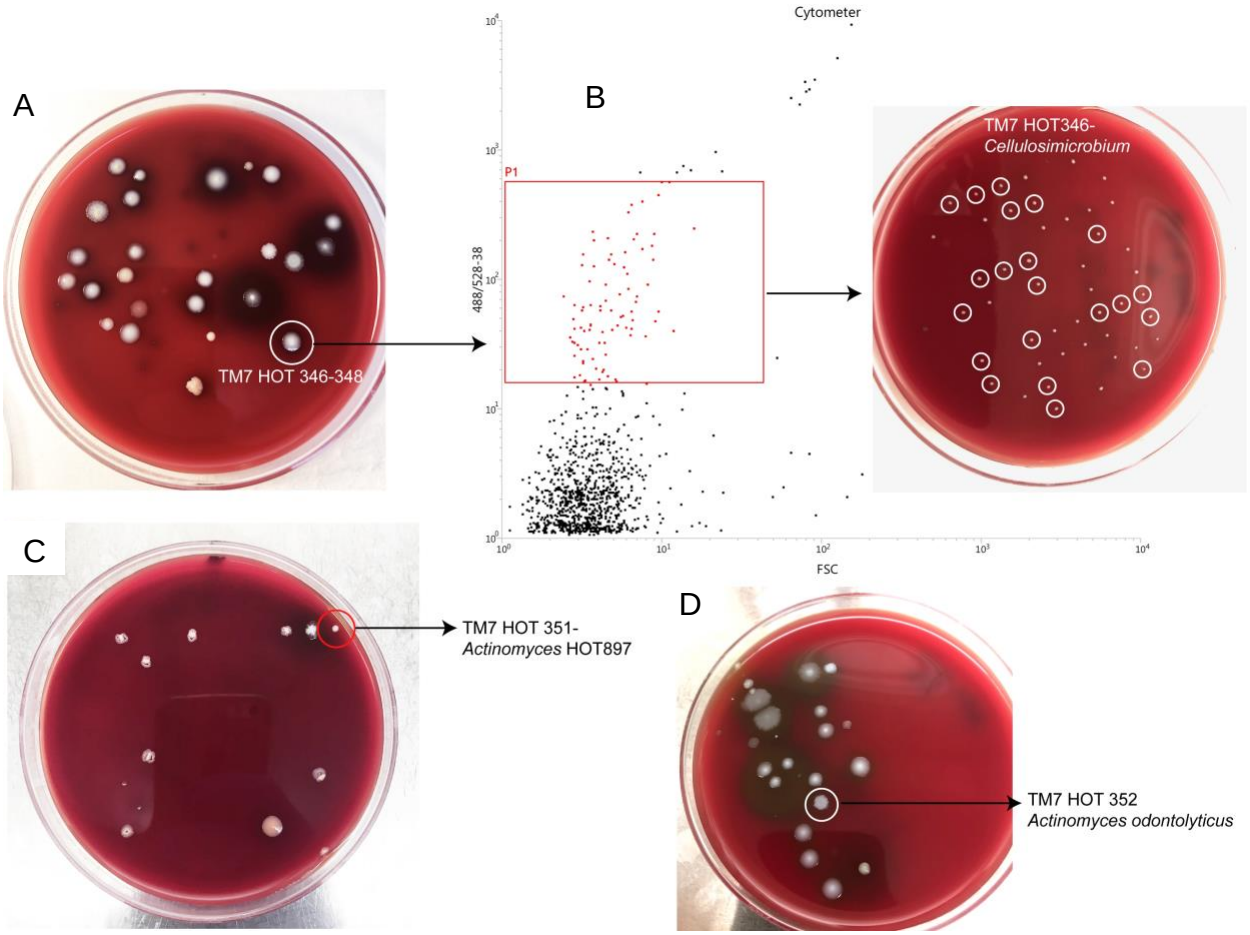


Figure A4.3. TM7-host colonies following single cell sorting and incubation at 37°C. **a.** Cells from primary TM7 HOT346-348 colony were re-stained with the antibody and **b.** sorted based on the gate shown in the flowgram on BHI-blood agar plate. The circled colonies contained a pure TM7 HOT346-*Cellulosimicrobium* co-culture. **c.** Primary colony of pure TM7 HOT351-*Actinomyces* sp. HOT897 co-culture. **d.** Primary colony of TM7 HOT352/952-*Actinomyces odontolyticus*, also containing *Streptococcus oris*.

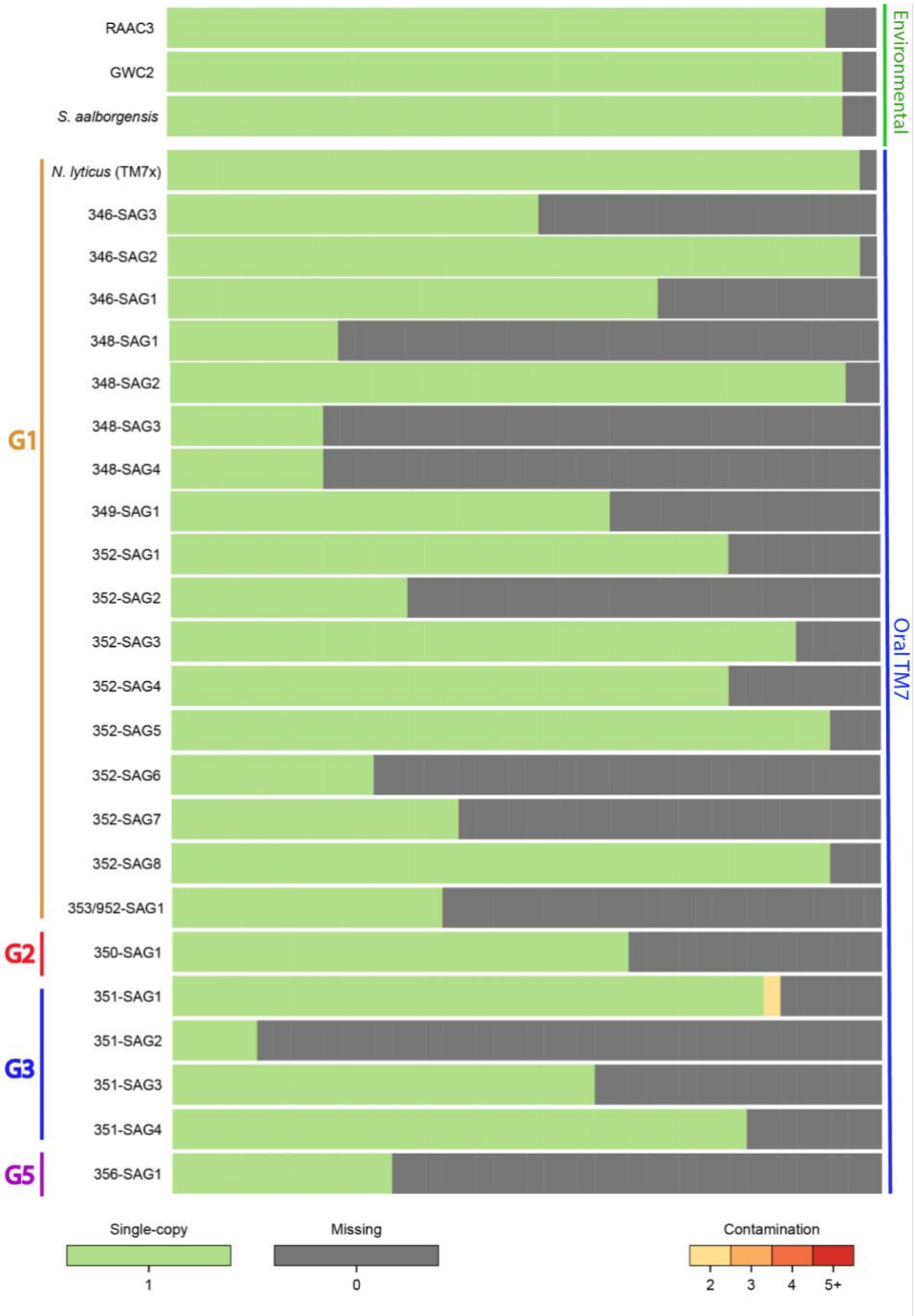


Figure A4.4. Completion level and contamination analysis of oral TM7 SAGs. Analysis was based on CheckM using 42 single copy genes conserved in bacteria from the "Candidate Phyla Radiation". Four closed TM7 genomes (TM7x, RAAC3, GWC2, and *S. aalborgensis*) were also analyzed for comparison. Each row represents a genome with colored bars representing the presence/absence/copies of single-copy genes. SAGs are organized based on which phylogenetic group they belong to.

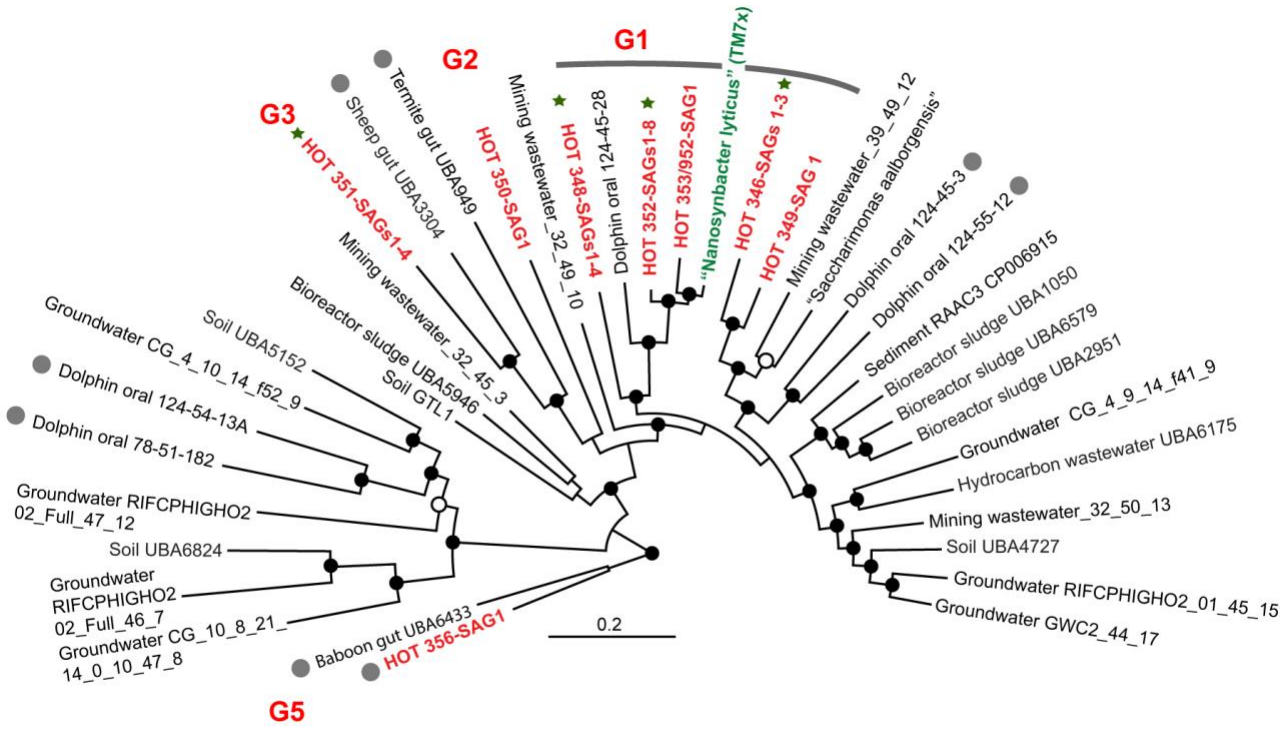


Figure A4.5. Maximum likelihood phylogeny of concatenated proteins from TM7 bacteria based on SAGs, MAGs, and TM7x. SAGs in red were sequenced as part of this study. Other host-associated TM7 are indicated by gray dots. Black dots at nodes indicate bootstrap support >80, white circles are support values of 50-80%. Scale bar is inferred number of substitutions per site.

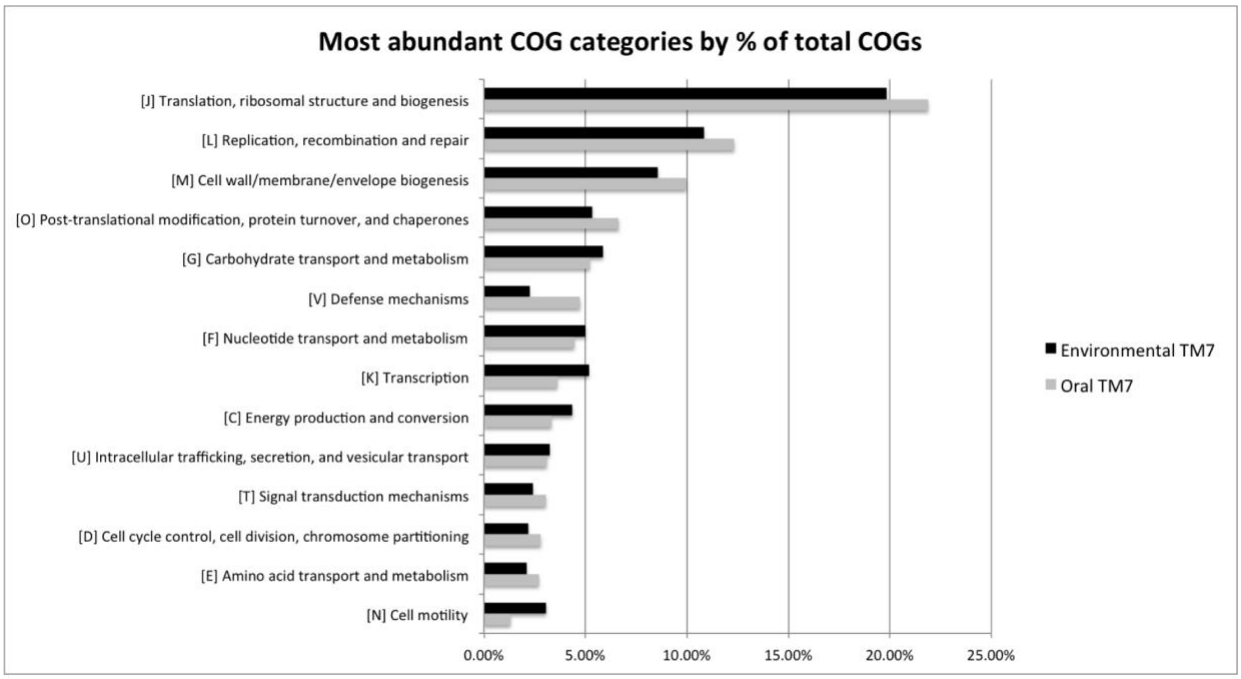


Figure A4.6. Abundance of COG categories between environmental and human oral TM7 bacteria.

