

Metagenomics: A New Realm of Uncultured Periodontal Microbes

L Pathak

Abstract

With over 400 bacterial species the periodontal microbiota has clearly a diverse ecology and uncultivated organisms are found to be associated with the periodontal diseases. Many efforts have been made to determine why so many bacterial species cannot be cultured on conventional bacteriological media and to develop new cultural methods for their isolation. In recent years, metagenomic analysis is available, based on the direct examination of nucleic acids from samples. This review will address the new dimensions of this novel procedure which will lead to a better understanding of the role of oral microbiota in periodontal disease.

Keywords: *Metagenomics, Human microbiome, Superorganism.*

Introduction

The microbes inhabiting the human mouth consists of several thousand bacterial types which are of different phylogeny. These microbes are organized into complex communities specifically adapted to inhabit different oral niches. In contrast to the traditional view of individual pathogens being responsible for disease onset, recent microbial ecosystem diversity analyses seem to point to a new perspective in which the transition from health to disease is attributed to a shift in the microbial flora rather than to the specific appearance of individual pathogens. Many bacteria also secrete factors that limit growth of other species, which substantially alters the entire oral ecosystem. Current quantitative analyses are heavily limited by the unavailability of relevant data. The term 'Metagenomics' which was coined by Handelsman in 1998 is an approach by which the analysis of the collective genomes of microbes within the entire oral microbial community can be made possible¹.

This approach generates billions of nucleotide bases at a fraction of the cost per base of traditional methods². Dental plaque pooled from 98 healthy adults revealed about about 10,000 microbial phylotypes which is of a magnitude higher than the previously reported 700 oral microbial phylotypes as identified by cultivation or traditional cloning and sequencing.

The total diversity of the global oral microbiome can be estimated to be around 25,000 phylotypes. An understanding of the metagenomic approach is essential to identify those strongly associated with periodontal disease, as well as oral flora specifically.

The Superorganism

Albert Einstein once said that "The true value of a human being can be found in the degree to which he has attained liberation from the self". For years our traditional view of 'self' was restricted to our own bodies; composed of eukaryote cells encoded by our genome. However, in the era of omics technologies and systems biology, this view now extends beyond the traditional limitations of our own core being to include our resident microbial communities. These prokaryote cells outnumber our own cells by a factor of ten and contain at least ten times more DNA than our own genome. In exchange for food and shelter, this symbiont provides the host, with metabolic functions far beyond the scope of our own physiological capabilities. In this respect the human body can be considered a superorganism; a communal group of human and microbial cells all working together for the benefit of the collective - a view which most certainly attains liberation from self.

In 2007, Human Microbiome Project was launched by the US National Institute (NIH) to characterize the extent and diversity of the human body in a site specific manner and, to identify its link to health and disease.³

Address of Correspondence:

Dr. Liza Pathak, MDS in Periodontology, Reader, Department of Periodontology and Oral Implantology Regional Dental College, Guwahati-781032, India.
E-mail: liza0368@gmail.com

This project included all the microbes, their genetic elements (genomes), and environmental interactions in a defined environment such as the gastrointestinal tract, skin, genitourinary tract and the oral cavity. The oral microflora forms an indispensable part of the human microbiome, and present a diverse and unique ecology in various niches such as the gingival sulcus and periodontal pockets, tongue dorsum and other mucosal surfaces including the hard tissue surfaces of teeth. The microorganisms in the human oral cavity have been referred to as the oral microflora, oral microbiota, or more recently as the oral microbiome. The investigators in the Human Microbiome Project believe that an understanding of human health and disease is impossible without fully understanding the collective microbiome where the entire human body space is regarded as the human “superorganism”⁴. The mere presence of pathogens in the periodontal pocket is not sufficient to trigger periodontitis. The disease might be correlated with the presence of specific virulence factors within the genomes of particular pathogens, or might be initiated once the abundance of one or more pathogens cross a specific threshold. The mechanisms that keep pathogenic bacteria ‘in check’ during health but allow them to bloom during disease are not yet understood. These observations support our suggestion that a full understanding of periodontal disease requires whole-genome and whole-system analyses.

There is a vast imbalance between what it is possible to hypothesize and test, and what is unknown which means that every microbial ecologist is on an epic voyage of equal importance to that of Darwin. There are many fundamental theories about microbial life that still need to be examined, and many of these can only be explored by intelligent sampling in an unrestricted environmental surveys. Restricted analysis such as laboratory based manipulation, culturing, PCR amplification and genome experimentation are very important to the understanding of microbial adaptation. However, lab experiments are artificial and hence it will always be necessary to contextualize these results with environmental observation, and DNA extraction. Metagenomics can be the most unrestricted and comprehensive approach.

