Metagenomic analysis of ancient dental calculus

Abstract:

In this study we analysed amplicon and shotgun metagenomic sequencing data obtained from ancient human skulls. Dental calculus – a cover on the dental surface caused by periodontitis – is able to preserve DNA for 1000s of years, which allows researchers to extract and sequence these DNA samples. We explored microbial diversity in such samples and compared ancient sequences with the modern ones in order to get information about evolution processes.

Introduction:

Although we can't usually see microbes, they are essential for every part of our lives. All living organisms have closely associated microbial communities that make necessary nutrients, metals, and vitamins available to their hosts. Understanding properties of these communities, such as structure, diversity and richness, is essential for unravelling the underlying processes that govern the organisation of those systems [1].

Two most popular strategies for studying host-associated and environmental microbiomes are amplicon sequencing and shotgun sequencing. The first approach targets specific genes (often ribosomal RNA genes, such as 16S/18S), which are present in the vast majority of organisms. Sequencing of these genes allows researchers to produce a profile of microbial biodiversity. Shotgun sequencing, in its turn, is a whole metagenome sequencing that provides researchers with a vast amount of information on different genes. This method is crucial to assemble new genomes, however it can be computationally complex [2].

In this study we used both methods to analyse DNA samples from dental calculus of ancient human skulls and compare results with modern microbiome.

Methods:

Samples were taken from Dental Calculus Metagenome study. The skeletons observed in this study had osteological signs of periodontal disease. We were using results of sequencing of V5 16S ribosomal RNA. Dental calculus (SRX351237) was an experimental sample, and the tooth root

(SRX351242) was the proxy for an environmental control. Microbiome analysis was performed by QIIME 2 [10]. We demultiplexed and quality filtered data using the *q2-demux* and denoised using q2-dada2. Then we classified the taxonomy to ASVs using the *q2-feature-classifier* classify-sklearn naive Bayes taxonomy classifier against the Greengenes 13_8 99% OTUs reference sequences.

Dental calculus (sample SRX351237) and bone skull (sample SRX351242) amplicon sequencing data were obtained from the Figshare repository [9]. Preassembled shotgun sequencing was loaded from Dropbox shared directory [8].

To align our metagenomics shotgun sequencing data to a reference database we used MetaPhlAn version 3.0.14. [3] Then we compared our results to the samples from different parts of the human body - the data was taken from the Human Microbiome Project. [4]

Then we used the Pavian web-tool to visualize our metagenomics classification results. To compare our ancient genome sample with modern *Tannerella forsythia* reference genome the sequences were aligned to a reference, the reference file was indexed via BWA v0.7.17 tool (*bwa mem* and *bwa index*). The data was taken from the NCBI database (*Tannerella forsythia* 92A2, complete sequence, NCBI Reference Sequence: NC_016610.1). To compress and sort the sam file, sort and index bam file were used samtools v1.13 package [5]. Then we searched for mutations using Integrated Genomics viewer (IGV). [6] To find out which of the modern strain regions do not overlap with the ancient sample, bedtools v2.30.0 package (*bamtobed* and *intersect* utility) was used. [7]

Results:

We obtained 5788 reads for the bones sample and 5362 for the calculus sequence sample. We filtered them from primer and adapter sequences and clustered. After the filtering we got 5589 and 5185 reads left respectively. Also we got 165 features (with length of 115 nts) by clustering into the amplicon sequence variant. The source of the samples was a human affected by periodontitis, as a result of taxonomic analysis shows. We found no evidence of red complex bacteria presence (*Porphyromonas gingivalis*, *Tannerella forsythia*, *Treponema denticola)* in the bone sample, but this bacterias appeared to be present in the calculus sample (Fig.2.)

Fig.1 Taxonomic range of bone and calculus samples.

