Practice

BIOINFORMATICS INSTITUTE

Dead Man's Teeth: taxonomic annotation of ancient human oral metagenome

Kristina Zheltova[1, 2] and Mikhail Filippov[1, 3]

[1]Bioinformatics institute
[2]ITMO University
[2]Herzen University

Abstract

Today we able to associate certain microbial markers with dental pathology, but there is little knowledge of how this data can be applied to paleopathology research. Microbial DNA can be preserved in paleonthological material by more that 1000 years, so modern metagenomic approaches theoretically can be used to estimate microbial abundance in ancient samples. In this study we will try to reproduce some steps from Christina Warriner's original research where state of art metagenomic approaches were applied to describe microbial features of 1100 A.D. human tooth samples from Germany. As a result, we characterised microbiomic differences between healthy and pathological ancient tooth samples and identified genomic differences between ancient and modern pathogenic bacteria.

Keywords: Oral microbiome, metagenome, paleopathology

Introduction

Metagenomic research, which refers to analysis not of the genome of one organism, but of the bunch of genomes from whole community (for example, soil sample), is the relatively new biological approach. This kind of study is able when DNA is more or less well preserved, because degradation severely affects output relative abundance of certain genomes in the sample. This becomes a great problem when we attempt to perform metagenomic study with old samples, such as ancient human teeth, which are the object of the current study. However, in some cases, DNA can resist degradation for thousands years. In the case of teeth, bacterial DNA can be found relatively intact in dental calculus - calcified dental plaques. In the study we try to reproduce (Warriner et al., 2014), dental calculus samples from the teeth with estimated age of 1100 A.D. were sequenced using two main approaches. One is the 16S rRNA V5 region sequencing, which allows researcher to estimate bacterial diversity of the sample and perform taxonomic annotation. Different taxonomic units in the sampled are referred as *operational taxonomic units* (OTUs). Another is the shotgun sequencing, where all DNA sequences are included it the analysis and assembled as *contigs*, continuous sequences of DNA, which usually interpreted as bacterial genomes. This approach allows to study genomic features of the organisms found in sample, for example, in our case, to compare modern and ancient bacterial genomes and identify their differences.

As the object of the study is teeth, we interested in microbial differences between healthy and diseased ones. Today we know, that certain bacteria is tightly correlated with peridontal pathology, and three most "evil" oral bacteria are *Porphyromonas gingivalis, Tannerella forsythia and Treponema denticola*, which referred as "red complex" (Rôças et al., 2001). We will try to identify, which samples contain all three of these species, and if their presence correlates with morphologically identified pathology of the teeth. Another aim is to compare ancient *Tannerella forsythia* genome from the sample with the modern one and identify the differences between them.

Figure 1 Teeth from Danhelm, Germany

Materials and methods

Data availability

All data used in this study is open for downloading:

Data from original Warriner's article is available at NCBI

Metagenome assembly stored at MG-RAST

T. forsynthia genome can be downloaded from Nuccore DB

There are total of 9 samples, which belong to 3 men, marked as B17, G12, B78 and B61. There are 2 samples for each tooth: one for root and one for calculus (except for G12, which has 2 samples for root)

Pipeline

We used QIIME2 (Bolyen et al., 2019) with DADA2 pipeline for performing microbiome analysis from raw DNA sequencing data including such stages as trimming, denoising, dereplication, quality control and clustering into amplicon sequence variants.

We aligned contigs to reference Tannarella forsythia genome with BWA (Li et al., 2009), then we sorted and indexed alignments with SAMtools (Li, Handsaker et al., 2009) and obtained alignment BAM file to BED with BEDtools (Quinlan et al., 2010). New regions in the modern strain absent in the ancient strain were obtained using BEDtools intersect.

Results

Taxonomic annotation resulted in 465 OTUs with median frequency of 4,862 per sample. All 9 samples are has different microbial patterns: they distinguished both by diversity and dominant bacterial phyla. As pictured on the *figure 4* (in the VERY end of the document), roots and calculus generally demonstrate different bacterial composition, as well as each sample is in general different from another. In terms of raw results, root is far more rich in Proteobacteria comparative to calculus, and Firmicutes (especially Clostridia) are more abundant in calculus. Samples from healthy teeth (which are B78 and B17) in average demonstrate more diverse composition, with less significant prevalence of one phylum, than samples from diseased teeth. Exception is B78 calculus, which has almost absent Proteobacterial component. Both 2 calculus samples from diseased teeth (G12 anf B61) have red complex bacteria DNA, and all healthy teeth, together with roots from diseased teeth, are absent with at least one of the red complex bacteria.