Metagenomics and Oral Microbiota

The first study of metagenomics for oral microbiota was performed over a decade ago and examined the cultivable and uncultivable bacterial species in dentoalveolar abscesses⁵.

Abscesses from three individuals were studied and although this study was not a true metagenomic analysis, it was seen that the number and the diversity of the bacterial species isolated through the method outnumbered the species isolated by cultivating the samples on traditional nonselective bacteriological media. This study identified previously undescribed lineages within the phyla *Bacteroidetes* and *Firmicutes* and also found that the numbers of cultivable organisms including *F. nucleatum* and *Porphyromonas endodontalis* were detected at higher levels.

Subsequent studies have confirmed that these are the predominant phyla found in oral niches. The HOMB database has allowed the most complete description of the periodontal microbiota to date. The genera found include representatives of the domain *Archaea*, the methanogenic species *Methanobrevibacter oralis* and related phylotypes, whose numbers have been shown to be raised in periodontitis⁶. A vast majority of oral bacteria belong to the domain Bacteria and include representatives of 11 phyla that can be detected from human oral samples. The phyla found are *Actinobacteria*, *Bacteroidetes*, *Chloroflexi*, *Firmicutes*, *Fusobacteria*, *Proteobacteria*, *Spirochaetes*, *Synergistetes* and *Tenericutes* and the unnamed Divisions TM7 and SR11. Members of the *Firmicutes* are the most frequently detected from periodontal samples and includes common oral genera such as *Streptococcus* and also a substantial number of as yet uncultured species, particularly within the family *Lachnospiraceae*. TM7 is a novel bacterial division with no cultivated representatives, and previous studies have shown microbes from this division to be commonly found in the human oral flora but at relatively low abundance, generally around 1% of the population though abundances as high as 8% were previously reported.

The high number of TM7 microbes found in plaque samples, in patients with periodontal disease, indicate that the prevalence of this poorly studied bacterial division within the oral cavity⁷. Its role in disease, have yet to be fully appreciated. The majority of bacterial phyla are Gram-negative diderms. Only the *Firmicutes*, the *Actinobacteria* and *Chloroflexi* are monoderms. Candidate phylum. TM7 is a close relative of the *Chloroflexi*. The abundance of TM7 rDNA was found to be higher in sites with mild periodontitis than in healthy sites.

Metagenomic Study Approaches

The metagenomic study starts with the isolation of genomic material from the organism present in a particular sample⁸. It involves the construction of clone libraries where the DNA is cloned into an appropriate large or small insert vector and it is transformed into a surrogate host which is typically *Escherichia coli*. Large insert libraries can be generated in bacterial chromosomes, fosmids or cosmids. These vector systems can accommodate DNA fragments varying from 20kb to 200kb in size. Small insert libraries are typically constructed using plasmid vectors like pUC19, pBlue-script SK+ or pSMART which contain tenfold more clones than their large insert counter parts. Resulting clone libraries are explored using Sequence and Function based metagenomic approaches.

Sequence Driven Analysis

The sequence driven approach requires no prior sequence knowledge for data generation. The vast amount of data has led to the discovery of millions of genes and deduced metabolic pathways and interactions among bacteria. The largest metagenomic sequencing project the Marine Microbiome Project predicted initially 6.12 million proteins representing all previous known families and ultimately added 1700 more families. The majority of the studies have used the shot-gun sequencing approach. This entails Sanger dideoxy sequencing of small-insert metagenomic clone libraries, producing sequence reads of 600-900 bp. Sequenced-based analysis can involve complete sequencing of clones containing phylogenetic anchors that indicate the taxonomic group that is the probable source of the DNA fragment. Alternatively, random sequencing can be conducted, and once a gene of interest is identified, phylogenetic anchors can be sought in the flanking DNA to provide a link of phylogeny with the functional gene.

A promising application of phylogenetic anchor-guided sequencing is to collect and sequence many genomic fragments from one taxon. In more complex environments and taxa, reassembly of a genome may not be feasible, but inference about the physiology and ecology of the members of the groups can be gleaned from sequence data. The collection of phylogenetic markers is growing, and as the diversity of markers increases, the power of this approach will also increase, making it possible to assign more fragments of anonymous DNA to the organisms from which they were isolated.

Moreover, as more genomes are reconstructed, more genes will be linked to phylogenetic markers even though they were not cloned initially on the same fragment. The sequence-based screening approach is limited to the identification of new members of known gene families. In general, target genes are identified either by PCR-based or hybridization-based approaches employing primers and probes derived from conserved regions of known genes and gene products.