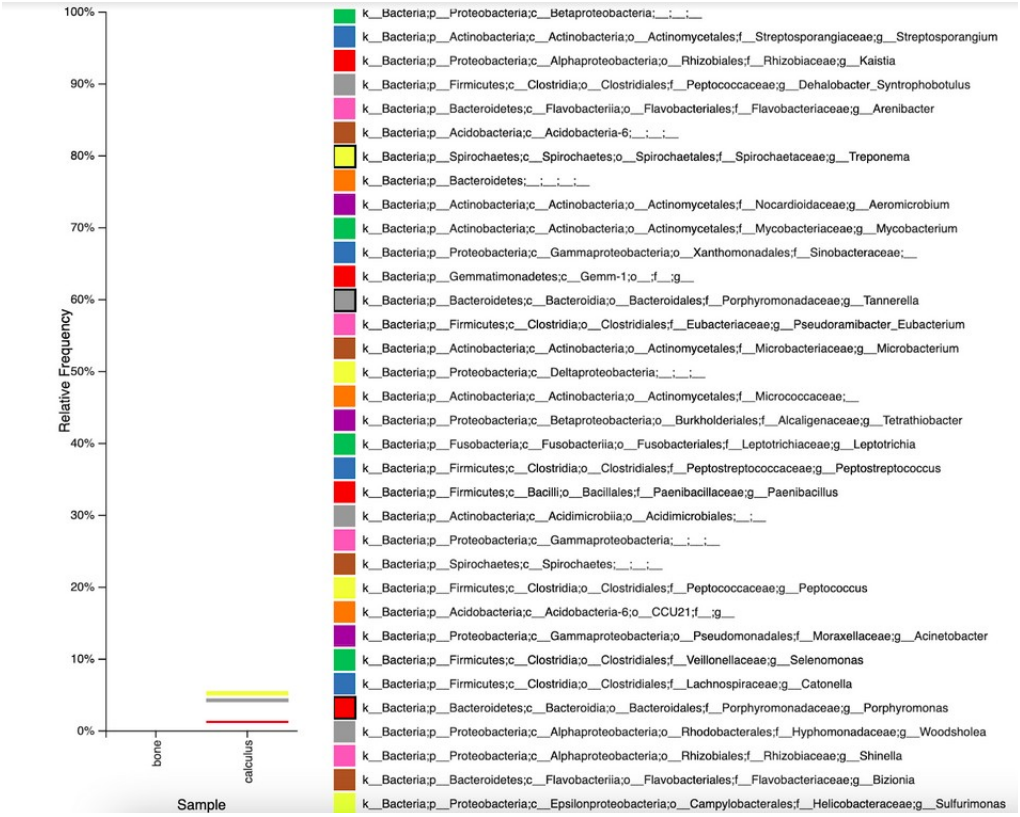

Fig.2. Evidence of presence of red complex bacteria in bone and calculus samples.

For the next step we profiled the whole genome shotgun sequencing data. We investigated the distribution of different taxa in various parts of human body. The microbiota diversity was obvious - for example at the oral surface we found signs of Streptococcus and Lactobacillus (Fig.4.) and in the stool sample the Clostridium and Faecalibacterium were widely spread (Fig.5.)