Figure A4.7. Structure modeling of TM7 PBP. **a.** Structural templates for the GT domain of PBP2 from TM7a and TM7x identified by HHsearch. **b.** Target sequences for structural modeling. Structural models of PBP2 were generated for the regions in black text. The epitope region in TM7a is boxed. **c.** Final models of the glycosyltransferase domain of PBP2 form TM7a (left) and TM7x (right). Residue contacts predicted from coevolution analysis are shown as yellow lines. The models of the GT domain satisfy the coevolution restrains well and show the overall folds of the two PBP variants are similar.

a

PDB	Organism	Resolution (Å)	% Identity	% Coverage
TM7a				
3vma, A	Escherichia coli K1	2.16	36	89
3udf, A	Acinetobacter baumannii	1.7	37	66
2oqo, A	Aquifex aeolicus VF5	2.1	40	95
2olv, B	Staphylococcus aureus	2.8	40	95
3vmt, A	Staphylococcus aureus	2.3	37	95
TM7x				
3vma, A	Escherichia coli K12	2.16	36	86
3udf, A	Acinetobacter baumannii	1.7	36	64
2oqo, A	Aquifex aeolicus VF5	2.1	42	92
2olv, B	Staphylococcus aureus	2.8	36	92
3vmt, A	Staphylococcus aureus	2.3	38	92

b

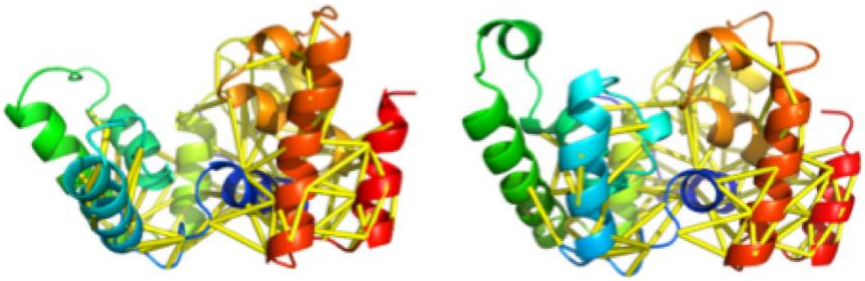
>TM7a

MRRVGKYTKNSKIRSARLTANKYVLWFKNLKAWQKVAVIVGPIAIIMVIVPLITYVYFARDIANKERLMNRNNTGVV
LFDKNGEVIYQTGNARHRDLVKLNDISDYTKKALLSSEDKDFYKHSGVSLTGVVAALYANIITGGKNFGGSTLTQQL
AKNTLLTNQKSFLRKYQELSIAIAIERTYSKDEILEMYLNSVYYGENSFGIHDAAKTYFNKTPAELNLAESAMLIGL
LPAPSAYSPISGNIKYAKQRQKTVLTRMVNNGAITEQQKQDAINTQLSYAPKVSPQDKSPAPHFVEMVMSDLYKKYD
EETLTRSGFQITTSLDLNMQRELQQMVSNNMRSIRLNGGSNASAIAIDPKTGGVRAMVGSYDWNDEKFGKVNMATAR
RQPGSSFKPIYYAEALANGVITPSTILEDAPTNFNGYQPKNALRNYNGDVTVRKALNWSLNIPAVKVMQKLGVSKSV
EAAKRMGVTTLGNTSDYGLSLALGAAEAKLIDMTHVYAAFANEGDQQPIKFIEDMKSKYDERVTIDNNKPKQVISKE
GAYLISNILSDNSARSGMFGSSLNIPGKKVAVKTGTTDDQRDAWTIGYTPDLAVGVWVGNNDNTPMLSGGGIQAGPI
WRSAMQKFSARQSGNPFTQPSGIVERDTCYGTGRLASRSGQNTYKEVYLSTALPSVGCNVKSEEDKKETKTDEVKPK
EETETTPTPRTLPRGQQSSGSNNSENNSSQDTS

>TM7x

MKRKTTINTAGHGSSNKKSPSKRMNLYANLAQKHRTKKDKDARERAEYLATLPKHPVKRFFYRLHPKRLAKYWFSKR
GGLMALKILGVGTLLVLLLIGGMFAYFRKDLDKIRPGELAKRVQTTVTKYYDRNDNLLWEDKGTGNYQLVVEADQIS
DYLKKATVAIEDRDFYKHHGISISGIMRAMLSTASRRQVQGGSTLTQQLVKQVFFADEAGDRSISGVPRKIKEIILA
IEVERMYSKDQILSLYLNESPYGGRRNGAESASQTYFGKHAKDLTLAEAALIASIPQNPTYYNPYNTAGHKALIARQ
HTTLDYMAEQGVISKDEAEKAKKVDILSTIKPQTEQLENIKAPHFLLMVRNQLSKELGESVVGRGGLTIKTTLDWRI
QEKLETEMKSFFATGRPDSVRISNGAATVEDVQTGQIVALVGSRDFNHAGYGQDNAATSYIQPGSTIKSLVFAQLFE
KHDNKQGYGSGTVLSDDNIDSLYGAKLRNWDGKFMGAINIRKALSFSRNIPAVKAAYIIGNGSAKPVVEGIRKMGNH
NYCRQEENAGGYGLGASIGACGTKQTELVNAYGTLARMGVQKNSSSVIEVKNSQGETLKKWKDDGKQVLDSQSAYIV
NDILSDRTTGLHGWMGVNGVRTSAKTGTSDKGSQPKDLWIVNYSPALVMGIWLGNSDTSVIGTSSSNFGMPVVRNVM
SFAHTQVYAKEGKWKSGQWYERPNGIQNVNGELYPSWWDKKQGQSTEKITFDKVSKKKATSCTPDGAKEEIEVTKVI
DPLTKKESLSVPSGYDANAEDDVHKCDDAKPQIGAVSVTGSGKNYTINVDVSAGTWGLAAIEISVDGKSVKSAEISS
GGKQSATVEIETAGTHTVSVTVRDSAYYTATSSSSFKSKDS

c



TM7a TM7x

CHAPTER FIVE – Summary, Conclusions, and Future Directions

Periodontitis is the leading cause of tooth loss worldwide and is hypothesized to contribute to systemic diseases such as cardiovascular disease and irritable bowel disease (20, 74, 262). With over 770 species level bacterial taxa identified in the human oral cavity, approximately 1/3rd represent uncultivated members (12, 13), their recalcitrance may be due to dependencies on co-evolving bacterial community members yet identified. Some of the periodontal pathobionts have been recognized but many disease-associated lineages remain uncultured (17). Isolation and cultivation of these disease-associated microorganisms is essential to fully understand their role in disease. Isolation allows us to move away from just mere correlation with disease to direct causation so that we can begin to piece together an appropriate model for how such microbiome-associated diseases are acquired. Additionally, discovering new information about these organisms would allow for the development of innovative treatments and possible means of prevention for periodontitis. This dissertation takes a dual approach towards elucidating novel disease-associated lineages of bacteria in the human oral cavity using cultivation-dependent and -independent approaches (**Figure 1.1**).

Sulfate-reducing bacteria (SRB) have been detected in periodontal pockets (55) and culture-independent studies have linked the SRB genus *Desulfobulbus* with progressive periodontal disease in the oral cavity (58, 59). Single-cell genomic analysis was able to probe the initial functions of the little known *Desulfobulbus* genus in the human oral cavity (33), but without a cultured representative, truly understanding the role of this organism in the microbial community remains difficult. **Chapter 2** describes the isolation and characterization of the first species of the *Desulfobulbus* genus from the human oral cavity. With collaborators from the Ohio State College of Dentistry, we received human periodontal samples and enriched for SRB and isolated *Desulfobulbus oralis* sp. nov. In working towards the isolation of *D. oralis*, we obtained a stable co-culture of *D. oralis* and *Fusobacterium nucleatum*, another native oral microbial species. Further physiological characterization identified that *D. oralis* required by-products produced by *F. nucleatum* for growth in pure culture to occur (186). *D. oralis* is the first human-associated representative of its genus and has provided unique insights into the genomic and physiological changes that have occurred as *D. oralis* co-evolved with its human host and to the emergence of pathogenicity. In addition to limited metabolic versatility and

niche specialization, we identified horizontal gene acquisitions from other oral microbiome members, such as toxins like leukotoxin and hemolysins, that can potentially trigger a proinflammatory response in oral epithelial cells. We showed that *D. oralis* is not merely a passive by-stander in disease progression but may actively contribute to disease etiology. This relationship could only be understood after isolation.

Furthermore, understanding the relationship between *D. oralis* and *F. nucleatum* will provide useful insights to aid in the cultivation of more disease-associated bacterial lineages. Other oral microorganisms are dependent on *F. nucleatum* for growth (39, 40) but it remains unknown what molecules produced by *F. nucleatum* stimulate growth Preliminary tests with defined additions of amino acids, vitamins, siderophores, and complex nutrients did not alleviate the dependency of *D. oralis* for *F. nucleatum* filtrate. Additionally, the growth promoting component(s) of the *F. nucleatum* filtrate could not be inactivated with heating or protease treatment, suggesting it is not a protein but rather a small molecule(s) of some kind. **Chapter 3** begins to elucidate the growth promoting characteristics of *F. nucleatum* using a metabolomics approach. Data analysis of mass spectrometry identified the metabolic by-products of *F. nucleatum* growth and we identified what *F. nucleatum* may be providing to stimulate *D. oralis* growth.

As emphasized with the isolation of *D. oralis*, cultivation is ideally the ultimate goal to allow one to fully understand the repertoire of bacteria that make up the human microbiome and their roles in health and disease. However, a recent view of the tree of life has revealed an even more expansive assessment of bacterial lineages identifying entire phyla of currently uncultured organisms (27). Some of these lineages exist within the human oral microbiome (7). Specifically, one such phylum, the TM7 phylum (Candidate Division Saccharibacteria), has been recalcitrant to cultivation for many years with only one TM7 species, TM7x, isolated from the human oral cavity (45). Although the isolation of TM7x has significantly influenced our view of the oral microbiome, it is only one isolate of a large taxonomic diversity of TM7 bacteria in the oral cavity (29) and provides limited insight into the physiological diversification and roles these bacteria play there. **Chapter 4** demonstrates a novel reverse-genomics approach using in-silico antibody design based on published sequencing data to target a wide diversity of TM7 bacteria in the human oral cavity for genomics and cultivation. This method allows for

cultivation of bacteria who must also be co-isolated with their respective obligate bacterial hosts; a bottleneck that has resulted from classical cultivation techniques.

We validate this with the cultivation of a previously uncultured lineage of TM7 bacteria HOT351 and its formerly uncultured host, *Actinomyces* HOT897. In addition to live cultivation, we also sequenced twenty-three human oral TM7 single-cell genomes covering a large taxonomic diversity (eight unique OTUs) of the TM7 phylum in the oral cavity, the largest collection so far at the individual OTU level. Comparative genomics of this large dataset has provided a robust analysis of a core human oral TM7 genome and comparisons with free-living TM7 genomes has identified what traits may be important for human host-association. As compared to free-living TM7 relatives, human oral TM7 bacteria are enriched for protein families of known virulence factors and novel viral defense mechanisms.

Next-generation sequencing has greatly informed our understanding of the vast array of microorganisms on Earth however the limited number of cultured representatives of entire phyla still hinders our ability to truly understand the roles these organisms really play in their environments. Inferences can be made regarding their functions based on genomics and phylogenetic placement but without being able to test those hypotheses in the lab, sequencing remains merely hypothetical. The dual approach of my dissertation in leveraging next-generation sequencing analysis to bring disease-associated bacterial lineages into culture will uncover new information about not only individual microorganisms but how these organisms contribute to microbial community architecture and the specific interactions that may lead to dysbiosis and disease.

Moving forward, figuring out the “mystery factor” that *F. nucleatum* is producing to stimulate growth of *D. oralis* is priority. To do so, one can utilize preparative HPLC to further fractionate the EtAc2 sample and continue to narrow in on what is in this supporting fraction. As it may be more than one compound working in synergy with hydroxybutyric acid, then this less complex fraction may be better resolved by LC-MS and NMR as GC-MS has limitations in the types of compounds that it can identify. Further, heavy labeled substrates (^{13}C) can be purchased or synthesized of the identifying compounds and fed to *D. oralis* to determine whether or not these are actually being utilized by *D. oralis*. If some growth stimulation is occurring, then it may be a concentration

problem (either not enough or inhibitory concentrations of the target compound) that needs to be further investigated. Lastly, it would be interesting to test whether the final identified compound(s) can support growth of other oral bacteria that have been shown to be dependent on *F. nucleatum* for growth.

Continued advancement and optimization of the “reverse genomics” is dually important. Culturing additional representatives of TM7 will uncover more information regarding host specificity and interactions. Likewise, expanding to environmental TM7 and other (currently uncultured) lineages of bacteria, such as SR1 and GN02, can identify whether all CPR organisms have actinobacterial hosts in host-associated or free-living environments or have evolved separately with different hosts. Understanding the range of hosts can also aid in understanding limitations of host association (such as only Gram positive organisms capable of being hosts or in identifying shared features not present in non-host vs. host bacteria).