Function Driven Analysis

The function-driven analysis is initiated by identification of clones that express a desired trait, followed by characterization of the active clones by sequence and biochemical analysis. This approach quickly identifies clones that have potential applications in medicine, agriculture or industry by focusing on natural products or proteins that have useful activities. The limitations of the approach are that it requires expression of the function of interest in the host cell and clustering of all of the genes required for the function. It also depends on the availability of an assay for the function of interest that can be performed efficiently on vast libraries, because the frequency of active clones is quite low. Many approaches are being developed to eliminate these limitations. Improved systems for heterologous gene expression are being developed with shuttle vectors that facilitate screening of the metagenomic DNA in diverse host species and with modifications of *Escherichia coli* to expand the range of gene expression. Although the genes encoding the enzymes required for synthesis of secondary metabolites are usually clustered on a contiguous piece of DNA, obtaining fragments of DNA large enough to contain the information required for synthesis of complex antibiotics, which can require over 100 genes, presents a challenge.

To address the challenge of detecting rare, active clones in large libraries, efforts are being directed toward the design of highly sensitive assays and robotic screens that efficiently detect low levels of activity in many samples. The most convenient traits to study are those that present a selectable phenotype, such as resistance to an antibiotic or growth on an unusual substrate, because selections are orders of magnitude more efficient than screens.

There is an inherent contradiction in this approach—genes are cloned from exotic organisms to discover new motifs in biology, and yet these genes are required to be expressed in *Escherichia coli* or another domesticated bacterium in order to be detected. The diversity of the organisms whose DNA has been successfully expressed in *E. coli* is surprising, but heterologous expression remains a barrier to extracting the maximum information from functional metagenomics analyses. One important attribute of the oral microbiota is its ability to act as a reservoir of antibiotic-resistant organisms. The bacteria from the oral niches can easily reach other body sites (by swallowing and via the bloodstream).

Therefore, antibiotic-resistant oral bacteria have the opportunity for rapid dissemination through the community and to transfer their resistance genes to other bacterial species. Virtually all investigations of the prevalence of antibiotic-resistant oral bacteria are based solely on the cultivable microbiota and there is no account of the large number of uncultured bacteria present in these complex communities. To overcome this, functional metagenomic approach can be used to identify the presence of genes encoding antibiotic resistance in the oral microbiota of adult humans⁹. This involves the generation of expression libraries from the total DNA present in saliva and plaque from the individuals and screening the resulting clones for antibiotic resistance. Function-driven screening of metagenomic libraries is not dependent on sequence information or sequence similarity to known genes. Thus, this is the only approach that bears the potential to discover new classes of genes that encode either known or new functions. A significant limitation of this technique is the dependence on expression of the target genes and production of functional gene products in a foreign host, which is in most studies *Escherichia coli*. Thus, the incapability to discover functional gene products or a low detection frequency during function-based screens of metagenomic libraries might be a result of the inability of the host to express the foreign genes and to form active recombinant proteins.

In addition, function-driven screening often requires the analysis of more clones than sequence-based screening for the recovery of a few positive clones. There is a major advantage of a function-based screening approach is that only full-length genes and functional gene products are detected.

16S rRNA gene

The 16S rRNA gene is the “gold-standard” of markers for culture-independent microbial studies because:

(1) It is found in all forms of cellular life: bacterial, archaeal and eukaryal

(2) The huge databases of 16S sequences from cultured and uncultured microbes, allow for easy identification and phylogenetic analyses

(3) This molecule contains both highly conserved regions necessary for universal primer design, and highly variable regions necessary for identifying the bacteria¹⁰. In the earlier studies, bacterial identification was based on the sequence data of relatively shorter segments of 300–400 bases. As the 16S rRNA gene is approximately 1550 bases in length, it could be surmised that more sensitive and specific information on taxonomical lineage and bacterial diversity might be obtained through analyzing longer lengths of sequences.

Direct sequencing of total environmental DNA has the potential to assess the true diversity of the environment. Metagenomic sequences can be assigned to taxa using their similarity to reference genomes based on either sequence similarity or genomic composition.