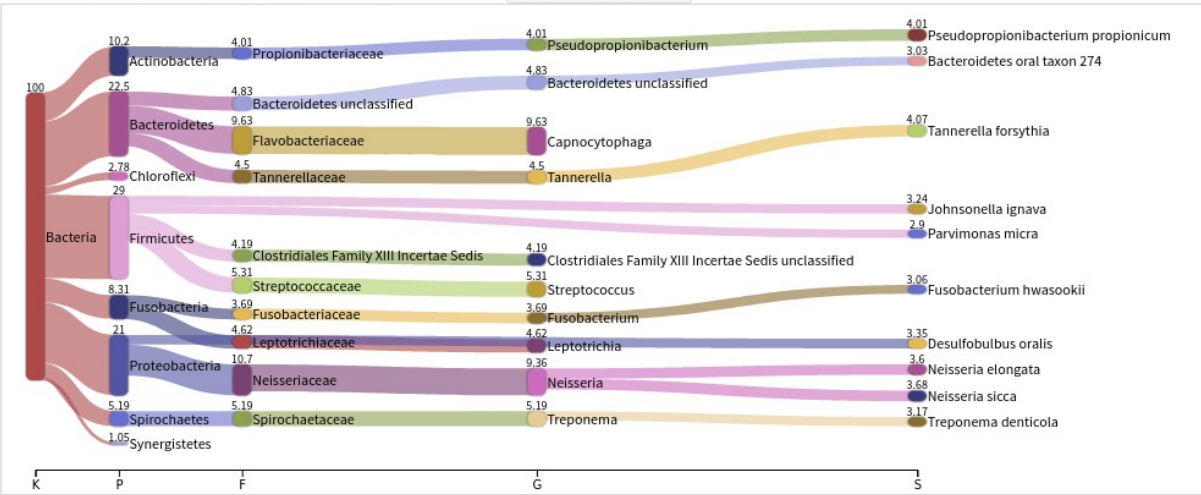

Fig.3. Taxon distribution among the G12 samples.

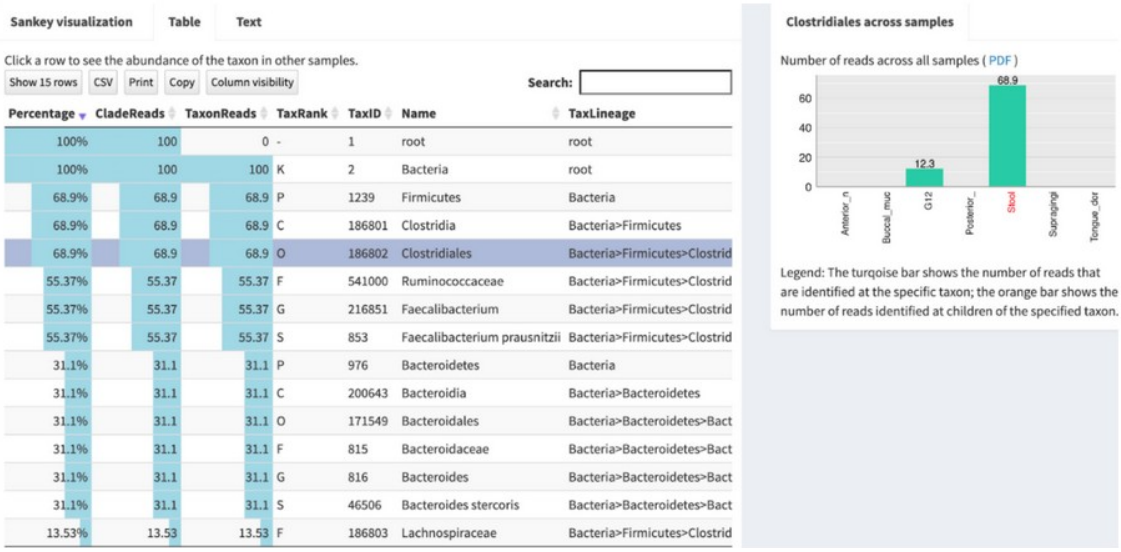

Fig.4. Number of reads for different taxa in oral surface sample.

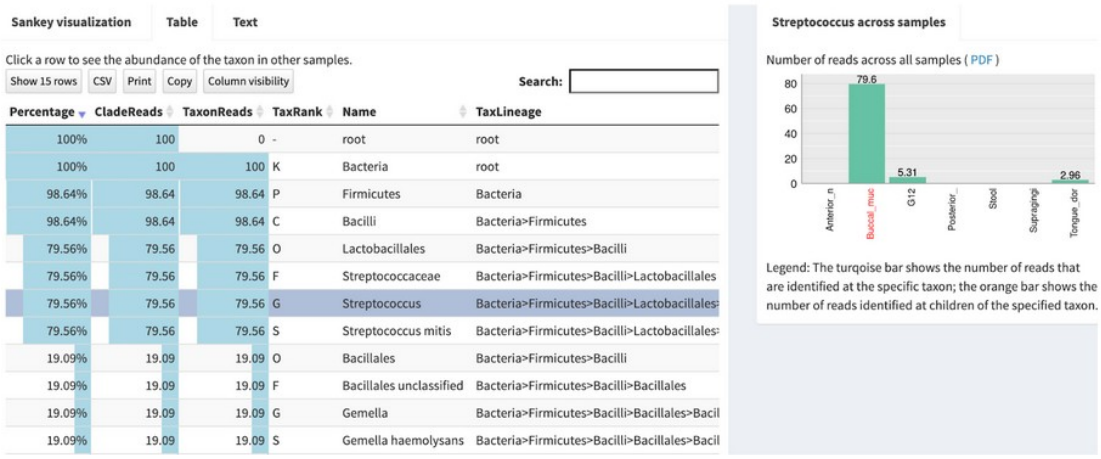

Fig.5. Number of reads for different taxa in stool sample.