List of References

1. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. Nature. 2007;449(7164):804-10. doi: 10.1038/nature06244. PubMed PMID: 17943116; PMCID: PMC3709439.
2. Human Microbiome Project C. Structure, function and diversity of the healthy human microbiome. Nature. 2012;486(7402):207-14. doi: 10.1038/nature11234. PubMed PMID: 22699609; PMCID: 3564958.
3. D'Argenio V. Human Microbiome Acquisition and Bioinformatic Challenges in Metagenomic Studies. Int J Mol Sci. 2018;19(2). Epub 2018/02/01. doi: 10.3390/ijms19020383. PubMed PMID: 29382070; PMCID: PMC5855605.
4. Collado MC, Rautava S, Aakko J, Isolauri E, Salminen S. Human gut colonisation may be initiated in utero by distinct microbial communities in the placenta and amniotic fluid. Sci Rep. 2016;6:23129. Epub 2016/03/24. doi: 10.1038/srep23129. PubMed PMID: 27001291; PMCID: PMC4802384.
5. Stinson LF, Boyce MC, Payne MS, Keelan JA. The Not-so-Sterile Womb: Evidence That the Human Fetus Is Exposed to Bacteria Prior to Birth. Frontiers in Microbiology. 2019;10. doi: 10.3389/fmicb.2019.01124.
6. Cho I, Blaser MJ. The human microbiome: at the interface of health and disease. Nat Rev Genet. 2012;13(4):260-70. doi: 10.1038/nrg3182. PubMed PMID: 22411464; PMCID: PMC3418802.
7. Griffen AL, Beall CJ, Campbell JH, Firestone ND, Kumar PS, Yang ZK, Podar M, Leys EJ. Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. ISME J. 2012;6(6):1176-85. doi: 10.1038/ismej.2011.191. PubMed PMID: 22170420; PMCID: PMC3358035.
8. Fodor AA, DeSantis TZ, Wylie KM, Badger JH, Ye Y, Hepburn T, Hu P, Sodergren E, Liolios K, Huot-Creasy H, Birren BW, Earl AM. The “Most Wanted” Taxa from the Human Microbiome for whole genome sequencing. PLoS One. 2012;7(7). doi: 10.1371/journal.pone.0041294.
9. Lloyd KG, Steen AD, Ladau J, Yin J, Crosby L. Phylogenetically Novel Uncultured Microbial Cells Dominate Earth Microbiomes. mSystems. 2018;3(5). Epub 2018/10/03. doi: 10.1128/mSystems.00055-18. PubMed PMID: 30273414; PMCID: PMC6156271.

10. Lloyd-Price J, Mahurkar A, Rahnavard G, Crabtree J, Orvis J, Hall AB, Brady A, Creasy HH, McCracken C, Giglio MG, McDonald D, Franzosa EA, Knight R, White O, Huttenhower C. Strains, functions and dynamics in the expanded Human Microbiome Project. Nature. 2017;550(7674):61-6. Epub 2017/09/28. doi: 10.1038/nature23889. PubMed PMID: 28953883; PMCID: PMC5831082.
11. Pasolli E, Asnicar F, Manara S, Zolfo M, Karcher N, Armanini F, Beghini F, Manghi P, Tett A, Ghensi P, Collado MC, Rice BL, DuLong C, Morgan XC, Golden CD, Quince C, Huttenhower C, Segata N. Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle. Cell. 2019;176(3):649-62 e20. Epub 2019/01/22. doi: 10.1016/j.cell.2019.01.001. PubMed PMID: 30661755; PMCID: PMC6349461.
12. Chen T, Yu WH, Izard J, Baranova OV, Lakshmanan A, Dewhirst FE. The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information. Database (Oxford). 2010;2010:baq013. doi: 10.1093/database/baq013. PubMed PMID: 20624719; PMCID: PMC2911848.
13. Dewhirst FE, Chen T, Izard J, Paster BJ, Tanner AC, Yu WH, Lakshmanan A, Wade WG. The human oral microbiome. J Bacteriol. 2010;192(19):5002-17. doi: 10.1128/JB.00542-10. PubMed PMID: 20656903; PMCID: PMC2944498.
14. Escapa IF, Chen T, Huang Y, Gajare P, Dewhirst FE, Lemon KP. New Insights into Human Nostril Microbiome from the Expanded Human Oral Microbiome Database (eHOMD): a Resource for the Microbiome of the Human Aerodigestive Tract. mSystems. 2018;3(6). Epub 2018/12/12. doi: 10.1128/mSystems.00187-18. PubMed PMID: 30534599; PMCID: PMC6280432.
15. Jenkinson HF, Lamont RJ. Oral microbial communities in sickness and in health. Trends Microbiol. 2005;13(12):589-95. doi: 10.1016/j.tim.2005.09.006. PubMed PMID: 16214341.
16. Darveau RP. Periodontitis: a polymicrobial disruption of host homeostasis. Nat Rev Microbiol. 2010;8(7):481-90. doi: 10.1038/nrmicro2337. PubMed PMID: 20514045.
17. Wade WG, Thompson BP, Rybalka A, Vartoukian SR. Uncultured members of the oral microbiome. California Dental Association Journal. 2016;44(7):447-56.

18. Sudhakara P, Gupta A, Bhardwaj A, Wilson A. Oral Dysbiotic Communities and Their Implications in Systemic Diseases. Dentistry Journal. 2018;6(2). doi: 10.3390/dj6020010.
19. Research NIoDaC. Peridontal (Gum) Disease: Causes, Symptoms, and Treatments: NIH Publication No. 13-1142; 2013 [cited 2015 December 14]. Available
20. Abusleme L, Dupuy AK, Dutzan N, Silva N, Burleson JA, Strausbaugh LD, Gamonal J, Diaz PI. The subgingival microbiome in health and periodontitis and its relationship with community biomass and inflammation. ISME J. 2013;7(5):1016-25. doi: 10.1038/ismej.2012.174. PubMed PMID: 23303375; PMCID: 3635234.
21. Hajishengallis G, Lamont RJ. Beyond the red complex and into more complexity: the polymicrobial synergy and dysbiosis (PSD) model of periodontal disease etiology. Mol Oral Microbiol. 2012;27(6):409-19. doi: 10.1111/j.2041-1014.2012.00663.x. PubMed PMID: 23134607; PMCID: PMC3653317.
22. Mason MR, Chambers S, Dabdoub SM, Thikkurissy S, Kumar PS. Characterizing oral microbial communities across dentition states and colonization niches. Microbiome. 2018;6(1):67. doi: 10.1186/s40168-018-0443-2. PubMed PMID: 29631628; PMCID: PMC5891995.
23. Szafranski SP, Wos-Oxley ML, Vilchez-Vargas R, Jauregui R, Plumeier I, Klawonn F, Tomasch J, Meisinger C, Kuhnisch J, Sztajer H, Pieper DH, Wagner-Dobler I. High-resolution taxonomic profiling of the subgingival microbiome for biomarker discovery and periodontitis diagnosis. Appl Environ Microbiol. 2015;81(3):1047-58. doi: 10.1128/AEM.03534-14. PubMed PMID: 25452281; PMCID: PMC4292489.
24. Espinoza J, Harkins D, Torralba M, Gomez A, Highlander S, Jones M, Leong P, Saffery R, Bockmann M, Kuelbs C, Inman J, Hughes T, Craig JM, Nelson KE, Dupont C. Supragingival plaque microbiome ecology and functional potential in the contect of health and disease2018. doi: 10.1101/325407.
25. Alneberg J, Karlsson CMG, Divne A, Begin C, Homa F, Lindh MV, Hugerth L, Ettema TJG, Bertilsson S, Andersson AF, Pinhassi J. Genomes of uncultivated prokaryotes: a comparison of metagenome-assembled and single-amplified genomes. Microbiome. 2018;6(173). doi: 10.1186/s40168-018-0550-0.

26. Woyke T, Doud DFR, Schulz F. The trajectory of microbial single-cell sequencing. Nat Methods. 2017;14(11):1045-54. Epub 2017/11/01. doi: 10.1038/nmeth.4469. PubMed PMID: 29088131.
27. Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, Butterfield CN, Hernsdorf AW, Amano Y, Ise K, Suzuki Y, Dudek N, Relman DA, Finstad KM, Amundson R, Thomas BC, Banfield JF. A new view of the tree of life. Nat Microbiol. 2016;1:16048. doi: 10.1038/nmicrobiol.2016.48. PubMed PMID: 27572647.
28. Marcy Y, Ouverney C, Bik EM, Losekann T, Ivanova N, Martin HG, Szeto E, Platt D, Hugenholtz P, Relman DA, Quake SR. Dissecting biological "dark matter" with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth. Proc Natl Acad Sci U S A. 2007;104(29):11889-94. Epub 2007/07/11. doi: 10.1073/pnas.0704662104. PubMed PMID: 17620602; PMCID: PMC1924555.
29. Brinig MM, Lepp PW, Ouverney CC, Armitage GC, Relman DA. Prevalence of Bacteria of Division TM7 in Human Subgingival Plaque and Their Association with Disease. Applied and Environmental Microbiology. 2003;69(3):1687-94. doi: 10.1128/aem.69.3.1687-1694.2003.
30. Christie PJ, Whitaker N, Gonzalez-Rivera C. Mechanism and structure of the bacterial type IV secretion systems. Biochim Biophys Acta. 2014;1843(8):1578-91. doi: 10.1016/j.bbamcr.2013.12.019. PubMed PMID: 24389247; PMCID: PMC4061277.
31. Green ER, Mecsas J. Bacterial Secretion Systems: An Overview. Microbiol Spectr. 2016;4(1). doi: 10.1128/microbiolspec.VMBF-0012-2015. PubMed PMID: 26999395; PMCID: PMC4804464.
32. Campbell JH, O'Donoghue P, Campbell AG, Schwientek P, Sczyrba A, Woyke T, Soll D, Podar M. UGA is an additional glycine codon in uncultured SR1 bacteria from the human microbiota. PNAS. 2013;110(14):5540-5.
33. Campbell AG, Campbell JH, Schwientek P, Woyke T, Sczyrba A, Allman S, Beall CJ, Griffen A, Leys E, Podar M. Multiple single-cell genomes provide insight into functions of uncultured Deltaproteobacteria in the human oral cavity. PLoS One. 2013;8(3):e59361. doi: 10.1371/journal.pone.0059361. PubMed PMID: 23555659; PMCID: PMC3608642.

34. Campbell AG, Schwientek P, Vishnivetskaya T, Woyke T, Levy S, Beall CJ, Griffen A, Leys E, Podar M. Diversity and genomic insights into the uncultured Chloroflexi from the human microbiota. Environ Microbiol. 2014;16(9):2635-43. doi: 10.1111/1462-2920.12461. PubMed PMID: 24738594; PMCID: PMC4149597.
35. Beall CJ, Campbell AG, Dayeh DM, Griffen AL, Podar M, Leys EJ. Single cell genomics of uncultured, health-associated Tannerella BU063 (Oral Taxon 286) and comparison to the closely related pathogen Tannerella forsythia. PLoS One. 2014;9(2):e89398. doi: 10.1371/journal.pone.0089398. PubMed PMID: 24551246; PMCID: PMC3925233.
36. Belda-Ferre P, Alcaraz LD, Cabrera-Rubio R, Romero H, Simon-Soro A, Pignatelli M, Mira A. The oral metagenome in health and disease. ISME J. 2012;6(1):46-56. Epub 2011/07/01. doi: 10.1038/ismej.2011.85. PubMed PMID: 21716308; PMCID: PMC3246241.
37. Jorth P, Turner KH, Gumus P, Nizam N, Buduneli N, Whiteley M. Metatranscriptomics of the human oral microbiome during health and disease. mBio. 2014;5(2):e01012-14. doi: 10.1128/mBio.01012-14. PubMed PMID: 24692635; PMCID: PMC3977359.
38. Nowicki EM, Shroff R, Singleton JA, Renaud DE, Wallace D, Drury J, Zirnheld J, Colleti B, Ellington AD, Lamont RJ, Scott DA, Whiteley M. Microbiota and Metatranscriptome Changes Accompanying the Onset of Gingivitis. MBio. 2018;9(2). Epub 2018/04/19. doi: 10.1128/mBio.00575-18. PubMed PMID: 29666288; PMCID: PMC5904416.
39. Vartoukian SR, Downes J, Palmer RM, Wade WG. Fretibacterium fastidiosum gen. nov., sp. nov., isolated from the human oral cavity. Int J Syst Evol Microbiol. 2013;63(Pt 2):458-63. doi: 10.1099/ijs.0.041038-0. PubMed PMID: 22493171.
40. Vartoukian SR, Moazzez RV, Paster BJ, Dewhirst FE, Wade WG. First Cultivation of Health-Associated Tannerella sp. HOT-286 (BU063). J Dent Res. 2016. doi: 10.1177/0022034516651078. PubMed PMID: 27193146.
41. Vartoukian SR, Adamowska A, Lawlor M, Moazzez R, Dewhirst FE, Wade WG. In Vitro Cultivation of 'Unculturable' Oral Bacteria, Facilitated by Community Culture and

Media Supplementation with Siderophores. PLoS One. 2016;11(1):e0146926. doi: 10.1371/journal.pone.0146926. PubMed PMID: 26764907; PMCID: 4713201.
42. Thompson H, Rybalka A, Moazzez R, Dewhirst FE, Wade WG. In-vitro culture of previously uncultured oral bacterial phylotypes. Appl Environ Microbiol. 2015. doi: 10.1128/AEM.02156-15. PubMed PMID: 26407883.
43. Tian Y, He X, Torralba M, Yooseph S, Nelson KE, Lux R, McLean JS, Yu G, Shi W. Using DGGE profiling to develop a novel culture medium suitable for oral microbial communities. Molecular Oral Microbiology. 2010;25.
44. Albertsen M, Hugenholtz P, Skarshewski A, Nielsen KL, Tyson GW, Nielsen PH. Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. Nat Biotechnol. 2013;31(6):533-8. doi: 10.1038/nbt.2579. PubMed PMID: 23707974.
45. He X, McLean JS, Edlund A, Yooseph S, Hall AP, Liu SY, Dorrestein PC, Esquenazi E, Hunter RC, Cheng G, Nelson KE, Lux R, Shi W. Cultivation of a human-associated TM7 phylotype reveals a reduced genome and epibiotic parasitic lifestyle. Proc Natl Acad Sci U S A. 2015;112(1):244-9. doi: 10.1073/pnas.1419038112. PubMed PMID: 25535390; PMCID: 4291631.
46. Bor B, Poweleit N, Bois JS, Cen L, Bedree JK, Zhou ZH, Gunsalus RP, Lux R, McLean JS, He X, Shi W. Phenotypic and Physiological Characterization of the Epibiotic Interaction Between TM7x and Its Basibiont Actinomyces. Microb Ecol. 2016;71(1):243-55. doi: 10.1007/s00248-015-0711-7. PubMed PMID: 26597961; PMCID: PMC4688200.
47. Kantor RS, Wrighton KC, Handley KM, Sharon I, Hug LA, Castelle CJ, Thomas BC, Banfield JF. Small genomes and sparse metabolisms of sediment-associated bacteria from four candidate phyla. MBio. 2013;4(5):e00708-13. doi: 10.1128/mBio.00708-13. PubMed PMID: 24149512; PMCID: PMC3812714.
48. Bedree JK, Bor B, Cen L, Edlund A, Lux R, McLean JS, Shi W, He X. Quorum Sensing Modulates the Epibiotic-Parasitic Relationship Between Actinomyces odontolyticus and Its Saccharibacteria epibiont, a Nanosynbacter lyticus Strain, TM7x. Frontiers in Microbiology. 2018;9. doi: 10.3389/fmicb.2018.02049.