Speaking of comparsion of modern and ancient *T. forsythia*, there are 126 new annotated genetic elements after exclusion of those which annotated as "hypothetical". Distribution of their functions are showed in *figure 2*.

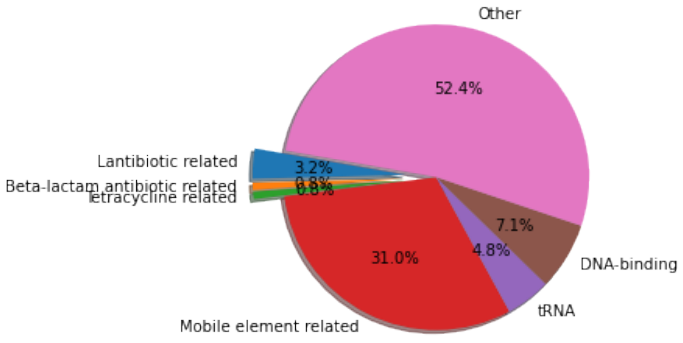

Figure 2 New elements in modern *T. forsythia* genome

Discussion

It is now well known, that oralmicrobial community become less diverse in pathologic conditions (Zhang et al., 2022). And it corresponds well to our data: two healthy teeth metagenomes are more diverse than ones from pathological teeth. Another marker of dental disease, the red complex, is too in accordance with expectations: calculus from corrupted teeth carried all 3 members of the complex. Overall, this is a brief demonstration that metagenomic approach can be applied to paleopathology research.

Interesting results were obtained from shotgun sequencing analysis. Not only modern *T. forsythia* has some new genetic regions, these regions are also have important biological functions. One large part of new genetic elements is related to transposones, and in our opinion this is an indicator of the fact of obtaining new genes in general, as transposase activity is a one of very few ways to get new genes at all. But most interesting part of new regions is the antibiotic resistance genes, which are, specifically, lanthionine synthetase C and lantibiotic dehydratase; glycosyltransferase; etracycline resistance ribosomal protection protein and TetR/AcrR family transcriptional regulator. Which are, respectively, genes responsible for resistance to lantibiotics, beta-lactam antibiotics and tetracycline. It is easy to explain why modern bacteria has these genes and ancient one doesn't: only in modern days usage of antibiotics created such selective power to force bacteria to evolve and gain antibiotics resistance. Rest new elements include tRNA, DNA-binding proteins, signal and transport proteins and other genes, which functions in modern bacteria is not so obvious.

References

- *Warinner C, Rodrigues JF, Vyas R, Trachsel C, Shved N, Grossmann J, Radini A, Hancock Y, Tito RY, Fiddyment S, Speller C, Hendy J, Charlton S, Luder HU, Salazar-García DC, Eppler E, Seiler R, Hansen LH, Castruita JA, Barkow-Oesterreicher S, Teoh KY, Kelstrup CD, Olsen JV, Nanni P, Kawai T, Willerslev E, von Mering C, Lewis CM Jr, Collins MJ, Gilbert MT, Rühli F, Cappellini E. Pathogens and host immunity in the ancient human oral cavity. Nat Genet. 2014 Apr;46(4):336-44. doi: 10.1038/ng.2906. Epub 2014 Feb 23. PMID: 24562188; PMCID: PMC3969750.*
- *Rôças IN, Siqueira JF, Santos KR, Coelho AM (April 2001). ""Red complex" (Bacteroides forsythus, Porphyromonas gingivalis, and Treponema denticola) in endodontic infections: a molecular approach". Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontics. 91 (4): 468–71. doi:10.1067/moe.2001.114379. PMID 11312465.*
- *Zhang JS, Chu CH, Yu OY. Oral Microbiome and Dental Caries Development. Dent J (Basel). 2022 Sep 30;10(10):184. doi: 10.3390/dj10100184. PMID: 36285994; PMCID: PMC9601200.*
- *Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu YX, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha*

R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, and Caporaso JG. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology 37: 852–857. https://doi.org/10.1038/s41587-019-0209-9

- Li *H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009 Jul 15;25(14):1754-60. doi: 10.1093/bioinformatics/btp324. Epub 2009 May 18. PMID: 19451168; PMCID: PMC2705234.*
- *Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078-9. doi: 10.1093/bioinformatics/btp352. Epub 2009 Jun 8. PMID: 19505943; PMCID: PMC2723002.*
- *Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010 Mar 15;26(6):841-2. doi: 10.1093/bioinformatics/btq033. Epub 2010 Jan 28. PMID: 20110278; PMCID: PMC2832824.*

Acknowledgments

We thank Mike Rayko for supervising and supporting us.