Then we compared the ancient genome of the *Tannerella forsythia* with modern strain and investigated how they evolutionally differ from each other. We have found several mobile transposable elements and specific proteins that probably regulate the genes expression.

Discussion:

Three bacterial species, also known as “the red complex”, are usually found together in periodontal pockets nowadays. According to our study our ancestor (G12) also suffered from the same bacteria, and the main ones are *Porphyromonas gingivalis*, *Tannerella forsythia* and *Treponema denticola*. We were able to trace the evolution of one of these bacteria, *T. forsythia*, by comparing the genome of “old” *T. forsythia,* obtained from G12 subject, with the modern one.

As a result of the comparison we obtained some information about how the evolution process of this “red complex” bacterial species went. For the period of more than 1000 years, the bacterium has received many mobile genetic elements (transposons) and special enzymes that catalyses movement of such MGEs to another part of the genome (transposases). In bacterial organisms, the transposition mechanism is important in creating genetic diversity within species and adaptability to changing living conditions [11]. It is also one of the causes of antibiotic resistance, and therefore it is to be expected that *T. forsythia* genome gained new antibiotic resistance mechanisms (eg. tetracycline resistance ribosomal protection protein). *T. forsythia* genome also gained some helix-turn-helix domain-containing proteins. HTH motif is known to be a major structural motif capable of binding DNA and occurs in many proteins that regulate gene expression. It is quite expected that during the

process of evolution new mechanisms to regulate gene expression are required and therefore new proteins with HTH motifs appear.

As we can see, bacteria evolve very quickly, especially because of horizontal gene transfer and viral insertions processes [12]. Such rapid evolution made it possible that people nowadays suffer from the same periodontal disease which is caused by the same bacteria that can be found in 1000y.o. samples, despite the difference in diet and huge time period.

Citations:

1. Thomas T, Gilbert J, Meyer F. Metagenomics - a guide from sampling to data analysis. Microb Inform Exp. 2012;2(1):3. Published 2012 Feb 9. doi:10.1186/2042-5783-2-3
2. Lind, A.L., Pollard, K.S. Accurate and sensitive detection of microbial eukaryotes from whole metagenome shotgun sequencing. Microbiome 9, 58 (2021).
3. Beghini F. et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. Elife. 2021 May 4;10:e65088. doi: 10.7554/eLife.65088.
4. http://hmpdacc.org/HMASM/
5. Li H.*, Handsaker B.*, Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. and 1000 Genome Project Data Processing Subgroup (2009) The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics, 25, 2078-9
6. James T. Robinson, Helga Thorvaldsdóttir, Wendy Winckler, Mitchell Guttman, Eric S. Lander, Gad Getz, Jill P. Mesirov. Integrative Genomics Viewer. Nature Biotechnology 29, 24–26 (2011).
7. 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.
8. https://www.mg-rast.org/mgmain.html?mgpage=download&metagenome=mgm4530391.3
9. https://figshare.com/articles/dataset/_Dead_man_s_teeth_dataset/12152040
10. Bolyen, E., Rideout, J.R., Dillon, M.R. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat*

Biotechnol 37, 852–857 (2019). https://doi.org/10.1038/s41587-019-0209-9

11. Rogers, Kara. "horizontal gene transfer". *Encyclopedia Britannica*, Invalid Date, https://www.britannica.com/science/horizontal-gene-transfer. Accessed 19 April 2022.
12. Irwin, N.A.T., Pittis, A.A., Richards, T.A. et al. Systematic evaluation of horizontal gene transfer between eukaryotes and viruses. Nat Microbiol 7, 327–336 (2022). https://doi.org/10.1038/s41564-021-01026-3