49. Mark Welch JL, Rossetti BJ, Rieken CW, Dewhirst FE, Borisy GG. Biogeography of a human oral microbiome at the micron scale. Proc Natl Acad Sci U S A. 2016. doi: 10.1073/pnas.1522149113. PubMed PMID: 26811460.
50. Marsh PD, Zaura E. Dental biofilm: ecological interactions in health and disease. J Clin Periodontol. 2017;44 Suppl 18:S12-S22. Epub 2017/03/08. doi: 10.1111/jcpe.12679. PubMed PMID: 28266111.
51. Edlund A, Yang Y, Hall AP, Guo L, Lux R, He X, Nelson KE, Nealson KH, Yooseph S, Shi W, McLean JS. An in vitro biofilm model system maintaining a highly reproducible species and metabolic diversity approaching that of the human oral microbiome. Microbiome. 2013;1(1):25. doi: 10.1186/2049-2618-1-25. PubMed PMID: 24451062; PMCID: PMC3971625.
52. Lam RH, Cui X, Guo W, Thorsen T. High-throughput dental biofilm growth analysis for multiparametric microenvironmental biochemical conditions using microfluidics. Lab Chip. 2016;16(9):1652-62. Epub 2016/04/06. doi: 10.1039/c6lc00072j. PubMed PMID: 27045372.
53. Nance WC, Dowd SE, Samarian D, Chludzinski J, Delli J, Battista J, Rickard AH. A high-throughput microfluidic dental plaque biofilm system to visualize and quantify the effect of antimicrobials. J Antimicrob Chemother. 2013;68(11):2550-60. Epub 2013/06/27. doi: 10.1093/jac/dkt211. PubMed PMID: 23800904; PMCID: PMC3797639.
54. Fitzsimons MS, Novotny M, Lo CC, Dichosa AE, Yee-Greenbaum JL, Snook JP, Gu W, Chertkov O, Davenport KW, McMurry K, Reitenga KG, Daughton AR, He J, Johnson SL, Gleasner CD, Wills PL, Parson-Quintana B, Chain PS, Detter JC, Lasken RS, Han CS. Nearly finished genomes produced using gel microdroplet culturing reveal substantial intraspecies genomic diversity within the human microbiome. Genome Res. 2013;23(5):878-88. Epub 2013/03/16. doi: 10.1101/gr.142208.112. PubMed PMID: 23493677; PMCID: PMC3638143.
55. Langendijk PS, Kulik EM, Sandmeier H, Meyer J, van der Hoeven JS. Isolation of Desulfomicrobium orale sp. nov. and Desulfovibrio strain NY682, oral sulfate-reducing bacteria involved in human periodontal disease. International Journal of Systematic and Evolutionary Microbiology. 2001;51:1035-44.

56. Rowan FE, Docherty NG, Coffey JC, O'Connell PR. Sulphate-reducing bacteria and hydrogen sulphide in the aetiology of ulcerative colitis. Br J Surg. 2009;96(2):151-8. doi: 10.1002/bjs.6454. PubMed PMID: 19160346.
57. Boopathy R, Robichaux M, LaFont D, Howell M. Activity of sulfate-reducing bacteria in human periodontal pocket. Canadian Journal of Microbiology. 2002;48(12):1099-103. doi: 10.1139/w02-104.
58. Kumar PS, Griffen AL, Moeschberger ML, Leys EJ. Identification of candidate periodontal pathogens and beneficial species by quantitative 16S clonal analysis. J Clin Microbiol. 2005;43(8):3944-55. doi: 10.1128/JCM.43.8.3944-3955.2005. PubMed PMID: 16081935; PMCID: PMC1233920.
59. Wade WG, Munson MA, de Lillo A, Weightman AJ. Specificity of the oral microflora in dentinal caries, endodontic infections and periodontitis. International Congress Series. 2005;1284:150-7. doi: 10.1016/j.ics.2005.06.097.
60. El Houari A, Ranchou-Peyruse M, Ranchou-Peyruse A, Dakdaki A, Guignard M, Idouhammou L, Bennisse R, Bouterfass R, Guyoneaud R, Qatibi AI. Desulfobulbus oligotrophicus sp. nov., a sulfate-reducing and propionate-oxidizing bacterium isolated from a municipal anaerobic sewage sludge digester. Int J Syst Evol Microbiol. 2017;67(2):275-81. doi: 10.1099/ijsem.0.001615. PubMed PMID: 27902225.
61. Kharrat H, Karray F, Bartoli M, Ben Hnia W, Mhiri N, Fardeau ML, Bennour F, Kamoun L, Alazard D, Sayadi S. Desulfobulbus aggregans sp. nov., a Novel Sulfate Reducing Bacterium Isolated from Marine Sediment from the Gulf of Gabes. Curr Microbiol. 2017;74(4):449-54. doi: 10.1007/s00284-017-1211-4. PubMed PMID: 28213662.
62. Lien T, Madsen M, Steen IH, Gjerdevik K. Desulfobulbus rhabdoformis sp. nov., a sulfate reducer from a water-oil seperation system. International Journal of Systematic Bacteriology. 1998(48):469-74.
63. Pagani I, Lapidus A, Nolan M, Lucas S, Hammon N, Deshpande S, Cheng JF, Chertkov O, Davenport K, Tapia R, Han C, Goodwin L, Pitluck S, Liolios K, Mavromatis K, Ivanova N, Mikhailova N, Pati A, Chen A, Palaniappan K, Land M, Hauser L, Chang YJ, Jeffries CD, Detter JC, Brambilla E, Kannan KP, Djao OD, Rohde M, Pukall R, Spring S, Goker M, Sikorski J, Woyke T, Bristow J, Eisen JA, Markowitz V, Hugenholtz

P, Kyrpides NC, Klenk HP. Complete genome sequence of Desulfobulbus propionicus type strain (1pr3). Stand Genomic Sci. 2011;4(1):100-10. doi: 10.4056/sigs.1613929. PubMed PMID: 21475592; PMCID: PMC3072085.

64. Samain E, Dubourguier HC, Albagnac G. Isolation and characterization of Desulfobulbus elongatus sp. nov. from a mesophilic industrial digester. Systematic and Applied Microbiology. 1984;5(3):391-401.

65. Sass A, Rutters H, Cypionka H, Sass H. Desulfobulbus mediterraneus sp. nov., a sulfate-reducing bacterium growing on mono- and disaccharides. Arch Microbiol. 2002;177(6):468-74. doi: 10.1007/s00203-002-0415-5. PubMed PMID: 12029392.

66. Sorokin DY, Tourova TP, Panteleeva AN, Muyzer G. Desulfonatronobacter acidivorans gen. nov., sp. nov. and Desulfobulbus alkaliphilus sp. nov., haloalkaliphilic heterotrophic sulfate-reducing bacteria from soda lakes. Int J Syst Evol Microbiol. 2012;62(Pt 9):2107-13. doi: 10.1099/ijs.0.029777-0. PubMed PMID: 22039002.

67. Suzuki D, Ueki A, Amaishi A, Ueki K. Desulfobulbus japonicus sp. nov., a novel Gram-negative propionate-oxidizing, sulfate-reducing bacterium isolated from an estuarine sediment in Japan. Int J Syst Evol Microbiol. 2007;57(Pt 4):849-55. doi: 10.1099/ijs.0.64855-0. PubMed PMID: 17392218.

68. Podar M, Gilmour CC, Brandt CC, Soren A, Brown SD, Crable BR, Palumbo AV, Somenahally AC, Elias DA. Global prevalence and distribution of genes and microorganisms involved in mercury methylation. Science Advantage. 2015:1-12. doi: e1500675.

69. Winsley TJ, Snape I, McKinlay J, Stark J, van Dorst JM, Ji M, Ferrari BC, Siciliano SD. The ecological controls on the prevalence of candidate division TM7 in polar regions. Front Microbiol. 2014;5:345. doi: 10.3389/fmicb.2014.00345. PubMed PMID: 25076941; PMCID: PMC4097103.

70. Podar M, Abulencia CB, Walcher M, Hutchison D, Zengler K, Garcia JA, Holland T, Cotton D, Hauser L, Keller M. Targeted access to the genomes of low-abundance organisms in complex microbial communities. Appl Environ Microbiol. 2007;73(10):3205-14. doi: 10.1128/AEM.02985-06. PubMed PMID: 17369337; PMCID: PMC1907129.

71. Dinis JM, Barton DE, Ghadiri J, Surendar D, Reddy K, Velasquez F, Chaffee CL, Lee MC, Gavrilova H, Ozuna H, Smits SA, Ouverney CC. In search of an uncultured human-associated TM7 bacterium in the environment. PLoS One. 2011;6(6):e21280. doi: 10.1371/journal.pone.0021280. PubMed PMID: 21701585; PMCID: PMC3118805.
72. Hanada A, Kurogi T, Giang NM, Yamada T, Kamimoto Y, Kiso Y, Hiraishi A. Bacteria of the Candidate Phylum TM7 are Prevalent in Acidophilic Nitrifying Sequencing-Batch Reactors. Microbes and Environments. 2014;29(4):353-62. doi: 10.1264/jsme2.ME14052.
73. Hugenholtz P, Tyson GW, Webb RI, Wagner AM, Blackall LL. Investigation of candidate division TM7, a recently recognized major lineage of the domain Bacteria with no known pure-culture representatives. Appl Environ Microbiol. 2001;67(1):411-9. doi: 10.1128/AEM.67.1.411-419.2001. PubMed PMID: 11133473; PMCID: PMC92593.
74. Kuehbacher T, Rehman A, Lepage P, Hellmig S, Folsch UR, Schreiber S, Ott SJ. Intestinal TM7 bacterial phylogenies in active inflammatory bowel disease. J Med Microbiol. 2008;57(Pt 12):1569-76. doi: 10.1099/jmm.0.47719-0. PubMed PMID: 19018031.
75. Fredricks DN, Fiedler TL, Marrazzo JM. Molecular Identification of Bacteria Associated with Bacterial Vaginosis. New England Journal of Medicine. 2005;353(18).
76. Rheims H, Rainey FA, Stackebrandt E. A molecular approach to search for diversity among bacteria in the environment. Journal of Industrial Microbiology. 1996;17:159-69.
77. Rheims H, Sproer C, Rainey FA, Stackebrandt E. Molecular biological evidence for the occurence of uncultured members of the actinomycete line of deceint in different enviornments and genographical locations. Microbiology. 1996;142:2863-70.
78. Liu B, Faller LL, Klitgord N, Mazumdar V, Ghodsi M, Sommer DD, Gibbons TR, Treangen TJ, Chang YC, Li S, Stine OC, Hasturk H, Kasif S, Segre D, Pop M, Amar S. Deep sequencing of the oral microbiome reveals signatures of periodontal disease. PLoS One. 2012;7(6):e37919. doi: 10.1371/journal.pone.0037919. PubMed PMID: 22675498; PMCID: PMC3366996.

79. Bentley SD, Parkhill J. Genomic perspectives on the evolution and spread of bacterial pathogens. Proc Biol Sci. 2015;282(1821). doi: 10.1098/rspb.2015.0488. PubMed PMID: 26702036.
80. Moran NA. Microbial minimalism: genome reduction in bacterial pathogens. Cell. 2002;108:583-86.
81. McCutcheon JP, Moran NA. Extreme genome reduction in symbiotic bacteria. Nat Rev Microbiol. 2011;10(1):13-26. doi: 10.1038/nrmicro2670. PubMed PMID: 22064560.
82. Nash AK, Auchtung TA, Wong MC, Smith DP, Gesell JR, Ross MC, Stewart CJ, Metcalf GA, Muzny DM, Gibbs RA, Ajami NJ, Petrosino JF. The gut mycobiome of the Human Microbiome Project healthy cohort. Microbiome. 2017;5(1):153. Epub 2017/11/28. doi: 10.1186/s40168-017-0373-4. PubMed PMID: 29178920; PMCID: PMC5702186.
83. Findley K, Oh J, Yang J, Conlan S, Deming C, Meyer JA, Schoenfeld D, Nomicos E, Park M, Program NIHISCCS, Kong HH, Segre JA. Topographic diversity of fungal and bacterial communities in human skin. Nature. 2013;498(7454):367-70. Epub 2013/05/24. doi: 10.1038/nature12171. PubMed PMID: 23698366; PMCID: PMC3711185.
84. Bradford LL, Ravel J. The vaginal mycobiome: A contemporary perspective on fungi in women's health and diseases. Virulence. 2017;8(3):342-51. Epub 2016/09/23. doi: 10.1080/21505594.2016.1237332. PubMed PMID: 27657355; PMCID: PMC5411243.
85. Ghannoum MA, Jurevic RJ, Mukherjee PK, Cui F, Sikaroodi M, Naqvi A, Gillevet PM. Characterization of the oral fungal microbiome (mycobiome) in healthy individuals. PLoS Pathog. 2010;6(1):e1000713. Epub 2010/01/15. doi: 10.1371/journal.ppat.1000713. PubMed PMID: 20072605; PMCID: PMC2795202.
86. Ahrendt SR, Quandt CA, Ciobanu D, Clum A, Salamov A, Andreopoulos B, Cheng JF, Woyke T, Pelin A, Henrissat B, Reynolds NK, Benny GL, Smith ME, James TY, Grigoriev IV. Leveraging single-cell genomics to expand the fungal tree of life. Nat Microbiol. 2018;3(12):1417-28. Epub 2018/10/10. doi: 10.1038/s41564-018-0261-0. PubMed PMID: 30297742.