Figure 3 Random cat picture to get barplots in desired position

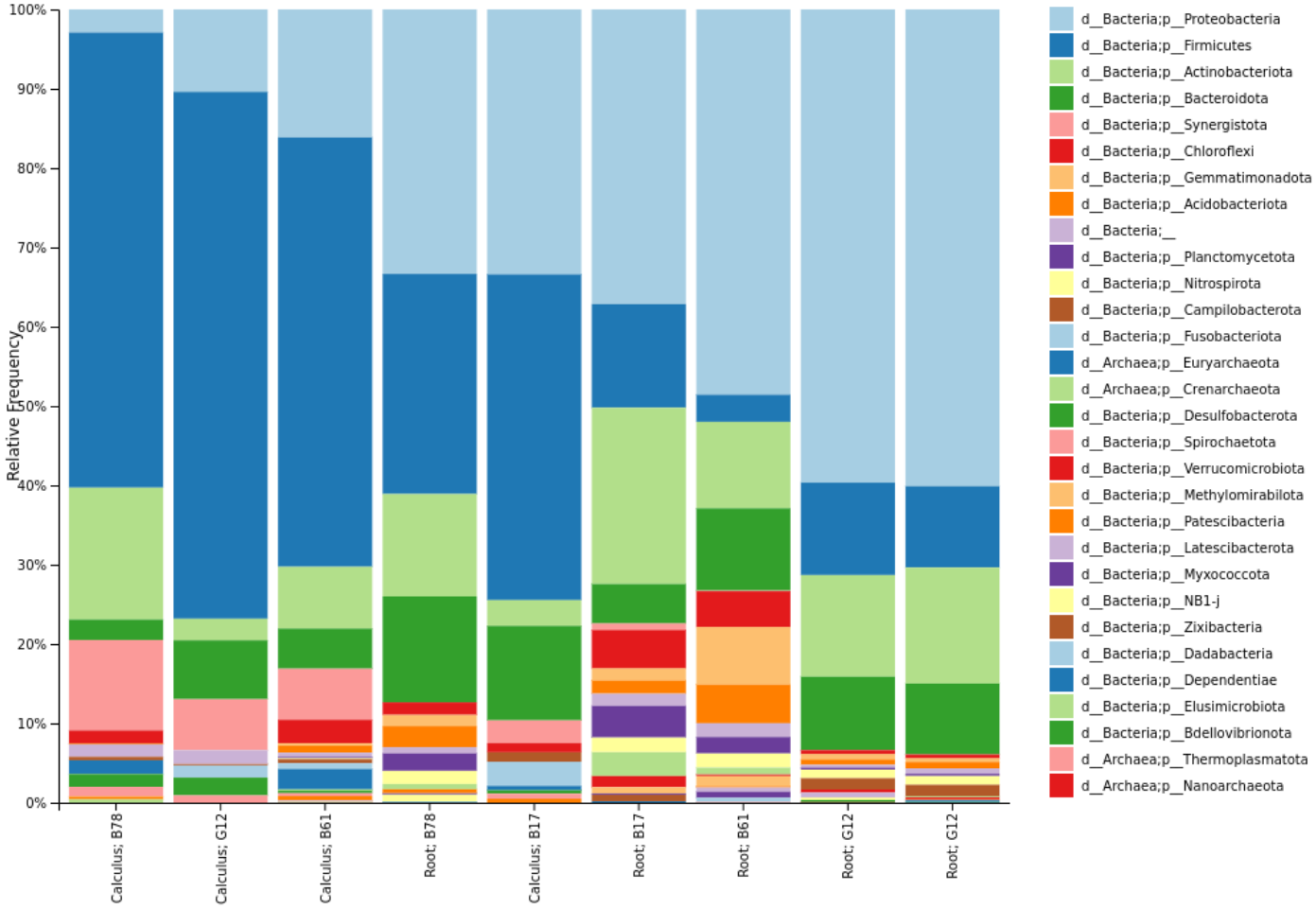

Figure 4 Relative abundance of bacteria phylum in different samples