In contrast, several comprehensive databases exist for the 16S rRNA gene that provide detailed phylogenetic trees and allow for taxonomic resolution down to the species level. Shotgun metagenomic datasets obviously also contain fragmented 16S rRNA genes and these have been directly assigned to taxa through BLAST-based comparisons or phylogenetic distance-based clustering. However, the short and random nature of metagenomic sequences may not contain the phylogenetically most informative regions of the 16S rRNA genes, thus diminishing the efficiency of taxonomic assignments. Sequence assembly can potentially increase the length of the 16S rRNA gene sequences recovered, but low sequence coverage may limit assembly success for 16S rRNA genes and low-stringency assemblies may result in chimeric sequences. The recent technology uses iterative mapping of short Illumina reads against reference sequences to reconstruct 16S rRNA genes. Although this approach has an explicit accuracy to single nucleotide difference, its potential to avoid chimeras is strongly dependent on the quality of the reference database. The metagenomic approach has been used to determine the subgingival bacterial diversity profiles under different circumstances¹¹.

In addition, 16r DNA amplicons from 47 subgingival plaque samples subjected to denaturing gradient gel electrophoresis (DGGE) and sequencing of each unique band have shown that *Aggregatibacter actinomycetemcomitans* was higher among individuals with periodontitis¹².

This method is not infallible, for the whole methodology operates under the premise that enough genomic DNA is recovered from the clinical sample, the amplification process (PCR) is unbiased and that the database sequences against which the nascent sequences are compared are accurate.

Conclusion

The human oral metagenome is currently under investigation as part of the Human Microbiome Project. Making sense of the vast volume of data will be a major challenge but new bioinformatic tools will undoubtedly allow the accumulation of much valuable data of relevance to periodontal disease.

Although there is still great scope for application of the basic metagenomic approach to microbial communities - in making spatial series and in population genomics for example - researchers are making concerted efforts to extend and enhance metagenomics using techniques such as flow sorting, microfluidics, transcriptomics and proteomics. There are many other recently developed methods that can similarly be applied to build on or complement the basic metagenomic approach, including stable isotope probing, stable isotope mass spectroscopy and subcellular high-resolution imaging, guaranteeing a rich and interesting future for those who study microbial ecology and evolution.

It is the fundamental question of microbial ecology that will focus future research, and the interplay of different technologies will be paramount in answering these questions. For example, high-throughput 16S rRNA gene studies alone can significantly increase our concept of the diversity of life. The ultimate future goal of scientific research is to provide a far more detailed understanding of microbial ecology. Improved exchange of ideas and data will inevitably improve and advance the theory, perhaps, even help to define the basic rules for biological systems beyond the constant of nucleic acid.

References

1. Handelsman, J., M. R. Rondon, S. F. Brady, J. Clardy, and R. M. Goodman. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* 1998;5: 245-249
2. Handelsman J . Metagenomics: application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* 2004; 68:669–685
3. Turnbaugh P, Ley R, Hamady M, Fraser-Liggett C, Knight R, et al. (2007) The Human Microbiome Project. *Nature* 449: 804–810.
4. Sleator RD. The Human Superorganism - of microbes and men .2010 Feb; 74(2):214-5.
5. Dymock D, Weightman AJ, Scully C, Wade WG (1996). Molecular analysis of microflora associated with dentoalveolar abscesses. *J Clin Microbiol* 34: 537–5426.
6. Lepp PW, Brinig MM, Ouverney CC et al (2004). Methanogenic Archaea and human periodontal disease. *Proc Natl Acad Sci USA* 101:6176–6181
7. Tringe SG, von Mering C, Kobayashi A, Salamov AA, Chen K, Chang HW, Podar M, Short JM, Mathur EJ, Detter JC, et al.: Comparative metagenomics of microbial communities. *Science* 2005, 308:554-557.
8. Woyke T, Teeling H, Ivanova NN, Huntemann M, Richter M, Gloeckner FO, Boffelli D, Anderson IJ, Barry KW, Shapiro HJ, et al.: Symbiosis insights through metagenomic analysis of a microbial consortium. 2006. *Nature* 443:950.
9. Zaura E, Keijsers BJ, Huse SM, Crielaard W (2009) Defining the healthy “core microbiome” of oral microbial communities. *BMC Microbiol* 9: 259
10. Aas, J. A., B. J. Paster, L. N. Stokes, I. Olsen, and F. E. Dewhirst. 2005. Defining the normal bacterial flora of the oral cavity. *J. Clin. Microbiol.* 43:5721-5732
11. Lemon KP, Klepac-Ceraj V, Schiffer HK, Brodie EL, Lynch SV, Kolter R. Comparative analyses of the bacterial microbiota of the human nostril and oropharynx. *MBio.* 2010; 1:e00129–00110
12. Ge X, Rodriguez R, Trinh M, Gunsolley J, Xu P. Oral microbiome of deep and shallow dental pockets in chronic periodontitis. 2013. *Plos One* .8(6):e65520.