87. Olm MR, West PT, Brooks B, Firek BA, Baker R, Morowitz MJ, Banfield JF. Genome-resolved metagenomics of eukaryotic populations during early colonization of premature infants and in hospital rooms. Microbiome. 2019;7(1). doi: 10.1186/s40168-019-0638-1.
88. Li CL, Liu DL, Jiang YT, Zhou YB, Zhang MZ, Jiang W, Liu B, Liang JP. Prevalence and molecular diversity of Archaea in subgingival pockets of periodontitis patients. Oral Microbiology Immunology. 2009;24:343-6.
89. Koskinen K, Pausan MR, Perras AK, Beck M, Bang C, Mora M, Schilhabel A, Schmitz R, M'oissl-Eichinger C. First insights into the diverse human archaeome: specific detection of archaea in the gastrointestinal tract, lung, and nose and on skin. mBio. 2017;8(6).
90. Nkamga VD, Henrissat B, Drancourt M. Archaea: Essential inhabitants of the human digestive microbiota. Human Microbiome Journal. 2017;3:1-8. doi: 10.1016/j.humic.2016.11.005.
91. Horz HP, Conrads G. Methanogenic Archaea and oral infections - ways to unravel the black box. J Oral Microbiol. 2011;3. Epub 2011/05/05. doi: 10.3402/jom.v3i0.5940. PubMed PMID: 21541092; PMCID: PMC3086593.
92. Mokili JL, Rohwer F, Dutilh BE. Metagenomics and future perspectives in virus discovery. Curr Opin Virol. 2012;2(1):63-77. Epub 2012/03/24. doi: 10.1016/j.coviro.2011.12.004. PubMed PMID: 22440968.
93. Ly M, Jones MB, Abeles SR, Santiago-Rodriguez TM, Gao J, Chan IC, Ghose C, Pride DT. Transmission of viruses via our microbiomes. Microbiome. 2016;4(1):64. Epub 2016/12/04. doi: 10.1186/s40168-016-0212-z. PubMed PMID: 27912785; PMCID: PMC5134127.
94. Oh J, Byrd AL, Park M, Program NCS, Kong HH, Segre JA. Temporal Stability of the Human Skin Microbiome. Cell. 2016;165(4):854-66. Epub 2016/05/08. doi: 10.1016/j.cell.2016.04.008. PubMed PMID: 27153496; PMCID: PMC4860256.
95. Abeles SR, Robles-Sikisaka R, Ly M, Lum AG, Salzman J, Boehm TK, Pride DT. Human oral viruses are personal, persistent and gender-consistent. ISME J. 2014;8(9):1753-67. Epub 2014/03/22. doi: 10.1038/ismej.2014.31. PubMed PMID: 24646696; PMCID: PMC4139723.

96. Wylie KM, Mihindukulasuriya KA, Zhou Y, Sodergren E, Storch GA, Weinstock GM. Metagenomic analysis of double-stranded DNA viruses in healthy adults. BMC Biol. 2014;12:71. Epub 2014/09/13. doi: 10.1186/s12915-014-0071-7. PubMed PMID: 25212266; PMCID: PMC4177058.
97. Dye BA. Global periodontal disease epidemiology. Periodontology 2000. 2012;58:10-25.
98. Kassebaum NJ, Bernabe E, Dahiya M, Bhandari B, Murray CJ, Marcenes W. Global burden of severe periodontitis in 1990-2010: a systematic review and meta-regression. J Dent Res. 2014;93(11):1045-53. doi: 10.1177/0022034514552491. PubMed PMID: 25261053; PMCID: PMC4293771.
99. Hajishengallis G, Liang S, Payne MA, Hashim A, Jotwani R, Eskan MA, McIntosh ML, Alsam A, Kirkwood KL, Lambris JD, Darveau RP, Curtis MA. Low-abundance biofilm species orchestrates inflammatory periodontal disease through the commensal microbiota and complement. Cell Host Microbe. 2011;10(5):497-506. doi: 10.1016/j.chom.2011.10.006. PubMed PMID: 22036469; PMCID: PMC3221781.
100. Diaz PI, Hoare A, Hong B-H. Subgingival Microbiome Shifts and Community Dynamics in Periodontal Diseases. California Dental Association Journal. 2016;44(7):421-36.
101. Khan SA, Kong EF, Meiller TF, Jabra-Rizk MA. Periodontal Diseases: Bug Induced, Host Promoted. PLoS Pathog. 2015;11(7):e1004952. doi: 10.1371/journal.ppat.1004952. PubMed PMID: 26226364; PMCID: PMC4520614.
102. Ohlrich EJ, Cullinan MP, Seymour GJ. The immunopathogenesis of periodontal disease. Australian dental journal. 2009;54 Suppl 1:2-10. doi: 10.1111/j.1834-7819.2009.01139.x. PubMed PMID: 19737265.
103. Socransky SS, AHaffajee AD, Cugini MA, Smith C, Kent Jr RL. Microbial complexes in subgingival plaque. J Clin Periodontol. 1998;25:134-44.
104. Griffen AL, Beall CJ, Campbell JH, Firestone ND, Kumar PS, Yang ZK, Podar M, Leys EJ. Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. ISME J. 2011;6:1176-85. Epub 2011/12/16. doi: ismej2011191 [pii]
10.1038/ismej.2011.191. PubMed PMID: 22170420.

105. Oliveira RR, Fermiano D, Feres M, Figueiredo LC, Teles FR, Soares GM, Faveri M. Levels of Candidate Periodontal Pathogens in Subgingival Biofilm. J Dent Res. 2016;95(6):711-8. doi: 10.1177/0022034516634619. PubMed PMID: 26936213; PMCID: PMC4924544.
106. Biyikoglu B, Ricker A, Diaz PI. Strain-specific colonization patterns and serum modulation of multi-species oral biofilm development. Anaerobe. 2012;18(4):459-70. doi: 10.1016/j.anaerobe.2012.06.003. PubMed PMID: 22771792; PMCID: PMC3410037.
107. Millhouse E, Jose A, Sherry L, Lappin DF, Patel N, Middleton AM, Pratten J, Culshaw S, Ramage G. Development of an in vitro periodontal biofilm model for assessing antimicrobial host modulatory effects of bioactive molecules. BMC Oral Health. 2014;14(80).
108. Wade WG, Thompson BP, Rybalka A, Vartoukian SR. Uncultured Members of the Oral Microbiome. California Dental Association Journal. 2016;44(7):447-55.
109. Langendijk PS, Kulik EM, Sandmeier H, Meyer J, van der Hoeven JS. Isolation of Desulfomicrobium orale sp nov and Desulfovibrio strain NY682, oral sulfate-reducing bacteria involved in human periodontal disease. International Journal of Systematic and Evolutionary Microbiology. 2001;51:1035-44. PubMed PMID: WOS:000168900000039.
110. Loubinoux J, Blsson-Bouteillez C, Miller N, Le Faou AE. Isolation of the provisionally named Desulfovibrio fairfieldensis from human peridontal pockets. Oral Microbiology Immunology. 2002;17:321-3.
111. Ichiishi S, Tanaka K, Nakao K, Izumi K, Mikamo H, Watanabe K. First isolation of Desulfovibrio from the human vaginal flora. Anaerobe. 2010;16(3):229-33. doi: 10.1016/j.anaerobe.2010.02.002. PubMed PMID: 20159048.
112. Loubinoux J, Valente FM, Pereira IA, Costa A, Grimont PA, Le Faou AE. Reclassification of the only species of the genus Desulfomonas, Desulfomonas pigra, as Desulfovibrio piger comb. nov. Int J Syst Evol Microbiol. 2002;52(Pt 4):1305-8. Epub 2002/08/01. PubMed PMID: 12148644.
113. Rowan F, Docherty NG, Murphy M, Murphy B, Calvin Coffey J, O'Connell PR. Desulfovibrio bacterial species are increased in ulcerative colitis. Diseases of the colon

and rectum. 2010;53(11):1530-6. doi: 10.1007/DCR.0b013e3181f1e620. PubMed PMID: 20940602.
114. Widdel F, Pfennig N. Studies on dissimilatory sulfate-reducing bacteria that decompose fatty-acids II. Incomplete oxidation of propionate by *Desulfobulbus propionicus* gen. nov., sp. nov. Archives of microbiology. 1982;131(4):360-5. PubMed PMID: WOS:A1982NW24500014.
115. Bentley SD, Parkhill J. Genomic perspectives on the evolution and spread of bacterial pathogens. Proc Biol Sci. 2015;282(1821):1-8. doi: 10.1098/rspb.2015.0488. PubMed PMID: 26702036; PMCID: 4707741.
116. Bliven KA, Maurelli AT. Evolution of Bacterial Pathogens Within the Human Host. Microbiol Spectr. 2016;4(1):1-8. doi: 10.1128/microbiolspec.VMBF-0017-2015. PubMed PMID: 26999399; PMCID: 4804625.
117. Thompson H, Rybalka A, Moazzez R, Dewhirst FE, Wade WG. In-vitro culture of previously uncultured oral bacterial phylotypes. Appl Environ Microbiol. 2015;81:8307-14. doi: 10.1128/AEM.02156-15. PubMed PMID: 26407883.
118. Paster BJ, Boches SK, Galvin JL, Ericson RE, Lau CN, Levanos VA, Sahasrabudhe A, Dewhirst FE. Bacterial diversity in human subgingival plaque. J Bacteriol. 2001;183(12):3770-83. doi: 10.1128/JB.183.12.3770-3783.2001. PubMed PMID: 11371542; PMCID: PMC95255.
119. Schwechheimer C, Kuehn MJ. Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions. Nat Rev Microbiol. 2015;13(10):605-19. doi: 10.1038/nrmicro3525. PubMed PMID: 26373371; PMCID: 5308417.
120. Xie H. Biogenesis and function of Porphyromonas gingivalis outer membrane vesicles. Future microbiology. 2015;10(9):1517-27. doi: 10.2217/fmb.15.63. PubMed PMID: 26343879; PMCID: 4603655.
121. Vartoukian SR, Palmer RM, Wade WG. Strategies for culture of 'unculturable' bacteria. FEMS Microbiol Lett. 2010;309(1):1-7. doi: 10.1111/j.1574-6968.2010.02000.x. PubMed PMID: 20487025.
122. Krebs HA. Chemical composition of blood and plasma serum1950;19:409-30.
123. Armitage GC. Analysis of gingival crevice fluid and risk of progression of periodontitis. Periodontol 2000. 2004;34:109-19.

124. Brown SA, Whiteley M. A novel exclusion mechanism for carbon resource partitioning in Aggregatibacter actinomycetemcomitans. J Bacteriol. 2007;189(17):6407-14. doi: 10.1128/JB.00554-07. PubMed PMID: 17586632; PMCID: PMC1951915.
125. Langendijk PS, Hanssen JTJ, Van der Hoeven JS. Sulfate-reducing bacteria in association with human periodontitis. Journal of Clinical Periodontology. 2000;27(12):943-50. doi: 10.1034/j.1600-051x.2000.027012943.x. PubMed PMID: WOS:000165359600010.
126. Dewhirst FE, Klein EA, Thomas EC, Blanton JM, Chen T, Milella L, Buckley CMF, Davis IJ, Bennett M-L, Marshall-Jones ZV. The Canine Oral Microbiome. PloS One. 2012;7(4). doi: 10.1371/.
127. Dewhirst FE, Klein EA, Bennett ML, Croft JM, Harris SJ, Marshall-Jones ZV. The feline oral microbiome: a provisional 16S rRNA gene based taxonomy with full-length reference sequences. Vet Microbiol. 2015;175(2-4):294-303. doi: 10.1016/j.vetmic.2014.11.019. PubMed PMID: 25523504; PMCID: PMC4297541.
128. Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M, Jensen LJ, von Mering C, Bork P. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res. 2016;44(D1):D286-93. doi: 10.1093/nar/gkv1248. PubMed PMID: 26582926; PMCID: PMC4702882.
129. Rabus R, Venceslau SS, Wohlbrand L, Voordouw G, Wall JD, Pereira IA. A Post-Genomic View of the Ecophysiology, Catabolism and Biotechnological Relevance of Sulphate-Reducing Prokaryotes. Adv Microb Physiol. 2015;66:55-321. doi: 10.1016/bs.ampbs.2015.05.002. PubMed PMID: 26210106.
130. Duarte AG, Santos AA, Pereira IA. Electron transfer between the QmoABC membrane complex and adenosine 5'-phosphosulfate reductase. Biochim Biophys Acta. 2016;1857(4):380-6. doi: 10.1016/j.bbabio.2016.01.001. PubMed PMID: 26768116.
131. Schut GJ, Boyd ES, Peters JW, Adams MW. The modular respiratory complexes involved in hydrogen and sulfur metabolism by heterotrophic hyperthermophilic archaea and their evolutionary implications. FEMS Microbiol Rev. 2013;37(2):182-203. doi: 10.1111/j.1574-6976.2012.00346.x. PubMed PMID: 22713092.

