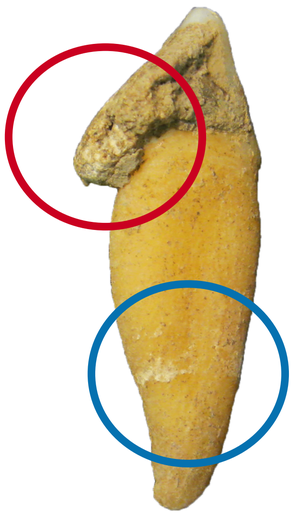
Differential preservation of endogenous human and microbial DNA in dental calculus and dentin

Allison E. Mann[1,2], Susanna Sabin[1], Kirsten Ziesemer[3], Åshild J. Vågene[1], Krithivasan Sankaranarayanan[2], Andrew Ozga[2], James Fellows Yates[1], Christopher Read[4], Alexander Herbig[1], Johannes Krause[1], Menno Hoogland[3], Corinne Hofman[3], Kirsten Bos[1], Christina Warinner[1,2]

[1]Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena, Germany
[2]Department of Anthropology, University of Oklahoma, Norman, OK, USA
[3]Faculty of Archaeology, Leiden University, Leiden, the Netherlands
[4]Institute of Technology Sligo, Sligo, Ireland

Introduction:

Ancient DNA (aDNA) provides unique insights into past human behavior, health, and evolution. Skeletal tissues (bone and dentin) and microbiome remains (dental calculus and paleofeces) can be rich sources of ancient biomolecules; however, inconsistent preservation and variable environmental contamination pose major challenges in recovering authentic aDNA. Recent studies suggest that dental calculus may provide a better preservation environment for aDNA than other skeletal tissues[1], but this hypothesis has not been systematically tested. In this study, paired dentin and calculus samples from sites representing an extensive geographical range and broad temporal depth are investigated to better understand aDNA preservation between calculus and dentin, and within calculus itself.



Does calculus provide a special environment for preservation?

Analytical Metrics:

- DNA yield
- Endogenous human content (%)
- Microbial community structure
- DNA fragmentation patterns

Figure 1: Geographic locations and corresponding age ranges of samples

Methods:

Sample sites: Middenbeemster, the Netherlands (n=2); Camino del Molino, Spain (n=2); Samdzong, Nepal (n=2); Hovsgol, Mongolia (n=2); Anse à la Gourde, Guadeloupe (n=2); Norris Farms, IL, USA (n=2); and Kilteasheen, Ireland (n=36) (*Figure 1*). All samples were Illumina shotgun sequenced using a 2x100 paired-end chemistry except for the Kilteasheen samples, the data of which are single-end, 75 bp. As such, the Kilteasheen samples were excluded from fragment length analyses. Data were quality filtered and mapped to the human genome (hg19) using EAGER[2] and taxonomically binned with MALT[3] against the full NCBI Nt database. Potential source contribution analysis was performed with Sourcetracker[6].

Results & Discussion:

Consistent with previous studies, calculus on average has higher overall DNA yield as compared to dentin (*Figure 2a*), though percent human endogenous content is generally lower (*Figure 2b*)[1,9]. Human endogenous content in dentin is variable while calculus is comparatively consistent. In addition, human reads recovered from calculus are distinctively fragmented as compared to dentin, independent of overall sample median fragment length (*Figure 2c*). It is possible that human DNA is vulnerable to hydrolytic or other damaging processes to the sugar-phosphate backbone during its incorporation into the calculus matrix. Immune cells that produce extracellular chromatin traps are highly active in the oral cavity to combat the formation of plaque[1,4]; we hypothesize that the incorporation of naked DNA from these traps, as well as bacterial surface nucleases[5] explain these fragmentation patterns.