132. Buckel W, Thauer RK. Energy conservation via electron bifurcating ferredoxin reduction and proton/Na(+) translocating ferredoxin oxidation. Biochim Biophys Acta. 2013;1827(2):94-113. doi: 10.1016/j.bbabio.2012.07.002. PubMed PMID: 22800682.
133. Ramos AR, Grein F, Oliveira GP, Venceslau SS, Keller KL, Wall JD, Pereira IA. The FlxABCD-HdrABC proteins correspond to a novel NADH dehydrogenase/heterodisulfide reductase widespread in anaerobic bacteria and involved in ethanol metabolism in Desulfovibrio vulgaris Hildenborough. Environ Microbiol. 2015;17(7):2288-305. doi: 10.1111/1462-2920.12689. PubMed PMID: 25367508.
134. Poock SR, Leach ER, Moir JW, Cole JA, Richardson DJ. Respiratory detoxification of nitric oxide by the cytochrome c nitrite reductase of Escherichia coli. J Biol Chem. 2002;277(26):23664-9. doi: 10.1074/jbc.M200731200. PubMed PMID: 11960983.
135. Weghoff MC, Bertsch J, Muller V. A novel mode of lactate metabolism in strictly anaerobic bacteria. Environ Microbiol. 2015;17(3):670-7. doi: 10.1111/1462-2920.12493. PubMed PMID: 24762045.
136. Seeliger S, Janssen PH, Schink B. Energetics and kinetics of lactate fermentation to acetate and propionate via methylmalonyl-CoA or acrylyl-CoA. FEMS Microbiol Lett. 2002;211(1):65-70. PubMed PMID: 12052552.
137. Roberts AP, Kreth J. The impact of horizontal gene transfer on the adaptive ability of the human oral microbiome. Front Cell Infect Microbiol. 2014;4:1-9. doi: 10.3389/fcimb.2014.00124. PubMed PMID: 25250243; PMCID: 4157583.
138. Smillie CS, Smith MB, Friedman J, Cordero OX, David LA, Alm EJ. Ecology drives a global network of gene exchange connecting the human microbiome. Nature. 2011;480(7376):241-4. doi: 10.1038/nature10571. PubMed PMID: 22037308.
139. Podar M, Gilmour CC, Brandt CC, Soren A, Brown SD, Crable BR, Palumbo AV, Somenahally AC, Elias DA. Global prevalence and distribution of genes and microorganisms involved in mercury methylation. Science Advantage. 2015. doi: e1500675.
140. Parks JM, Johs A, Podar M, Bridou R, Hurt RA, Jr., Smith SD, Tomanicek SJ, Qian Y, Brown SD, Brandt CC, Palumbo AV, Smith JC, Wall JD, Elias DA, Liang L. The

genetic basis for bacterial mercury methylation. Science. 2013;339(6125):1332-5. doi: 10.1126/science.1230667. PubMed PMID: 23393089.
141. Al-Khodor S, Price CT, Kalia A, Abu Kwaik Y. Functional diversity of ankyrin repeats in microbial proteins. Trends Microbiol. 2010;18(3):132-9. doi: 10.1016/j.tim.2009.11.004. PubMed PMID: 19962898; PMCID: 2834824.
142. Konstantinidis KT, Tiedje JM. Trends between gene content and genome size in prokaryotic species with larger genomes. Proc Natl Acad Sci U S A. 2004;101(9):3160-5. doi: 10.1073/pnas.0308653100. PubMed PMID: 14973198; PMCID: 365760.
143. Ulrich LE, Zhulin IB. The MiST2 database: a comprehensive genomics resource on microbial signal transduction. Nucleic Acids Res. 2010;38(Database issue):D401-7. doi: 10.1093/nar/gkp940. PubMed PMID: 19900966; PMCID: 2808908.
144. Ulrich LE, Koonin EV, Zhulin IB. One-component systems dominate signal transduction in prokaryotes. Trends Microbiol. 2005;13(2):52-6. doi: 10.1016/j.tim.2004.12.006. PubMed PMID: 15680762; PMCID: 2756188.
145. Mittl PR, Schneider-Brachert W. Sel1-like repeat proteins in signal transduction. Cell Signal. 2007;19(1):20-31. doi: 10.1016/j.cellsig.2006.05.034. PubMed PMID: 16870393.
146. Hubber A, Roy CR. Modulation of host cell function by Legionella pneumophila type IV effectors. Annual review of cell and developmental biology. 2010;26:261-83. doi: 10.1146/annurev-cellbio-100109-104034. PubMed PMID: 20929312.
147. Swann R, Holmes B. Infective endocarditis caused by Kingella denitrificans. J Clin Pathol. 1984;37:1384-7.
148. Minamoto GY, Sordillo EM. Kingella denitrificans as a cause of granulomatous disease in a patient with AIDS. Clinical Infectious Diseases. 1992;15(6):1052-3.
149. Rajanna DM, Manickavasagam J, Jewes L, Capper R. Retropharyngeal abscess from an unusual organism - Kingella denitrificans - in a patient on low-dose methotrexate. Ear, Nose & Throat Journal. 2011;90(7):15-7.
150. Chiang WF, Cheng CJ, Chau T, Hsiao PJ, Lin SH. Peritoneal dialysis-related peritonitis due to Kingella denitrificans: the first case report. Perit Dial Int. 2014;34(7):817-9. doi: 10.3747/pdi.2013.00184. PubMed PMID: 25520494; PMCID: PMC4269516.

151. Moriel DG, Bertoldi I, Spagnuolo A, Marchi S, Rosini R, Nesta B, Pastorello I, Corea VA, Torricelli G, Cartocci E, Savino S, Scarselli M, Dobrindt U, Hacker J, Tettelin H, Tallon LJ, Sullivan S, Wieler LH, Ewers C, Pickard D, Dougan G, Fontana MR, Rappuoli R, Pizza M, Serino L. Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic Escherichia coli. Proc Natl Acad Sci U S A. 2010;107(20):9072-7. doi: 10.1073/pnas.0915077107. PubMed PMID: 20439758; PMCID: PMC2889118.
152. van Egmond M, Damen CA, van Spriel AB, Vidarsson G, Van Garderen E, van de Winkel JGJ. IgA and IgA Fc receptor. Trends in Immunology. 2001;22(4):205-11.
153. Pastorello I, Rossi Paccani S, Rosini R, Mattera R, Ferrer Navarro M, Urosev D, Nesta B, Lo Surdo P, Del Vecchio M, Rippa V, Bertoldi I, Gomes Moriel D, Laarman AJ, van Strijp JA, Daura X, Pizza M, Serino L, Soriani M. EsiB, a novel pathogenic Escherichia coli secretory immunoglobulin A-binding protein impairing neutrophil activation. mBio. 2013;4(4): e00206-13. doi: 10.1128/mBio.00206-13. PubMed PMID: 23882011; PMCID: PMC3735183.
154. Taichman NS, Dean RT, Sanderson CJ. Biochemical and morphological characterization of the killing of human monocytes by a leukotoxin derived from Actinobacillus actinomycetemcomitans. Infect Immun. 1980;28(1):258-68.
155. Mangan DF, Taichman NS, Lally ET, Wahl SM. Lethal effects of Actinobacillus actinomycetemcomitans leukotoxin on human T lymphocyes. Infect Immun. 1991;59(9):3267-73.
156. Slots J, Reynolds HS, Genco RJ. Actinobacillus actinomycetemcomitans in human periodontal disease: a cross-sectional microbiological investigation. Infect Immun. 1980;29(3):1013-20.
157. Linhartova I, Bumba L, Masin J, Basler M, Osicka R, Kamanova J, Prochazkova K, Adkins I, Hejnova-Holubova J, Sadilkova L, Morova J, Sebo P. RTX proteins: a highly diverse family secreted by a common mechanism. FEMS Microbiol Rev. 2010;34(6):1076-112. doi: 10.1111/j.1574-6976.2010.00231.x. PubMed PMID: 20528947; PMCID: PMC3034196.
158. Aberg CH, Kelk P, Johansson A. Aggregatibacter actinomycetemcomitans: virulence of its leukotoxin and association with aggressive periodontitis. Virulence.

2015;6(3):188-95. doi: 10.4161/21505594.2014.982428. PubMed PMID: 25494963; PMCID: 4601274.
159. Confer DL, Eaton JW. Phagocyte impotence caused by an invasive bacterial adenylate cyclase Science. 1982;217(4563):948-50.
160. Hanski E. Invasive adenylate cyclase toxin of Bordetella pertussis. Trends Biochem. 1989;14:459-63.
161. Mouallem M, Farfel Z, Hanski E. Bordetella pertussis adenylate cyclase toxin: intoxication of host cells by bacterial invasion. Infect Immun. 1990;58(11):3759-64.
162. Zanardo RC, Brancaleone V, Distrutti E, Fiorucci S, Cirino G, Wallace JL. Hydrogen sulfide is an endogenous modulator of leukocyte-mediated inflammation. FASEB J. 2006;20(12):2118-20. doi: 10.1096/fj.06-6270fje. PubMed PMID: 16912151.
163. Greabu M, Totan A, Miricescu D, Radulescu R, Virlan J, Calenic B. Hydrogen Sulfide, Oxidative Stress and Periodontal Diseases: A Concise Review. Antioxidants (Basel). 2016;5(1):1-13. doi: 10.3390/antiox5010003. PubMed PMID: 26805896; PMCID: PMC4808752.
164. Wurch L, Giannone RJ, Belisle BS, Swift C, Utturkar S, Hettich RL, Reysenbach AL, Podar M. Genomics-informed isolation and characterization of a symbiotic Nanoarchaeota system from a terrestrial geothermal environment. Nature communications. 2016;7:1-10. doi: 10.1038/ncomms12115. PubMed PMID: 27378076; PMCID: 4935971.
165. Giannone RJ, Huber H, Karpinets T, Heimerl T, Kuper U, Rachel R, Keller M, Hettich RL, Podar M. Proteomic characterization of cellular and molecular processes that enable the Nanoarchaeum equitans--Ignicoccus hospitalis relationship. PLoS One. 2011;6(8):e22942. Epub 2011/08/10. doi: 10.1371/journal.pone.0022942 PONE-D-11-09749 [pii]. PubMed PMID: 21826220; PMCID: 3149612.
166. Otwell AE, Callister SJ, Zink EM, Smith RD, Richardson RE. Comparative Proteomic Analysis of Desulfotomaculum reducens MI-1: Insights into the Metabolic Versatility of a Gram-Positive Sulfate- and Metal-Reducing Bacterium. Front Microbiol. 2016;7:191. doi: 10.3389/fmicb.2016.00191. PubMed PMID: 26925055; PMCID: PMC4759654.

167. Sahingur SE, Yeudall WA. Chemokine function in periodontal disease and oral cavity cancer. Front Immunol. 2015;6:214. doi: 10.3389/fimmu.2015.00214. PubMed PMID: 25999952; PMCID: PMC4419853.
168. Henderson B, Nair SP, Ward JM, Wilson M. Molecular pathogenicity of the oral opportunistic pathogen Actinobacillus actinomycetemcomitans. Annu Rev Microbiol. 2003;57:29-55. doi: 10.1146/annurev.micro.57.030502.090908. PubMed PMID: 14527274.
169. Ramage G, Lappin DF, Millhouse E, Malcolm J, Jose A, Yang J, Bradshaw DJ, Pratten JR, Culshaw S. The epithelial cell response to health and disease associated oral biofilm models. J Periodontal Res. 2017;52(3):325-33. doi: 10.1111/jre.12395. PubMed PMID: 27330034; PMCID: PMC5412879.
170. Stathopoulou PG, Benakanakere MR, Galicia JC, Kinane DF. Epithelial cell pro-inflammatory cytokine response differs across dental plaque bacterial species. J Clin Periodontol. 2010;37(1):24-9. doi: 10.1111/j.1600-051X.2009.01505.x. PubMed PMID: 20096064; PMCID: PMC2900159.
171. Darveau RP, Belton CM, Reife RA, Lamont RJ. Local chemokine paralysis, a novel pathogenic mechanism for Porphyromonas gingivalis. Infect Immun. 1998;66(4):1660-5. PubMed PMID: 9529095; PMCID: 108102.
172. Wollert T, Rollenhagen C, Langford GM, Sundstrom P. Human oral keratinocytes: a model system to analyze host-pathogen interactions. Methods Mol Biol. 2012;845:289-302. doi: 10.1007/978-1-61779-539-8_19. PubMed PMID: 22328382.
173. Li Y, He J, He Z, Zhou Y, Yuan M, Xu X, Sun F, Liu C, Li J, Xie W, Deng Y, Qin Y, VanNostrand JD, Xiao L, Wu L, Zhou J, Shi W, Zhou X. Phylogenetic and functional gene structure shifts of the oral microbiomes in periodontitis patients. ISME J. 2014;8(9):1879-91. doi: 10.1038/ismej.2014.28. PubMed PMID: 24671083; PMCID: PMC4139721.
174. Langendijk PS, Hagemann J, van der Hoeven JS. Sulfate-reducing bacteria in periodontal pockets and in healthy oral sites. J Clin Periodontol. 1999;26:596-9.
175. Calenic B, Yaegaki K, Murata T, Imai T, Aoyama I, Sato T, Ii H. Oral malodorous compound triggers mitochondrial-dependent apoptosis and causes genomic DNA

damage in human gingival epithelial cells. J Periodontal Res. 2010;45(1):31-7. doi: 10.1111/j.1600-0765.2008.01199.x. PubMed PMID: 19602115.
176. Dufton N, Natividad J, Verdu EF, Wallace JL. Hydrogen sulfide and resolution of acute inflammation: A comparative study utilizing a novel fluorescent probe. Sci Rep. 2012;2:499. doi: 10.1038/srep00499. PubMed PMID: 22787557; PMCID: PMC3391661.
177. Chi XP, Ouyang XY, Wang YX. Hydrogen sulfide synergistically upregulates Porphyromonas gingivalis lipopolysaccharide-induced expression of IL-6 and IL-8 via NF-kappaB signalling in periodontal fibroblasts. Arch Oral Biol. 2014;59(9):954-61. doi: 10.1016/j.archoralbio.2014.05.022. PubMed PMID: 24927331.
178. Szafranski SP, Deng ZL, Tomasch J, Jarek M, Bhuju S, Meisinger C, Kuhnisch J, Sztajer H, Wagner-Dobler I. Functional biomarkers for chronic periodontitis and insights into the roles of Prevotella nigrescens and Fusobacterium nucleatum; a metatranscriptome analysis. NPJ Biofilms Microbiomes. 2015;1:15017. doi: 10.1038/npjbiofilms.2015.17. PubMed PMID: 28721234; PMCID: PMC5515211.
179. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012;28(12):1647-9. doi: 10.1093/bioinformatics/bts199. PubMed PMID: 22543367; PMCID: PMC3371832.
180. Krzywinski MI, Schein JE, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. Circos: an information aesthetic for comparative genomics. Genome Res. 2009;19:1639-45. doi: 10.1101/gr.092759.109.
181. Clarkson SM, Giannone RJ, Kridelbaugh DM, Elkins JG, Guss AM, Michener JK. Construction and Optimization of a Heterologous Pathway for Protocatechuate Catabolism in Escherichia coli Enables Bioconversion of Model Aromatic Compounds. Appl Environ Microbiol. 2017;83(18):e01313-17. doi: 10.1128/AEM.01313-17. PubMed PMID: 28733280; PMCID: 5583500.
182. Tabb DL, Fernando CG, Chambers MC. MyriMatch: highly accurate tandem mass spectral peptide identification by multivariate hypergeometric analysis. J Proteome Res. 2007;6(2):654-61. doi: 10.1021/pr0604054. PubMed PMID: 17269722; PMCID: PMC2525619.

183. Ma ZQ, Dasari S, Chambers MC, Litton MD, Sobecki SM, Zimmerman LJ, Halvey PJ, Schilling B, Drake PM, Gibson BW, Tabb DL. IDPicker 2.0: Improved protein assembly with high discrimination peptide identification filtering. J Proteome Res. 2009;8(8):3872-81. doi: 10.1021/pr900360j. PubMed PMID: 19522537; PMCID: PMC2810655.
184. Taverner T, Karpievitch YV, Polpitiya AD, Brown JN, Dabney AR, Anderson GA, Smith RD. DanteR: an extensible R-based tool for quantitative analysis of -omics data. Bioinformatics. 2012;28(18):2404-6. doi: 10.1093/bioinformatics/bts449. PubMed PMID: 22815360; PMCID: PMC3436848.
185. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, Cox J. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods. 2016;13(9):731-40. doi: 10.1038/nmeth.3901. PubMed PMID: 27348712.
186. Cross KL, Chirania P, Xiong W, Beall CJ, Elkins JG, Giannone RJ, Griffen AL, Guss AM, Hettich RL, Joshi SS, Mokrzan EM, Martin RK, Zhulin IB, Leys EJ, Podar M. Insights into the Evolution of Host Association through the Isolation and Characterization of a Novel Human Periodontal Pathobiont, Desulfobulbus oralis. mBio. 2018;9(2). doi: 10.1128/mBio.02061-17.
187. van der Hoeven JS, de Jong MH, Camp PJM, van den Kieboom CWA. Competition between oral streptococcus species in the chemostat under alternating conditions of glucose limitation and excess. FEMS Microbiol Ecology. 1985;31:373-9.
188. D'zink JL, Socransky SS. Amino acid utilization by Fusobacterium nucleatum grown ina chemically defined medium. Oral Microbiology Immunology. 1990;5:172-4.
189. Rogers AH, Zilm PS, Gully NJ, Pfennig AL, Marsh PD. Aspects of the growth and metabolism of Fusobacterium nucleatum ATCC 10953. Oral Microbiology Immunology. 1991;6:250-5.
190. Zilm PS, Neville JG, Rogers AH. Growth pH and transient increases in amino acid availability influence polyglucose synthesis by Fusobacterium nucleatum grown in continuous culture. FEMS Microbiol Lett. 2002;215:203-8.
191. DeLong EF, Pace NR. Environmental diversity of bacteria and archaea. Syst Biol. 2001;50(4):470-8. Epub 2002/07/16. PubMed PMID: 12116647.

192. Zaremba-Niedzwiedzka K, Caceres EF, Saw JH, Backstrom D, Juzokaite L, Vancaester E, Seitz KW, Anantharaman K, Starnawski P, Kjeldsen KU, Stott MB, Nunoura T, Banfield JF, Schramm A, Baker BJ, Spang A, Ettema TJ. Asgard archaea illuminate the origin of eukaryotic cellular complexity. Nature. 2017;541(7637):353-8. doi: 10.1038/nature21031. PubMed PMID: 28077874.
193. Parks DH, Rinke C, Chuvochina M, Chaumeil PA, Woodcroft BJ, Evans PN, Hugenholtz P, Tyson GW. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. Nature microbiology. 2017;2(11):1533-42. Epub 2017/09/13. doi: 10.1038/s41564-017-0012-7. PubMed PMID: 28894102.
194. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng JF, Darling A, Malfatti S, Swan BK, Gies EA, Dodsworth JA, Hedlund BP, Tsiamis G, Sievert SM, Liu WT, Eisen JA, Hallam SJ, Kyrpides NC, Stepanauskas R, Rubin EM, Hugenholtz P, Woyke T. Insights into the phylogeny and coding potential of microbial dark matter. Nature. 2013;499(7459):431-7. doi: 10.1038/nature12352. PubMed PMID: 23851394.
195. Gutleben J, Chaib De Mares M, van Elsas JD, Smidt H, Overmann J, Sipkema D. The multi-omics promise in context: from sequence to microbial isolate. Critical reviews in microbiology. 2017:1-18. doi: 10.1080/1040841X.2017.1332003. PubMed PMID: 28562180.
196. Huber H, Hohn MJ, Rachel R, Fuchs T, Wimmer VC, Stetter KO. A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. Nature. 2002;417(6884):63-7. PubMed PMID: 11986665.
197. St. John E, Liu Y, Podar M, Stott MB, Meneghin J, Chen Z, Lagutin K, Mitchell K, Reysenbach A-L. A new symbiotic nanoarchaeote (Candidatus Nanoclepta minutus) and its host (Zestosphaera tikiterensis gen. nov., sp. nov.) from a New Zealand hot spring. Systematic and applied microbiology. 2018. doi: 10.1016/j.syapm.2018.08.005.
198. Berdy B, Spoering AL, Ling LL, Epstein SS. In situ cultivation of previously uncultivable microorganisms using the ichip. Nat Protoc. 2017;12(10):2232-42. doi: 10.1038/nprot.2017.074. PubMed PMID: 29532802.
199. D'Onofrio A, Crawford JM, Stewart EJ, Witt K, Gavrish E, Epstein S, Clardy J, Lewis K. Siderophores from neighboring organisms promote the growth of uncultured

bacteria. Chemistry & biology. 2010;17(3):254-64. doi: 10.1016/j.chembiol.2010.02.010. PubMed PMID: 20338517; PMCID: 2895992.
200. Lagier JC, Khelaifia S, Alou MT, Ndongo S, Dione N, Hugon P, Caputo A, Cadoret F, Traore SI, Seck EH, Dubourg G, Durand G, Mourembou G, Guilhot E, Togo A, Bellali S, Bachar D, Cassir N, Bittar F, Delerce J, Mailhe M, Ricaboni D, Bilen M, Dangui Nieko NP, Dia Badiane NM, Valles C, Mouelhi D, Diop K, Million M, Musso D, Abrahao J, Azhar EI, Bibi F, Yasir M, Diallo A, Sokhna C, Djossou F, Vitton V, Robert C, Rolain JM, La Scola B, Fournier PE, Levasseur A, Raoult D. Culture of previously uncultured members of the human gut microbiota by culturomics. Nat Microbiol. 2016;1:16203. Epub 2016/11/08. doi: 10.1038/nmicrobiol.2016.203. PubMed PMID: 27819657.
201. Strandwitz P, Kim KH, Terekhova D, Liu JK, Sharma A, Levering J, McDonald D, Dietrich D, Ramadhar TR, Lekbua A, Mroue N, Liston C, Stewart EJ, Dubin MJ, Zengler K, Knight R, Gilbert JA, Clardy J, Lewis K. GABA-modulating bacteria of the human gut microbiota. Nature microbiology. 2018. Epub 2018/12/12. doi: 10.1038/s41564-018-0307-3. PubMed PMID: 30531975.
202. Zengler K, Toledo G, Rappe M, Elkins J, Mathur EJ, Short JM, Keller M. Cultivating the uncultured. Proc Natl Acad Sci U S A. 2002;99(24):15681-6. Epub 2002/11/20. doi: 10.1073/pnas.252630999. PubMed PMID: 12438682.
203. Hugenholtz P, Goebel BM, Pace NR. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. J Bacteriol. 1998;180(18):4765-74. PubMed PMID: 9733676.
204. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, Hugenholtz P. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol. 2018. Epub 2018/08/28. doi: 10.1038/nbt.4229. PubMed PMID: 30148503.
205. Dewhirst FE, Chen T, Izard J, Paster BJ, Tanner ACR, Yu W-H, Lakshmanan A, Wade WG. The Human Oral Microbiome. Journal of Bacteriology. 2010;192(19):5002-17. doi: 10.1128/jb.00542-10. PubMed PMID: WOS:000281866900022.
206. Brinig MM, Lepp PW, Ouverney CC, Armitage GC, Relman DA. Prevalence of bacteria of division TM7 in human subgingival plaque and their association with

disease. Appl Environ Microbiol. 2003;69(3):1687-94. Epub 2003/03/07. PubMed PMID: 12620860.

207. Dewhirst FE, Klein EA, Thompson EC, Blanton JM, Chen T, Milella L, Buckley CM, Davis IJ, Bennett ML, Marshall-Jones ZV. The canine oral microbiome. PLoS One. 2012;7(4):e36067. doi: 10.1371/journal.pone.0036067. PubMed PMID: 22558330; PMCID: 3338629.

208. Dudek NK, Sun CL, Burstein D, Kantor RS, Aliaga Goltsman DS, Bik EM, Thomas BC, Banfield JF, Relman DA. Novel Microbial Diversity and Functional Potential in the Marine Mammal Oral Microbiome. Curr Biol. 2017. doi: 10.1016/j.cub.2017.10.040. PubMed PMID: 29153320.

209. Haghighat S, Siadat SD, Sorkhabadi SMR, Sepahi AA, Mahdavi M. A novel recombinant vaccine candidate comprising PBP2a and autolysin against Methicillin Resistant Staphylococcus aureus confers protection in the experimental mice. Mol Immunol. 2017;91:1-7. Epub 2017/09/01. doi: 10.1016/j.molimm.2017.08.013. PubMed PMID: 28858628.

210. Lovering AL, de Castro LH, Lim D, Strynadka NC. Structural insight into the transglycosylation step of bacterial cell-wall biosynthesis. Science. 2007;315(5817):1402-5. Epub 2007/03/10. doi: 10.1126/science.1136611. PubMed PMID: 17347437.

211. Simons KT, Ruczinski I, Kooperberg C, Fox BA, Bystroff C, Baker D. Improved recognition of native-like protein structures using a combination of sequence-dependent and sequence-independent features of proteins. Proteins. 1999;34(1):82-95. Epub 1999/05/21. PubMed PMID: 10336385.

212. Palmer RJ, Jr., Shah N, Valm A, Paster B, Dewhirst F, Inui T, Cisar JO. Interbacterial Adhesion Networks within Early Oral Biofilms of Single Human Hosts. Appl Environ Microbiol. 2017;83(11). Epub 2017/03/28. doi: 10.1128/AEM.00407-17. PubMed PMID: 28341674; PMCID: PMC5440702.

213. Jakubovics NS, Yassin SA, Rickard AH. Community interactions of oral streptococci. Adv Appl Microbiol. 2014;87:43-110. Epub 2014/03/04. doi: 10.1016/B978-0-12-800261-2.00002-5. PubMed PMID: 24581389.

214. Heym B, Gehanno P, Friocourt V, Bougnoux ME, Le Moal M, Husson C, Leibowitch J, Nicolas-Chanoine MH. Molecular detection of Cellulosimicrobium cellulans as the etiological agent of a chronic tongue ulcer in a human immunodeficiency virus-positive patient. J Clin Microbiol. 2005;43(8):4269-71. Epub 2005/08/06. doi: 10.1128/JCM.43.8.4269-4271.2005. PubMed PMID: 16081997; PMCID: PMC1234009.
215. Castelle CJ, Brown CT, Anantharaman K, Probst AJ, Huang RH, Banfield JF. Biosynthetic capacity, metabolic variety and unusual biology in the CPR and DPANN radiations. Nature Reviews Microbiology. 2018;16(10):629-45. doi: 10.1038/s41579-018-0076-2.
216. McLean JS, Bor B, To TT, Liu Q, Kerns KA, Solden L, Wrighton KC, He X, Shi W. Evidence of independent acquisition and adaption of ultra-small bacteria to human hosts across the highly diverse yet reduced genomes of the phylum Saccharibacteria. bioRxiv. 2018.
217. Delmont TO, Eren AM. Linking pangenomes and metagenomes: the Prochlorococcus metapangenome. PeerJ. 2018;6:e4320. doi: 10.7717/peerj.4320. PubMed PMID: 29423345; PMCID: PMC5804319.
218. Eren AM, Esen OC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ. 2015;3:e1319. Epub 2015/10/27. doi: 10.7717/peerj.1319. PubMed PMID: 26500826; PMCID: PMC4614810.
219. Heimerl T, Flechsler J, Pickl C, Heinz V, Salecker B, Zweck J, Wanner G, Geimer S, Samson RY, Bell SD, Huber H, Wirth R, Wurch L, Podar M, Rachel R. A Complex Endomembrane System in the Archaeon Ignicoccus hospitalis Tapped by Nanoarchaeum equitans. Frontiers in Microbiology. 2017;8. doi: 10.3389/fmicb.2017.01072.
220. Mergaert P, Kikuchi Y, Shigenobu S, Nowack ECM. Metabolic Integration of Bacterial Endosymbionts through Antimicrobial Peptides. Trends Microbiol. 2017. doi: 10.1016/j.tim.2017.04.007. PubMed PMID: 28549825.
221. Epstein SS. The phenomenon of microbial uncultivability. Curr Opin Microbiol. 2013;16:636-42. doi: 10.1016/j.mib.2013.08.003.