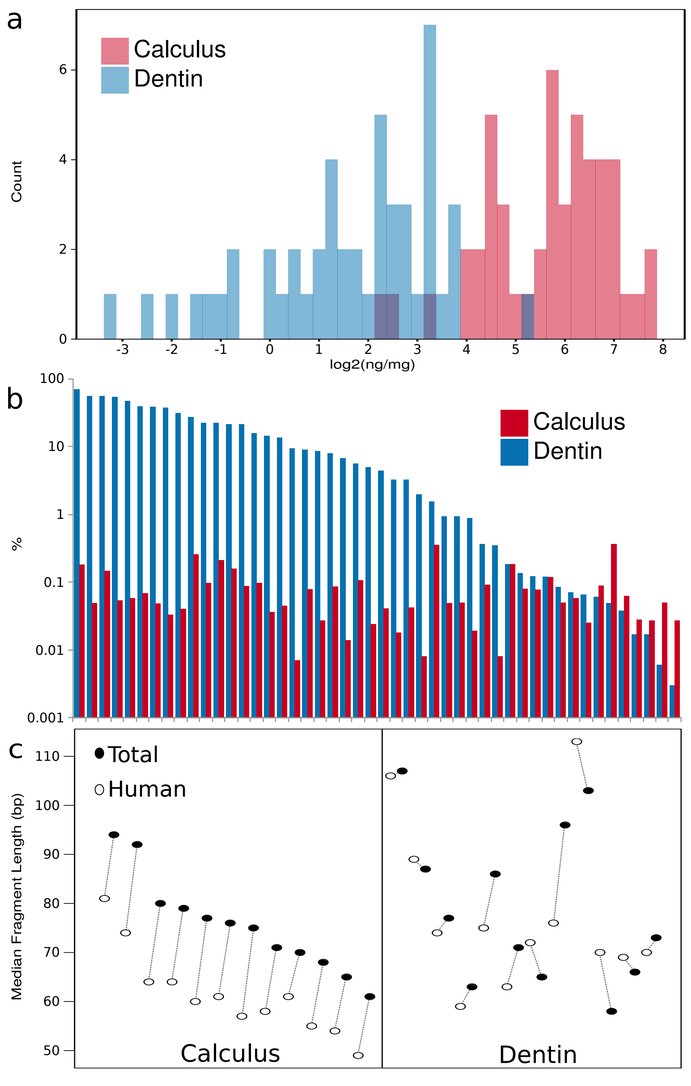


Figure 2 (right): (a) total DNA yield of dentin (n=49) versus calculus (n=49) post extraction; (b) human endogenous content for all paired calculus and dentin samples (n=98); (c) overall median fragment length versus human median fragment length for a subset of samples (calculus n=12, dentin n=12)

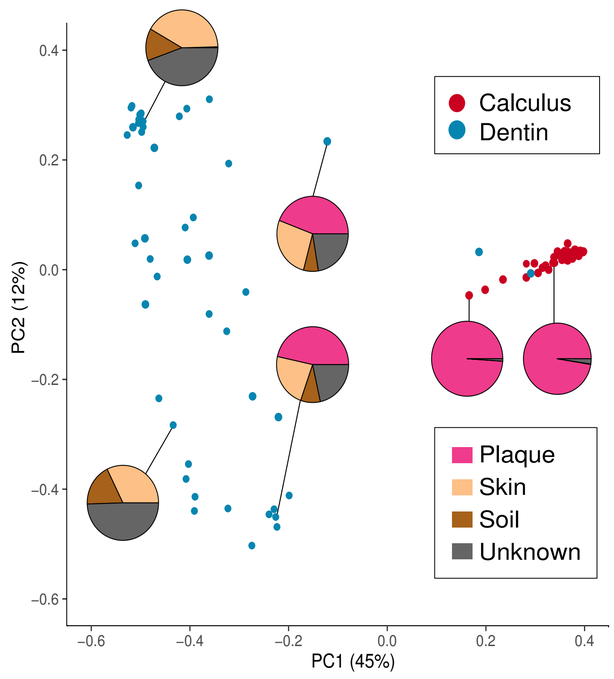


Figure 3: PCoA (Bray-Curtis) of all bacterial and archaeal species hits to the NCBI Nt database in dentin and calculus. Pie charts indicate proportion of potential source contribution of select samples[6].