222. Overmann J, Abt B, Sikorski J. Present and Future of Culturing Bacteria. Annu Rev Microbiol. 2017;71:711-30. doi: 10.1146/annurev-.
223. Markowitz VM, Korzeniewski F, Palaniappan K, Szeto E, Werner G, Padki A, Zhao X, Dubchak I, Hugenholtz P, Anderson I, Lykidis A, Mavromatis K, Ivanova N, Kyrpides NC. The integrated microbial genomes (IMG) system. Nucleic Acids Res. 2006;34(Database issue):D344-8. Epub 2005/12/31. PubMed PMID: 16381883.
224. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 2001;305(3):567-80. Epub 2001/01/12. doi: 10.1006/jmbi.2000.4315. PubMed PMID: 11152613.
225. Sauvage E, Kerff F, Terrak M, Ayala JA, Charlier P. The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. FEMS Microbiology Reviews. 2008;32(2):234-58. doi: 10.1111/j.1574-6976.2008.00105.x.
226. Haghighat S, Siadat SD, Sorkhabadi SM, Sepahi AA, Mahdavi M. Cloning, Expression and Purification of Penicillin Binding Protein2a (PBP2a) from Methicillin Resistant Staphylococcus aureus: A Study on Immunoreactivity in Balb/C Mouse. Avicenna J Med Biotechnol. 2013;5(4):204-11. Epub 2013/11/29. PubMed PMID: 24285994; PMCID: PMC3838764.
227. Zarantonelli ML, Antignac A, Lancellotti M, Guiyoule A, Alonso JM, Taha MK. Immunogenicity of meningococcal PBP2 during natural infection and protective activity of anti-PBP2 antibodies against meningococcal bacteraemia in mice. J Antimicrob Chemother. 2006;57(5):924-30. Epub 2006/03/04. doi: 10.1093/jac/dkl066. PubMed PMID: 16513914.
228. Byrne JP, Morona JK, Paton JC, Morona R. Identification of Streptococcus pneumoniae Cps2C residues that affect capsular polysaccharide polymerization, cell wall ligation, and Cps2D phosphorylation. J Bacteriol. 2011;193(9):2341-6. Epub 2011/03/08. doi: 10.1128/JB.00074-11. PubMed PMID: 21378192; PMCID: PMC3133057.
229. Toniolo C, Balducci E, Romano MR, Proietti D, Ferlenghi I, Grandi G, Berti F, Ros IM, Janulczyk R. Streptococcus agalactiae capsule polymer length and attachment is determined by the proteins CpsABCD. J Biol Chem. 2015;290(15):9521-32. Epub

2015/02/11. doi: 10.1074/jbc.M114.631499. PubMed PMID: 25666613; PMCID: PMC4392257.
230. Morona R, Purins L, Tocilj A, Matte A, Cygler M. Sequence-structure relationships in polysaccharide co-polymerase (PCP) proteins. Trends Biochem Sci. 2009;34(2):78-84. Epub 2008/12/09. doi: 10.1016/j.tibs.2008.11.001. PubMed PMID: 19058968.
231. Saha S, Raghava GP. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins. 2006;65(1):40-8. Epub 2006/08/09. doi: 10.1002/prot.21078. PubMed PMID: 16894596.
232. Larsen JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. Immunome Res. 2006;2:2. Epub 2006/04/26. doi: 10.1186/1745-7580-2-2. PubMed PMID: 16635264; PMCID: PMC1479323.
233. Heimerl T, Flechsler J, Pickl C, Heinz V, Salecker B, Zweck J, Wanner G, Geimer S, Samson RY, Bell SD, Huber H, Wirth R, Wurch L, Podar M, Rachel R. A Complex Endomembrane System in the Archaeon Ignicoccus hospitalis Tapped by Nanoarchaeum equitans. Front Microbiol. 2017;8:1072. Epub 2017/07/01. doi: 10.3389/fmicb.2017.01072. PubMed PMID: 28659892; PMCID: PMC5468417.
234. Miller LD, Mosher JJ, Venkateswaran A, Yang ZK, Palumbo AV, Phelps TJ, Podar M, Schadt CW, Keller M. Establishment and metabolic analysis of a model microbial community for understanding trophic and electron accepting interactions of subsurface anaerobic environments. BMC Microbiol. 2010;10:149. Epub 2010/05/26. doi: 10.1186/1471-2180-10-149. PubMed PMID: 20497531; PMCID: PMC2906461.
235. Podar M, Makarova KS, Graham DE, Wolf YI, Koonin EV, Reysenbach AL. Insights into archaeal evolution and symbiosis from the genomes of a nanoarchaeon and its inferred crenarchaeal host from Obsidian Pool, Yellowstone National Park. Biol Direct. 2013;8:9. Epub 2013/04/24. doi: 10.1186/1745-6150-8-9. PubMed PMID: 23607440; PMCID: PMC3655853.
236. Rinke C, Lee J, Nath N, Goudeau D, Thompson B, Poulton N, Dmitrieff E, Malmstrom R, Stepanauskas R, Woyke T. Obtaining genomes from uncultivated environmental microorganisms using FACS-based single-cell genomics. Nat Protoc. 2014;9(5):1038-48. doi: 10.1038/nprot.2014.067. PubMed PMID: 24722403.

237. Wong L, Sissons C. A comparison of human dental plaque microcosm biofilms grown in an undefined medium and a chemically defined artificial saliva. Arch Oral Biol. 2001;46(6):477-86. Epub 2001/04/20. PubMed PMID: 11311195.
238. Pernthaler A, Pernthaler J. Fluorescence in situ hybridization for the identification of environmental microbes. Methods Mol Biol. 2007;353:153-64. Epub 2007/03/03. doi: 1-59745-229-7:153 [pii]. PubMed PMID: 17332640.
239. Lundberg DS, Yourstone S, Mieczkowski P, Jones CD, Dangl JL. Practical innovations for high-throughput amplicon sequencing. Nat Methods. 2013;10(10):999-1002. doi: 10.1038/nmeth.2634. PubMed PMID: 23995388.
240. Martin M. Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. EMBnetjournal. doi: https://doi.org/10.14806/ej.17.1.200.
241. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26(19):2460-1. doi: 10.1093/bioinformatics/btq461. PubMed PMID: 20709691.
242. Caporaso J, Kuczynski J, Stombaugh J, Bittinger K. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010;7:335-6. PubMed PMID: 742666915146722492 8related:YD-OvT7UEGcJ.
243. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19(5):455-77. doi: 10.1089/cmb.2012.0021. PubMed PMID: 22506599; PMCID: PMC3342519.
244. Brown CT, Hug LA, Thomas BC, Sharon I, Castelle CJ, Singh A, Wilkins MJ, Wrighton KC, Williams KH, Banfield JF. Unusual biology across a group comprising more than 15% of domain Bacteria. Nature. 2015;523(7559):208-11. doi: 10.1038/nature14486. PubMed PMID: 26083755.
245. Hug LA, Thomas BC, Sharon I, Brown CT, Sharma R, Hettich RL, Wilkins MJ, Williams KH, Singh A, Banfield JF. Critical biogeochemical functions in the subsurface are associated with bacteria from new phyla and little studied lineages. Environ Microbiol. 2015. doi: 10.1111/1462-2920.12930. PubMed PMID: 26033198.

246. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25(7):1043-55. doi: 10.1101/gr.186072.114. PubMed PMID: 25977477; PMCID: PMC4484387.
247. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068-9. doi: 10.1093/bioinformatics/btu153. PubMed PMID: 24642063.
248. Anantharaman K, Brown CT, Hug LA, Sharon I, Castelle CJ, Probst AJ, Thomas BC, Singh A, Wilkins MJ, Karaoz U, Brodie EL, Williams KH, Hubbard SS, Banfield JF. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. Nat Commun. 2016;7:13219. doi: 10.1038/ncomms13219. PubMed PMID: 27774985; PMCID: PMC5079060.
249. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer EL, Tate J, Punta M. Pfam: the protein families database. Nucleic Acids Res. 2014;42(Database issue):D222-30. doi: 10.1093/nar/gkt1223. PubMed PMID: 24288371; PMCID: 3965110.
250. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;30(9):1236-40. doi: 10.1093/bioinformatics/btu031. PubMed PMID: 24451626; PMCID: PMC3998142.
251. Balakrishnan S, Kamisetty H, Carbonell JG, Lee SI, Langmead CJ. Learning generative models for protein fold families. Proteins. 2011;79(4):1061-78. Epub 2011/01/27. doi: 10.1002/prot.22934. PubMed PMID: 21268112.
252. Kamisetty H, Ovchinnikov S, Baker D. Assessing the utility of coevolution-based residue-residue contact predictions in a sequence- and structure-rich era. Proc Natl Acad Sci U S A. 2013;110(39):15674-9. Epub 2013/09/07. doi: 10.1073/pnas.1314045110. PubMed PMID: 24009338; PMCID: PMC3785744.
253. Ovchinnikov S, Park H, Varghese N, Huang PS, Pavlopoulos GA, Kim DE, Kamisetty H, Kyrpides NC, Baker D. Protein structure determination using metagenome sequence data. Science. 2017;355(6322):294-8. doi: 10.1126/science.aah4043. PubMed PMID: 28104891.

254. Remmert M, Biegert A, Hauser A, Soding J. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat Methods. 2011;9(2):173-5. Epub 2011/12/27. doi: 10.1038/nmeth.1818. PubMed PMID: 22198341.
255. Johnson LS, Eddy SR, Portugaly E. Hidden Markov model speed heuristic and iterative HMM search procedure. BMC Bioinformatics. 2010;11:431. Epub 2010/08/20. doi: 10.1186/1471-2105-11-431. PubMed PMID: 20718988; PMCID: PMC2931519.
256. Burley SK, Berman HM, Kleywegt GJ, Markley JL, Nakamura H, Velankar S. Protein Data Bank (PDB): The Single Global Macromolecular Structure Archive. Methods Mol Biol. 2017;1607:627-41. Epub 2017/06/03. doi: 10.1007/978-1-4939-7000-1_26. PubMed PMID: 28573592; PMCID: PMC5823500.
257. Sung MT, Lai YT, Huang CY, Chou LY, Shih HW, Cheng WC, Wong CH, Ma C. Crystal structure of the membrane-bound bifunctional transglycosylase PBP1b from Escherichia coli. Proc Natl Acad Sci U S A. 2009;106(22):8824-9. Epub 2009/05/22. doi: 10.1073/pnas.0904030106. PubMed PMID: 19458048; PMCID: PMC2689995.
258. Han S, Caspers N, Zaniewski RP, Lacey BM, Tomaras AP, Feng X, Geoghegan KF, Shanmugasundaram V. Distinctive attributes of beta-lactam target proteins in Acinetobacter baumannii relevant to development of new antibiotics. Journal of the American Chemical Society. 2011;133(50):20536-45. Epub 2011/11/05. doi: 10.1021/ja208835z. PubMed PMID: 22050378.
259. Yuan Y, Barrett D, Zhang Y, Kahne D, Sliz P, Walker S. Crystal structure of a peptidoglycan glycosyltransferase suggests a model for processive glycan chain synthesis. Proc Natl Acad Sci U S A. 2007;104(13):5348-53. Epub 2007/03/16. doi: 10.1073/pnas.0701160104. PubMed PMID: 17360321; PMCID: PMC1817829.
260. Huang CY, Shih HW, Lin LY, Tien YW, Cheng TJ, Cheng WC, Wong CH, Ma C. Crystal structure of Staphylococcus aureus transglycosylase in complex with a lipid II analog and elucidation of peptidoglycan synthesis mechanism. Proc Natl Acad Sci U S A. 2012;109(17):6496-501. Epub 2012/04/12. doi: 10.1073/pnas.1203900109. PubMed PMID: 22493270; PMCID: PMC3340074.
261. Gront D, Kulp DW, Vernon RM, Strauss CE, Baker D. Generalized fragment picking in Rosetta: design, protocols and applications. PLoS One. 2011;6(8):e23294.

Epub 2011/09/03. doi: 10.1371/journal.pone.0023294. PubMed PMID: 21887241; PMCID: PMC3160850.

262. Moutsopoulos NM, Madianos PN. Low-grade inflammation in chronic infectious diseases: paradigm of periodontal infections. Ann N Y Acad Sci. 2006;1088:251-64. doi: 10.1196/annals.1366.032. PubMed PMID: 17192571.

Vita

Karissa Lynn Cross was born in Phoenix, Arizona. At the age of five she and her twin sister moved to Murphy, North Carolina where they were raised by their grandparents Nancy and Lowell Williamson. She attended a small K-12 school named Hiwassee Dam in beautiful Cherokee County. After graduating high school, Karissa attended Young Harris College in Georgia where she was a biology major. In her coursework there, she fell in love with microbiology which prompted her to transfer to North Carolina State University in Raleigh, NC to major in their microbiology program. While there, Karissa held two separate undergraduate research position with Drs. Jennifer Miller and H.T. Banks that inspired her to pursue a PhD. Following graduation with a BS in Microbiology, Karissa moved to Knoxville, TN where she first interned at Oak Ridge National Laboratory with Dr. Jim Elkins before beginning her PhD at the University of Tennessee in the Department of Microbiology under the mentorship of Dr. Mircea Podar.