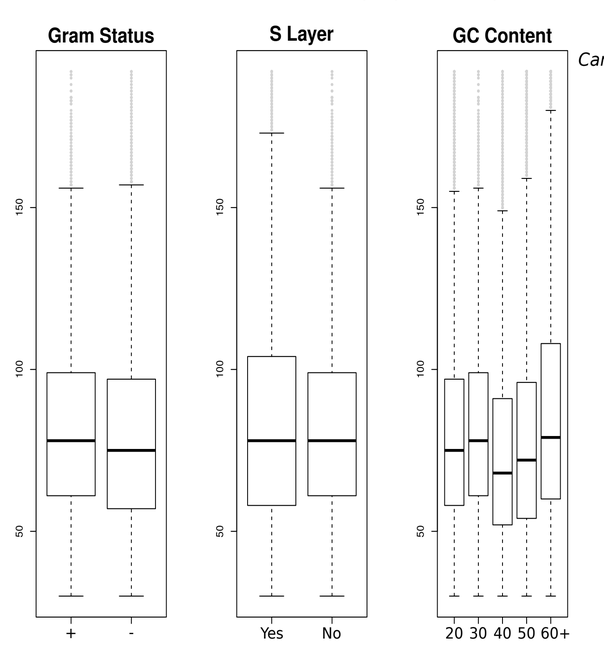


Figure 4: Fragment length distribution of the top 55 bacterial species in calculus grouped into three metadata categories: gram status, presence of a surface layer, and overall genomic GC content.

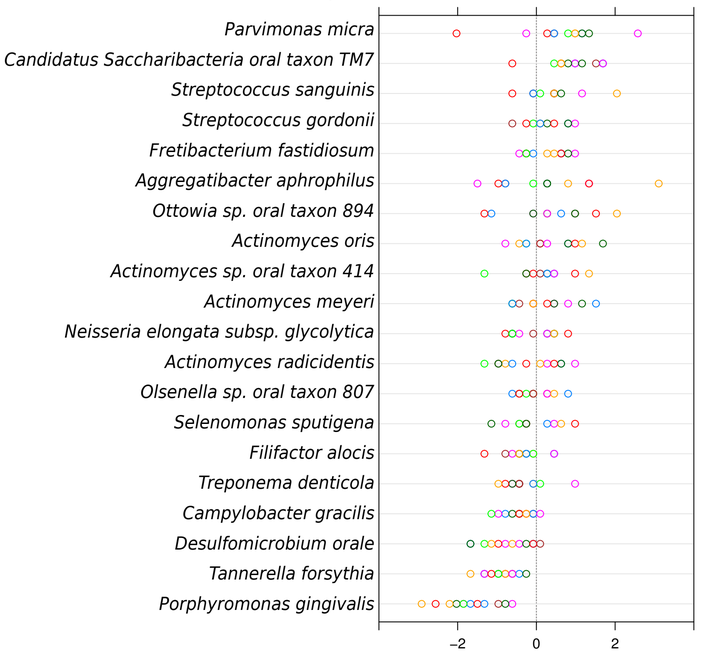


Figure 5: Deviation of median fragment length from overall sample median fragment length of 20 top oral bacteria. Colored rings represent different samples.

Next, we investigated the impact of cellular and genomic structure on fragment length patterns, which has been argued to skew bacterial community profiles in ancient dental calculus[7]. While Gram status, the presence of a surface layer (S layer), and overall genomic GC content do not appear to impact preservation of the top 55 species within calculus (*Figure 4*), four individual species, including two members of the "Red Complex"[8], have consistently shorter median fragment lengths than expected, independent of overall sample median fragment size (*Figure 5*). Because fragment size impacts reference genome mapping efficiency, taxonomic-specific deviations have the potential to skew microbial community reconstruction.

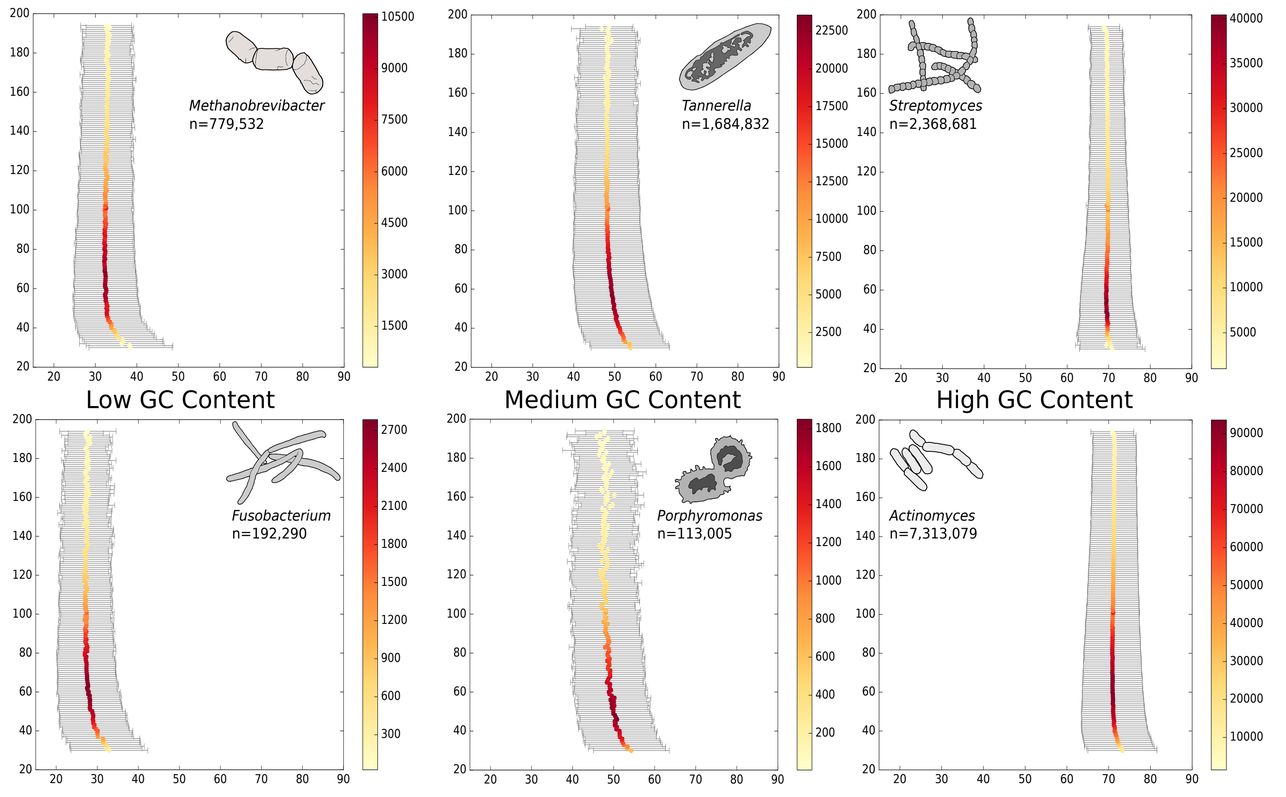


Figure 6: Relationship of GC content to fragment length in five common oral genera and one high GC content soil genus (Streptomyces). X axis = mean GC content of reads mapped to a particular genus, categorized as low expected GC content (<40%), medium expected GC content (40%-59%), and high expected GC content (>60%). Y axis = fragment length. Heat map colors indicate where the bulk of reads fall within the total length distribution.

Finally, we detect a loss of AT rich reads at shorter fragment lengths in both calculus and dentin, likely related to the relative weakness of hydrogen bonds of AT rich sequences (*Figure 6*). This effect is exacerbated in lower overall genomic GC taxa but minimal in high GC rich genomes which include many potential soil contaminates. As the ultimate goal of microbiome studies is to reliably compare modern and ancient samples, understanding potential taxonomic biases is of paramount importance for the field of ancient microbiomes.

References Cited:

1. Warinner et al. 2014. Nature Genetics, 46: 336-344
2. Peltzer et al. 2016. Genome Biology 17:60
3. Herbig et al. BioRxiv, April 27, 2016
4. Ryder. 2000. Periodontology 53: 124-137
5. Palmer et al. 2012. J Periodont Res 47:439-445
6. Knights et al. 2011. Nature Methods 8: 761-763
7. Adler et al. 2013. Nature Genetics 45:450-455
8. Socransky et al. 1998. J Clin Periodontol 25: 134-144
9. Ozga et al. 2016. American Journal of Physical Anthropology 160(2): 220-228

Acknowledgements:

This research has received funding from the ERC under the European Union's Seventh Framework Programme (FP7/2007-2013) ERC grant agreement n° 319209 "NEXUS 1492" under the direction of Prof. Dr. C.L. Hofman as well as an ERC starting grant under the direction of Johannes Krause, a Social Sciences and Humanities Research Council of Canada postdoctoral fellowship under the direction of Kirsten I. Bos, and the Max Planck Society. In addition, the authors would like to thank Mark Aldenderfer, Bruno Frohlich, and Domingo Salazar Garcia.

Contact: mann@shh.mpg.de