nature research

Peer Review Information

Journal: Nature Microbiology
Manuscript Title: Ancient dental calculus reveals oral microbiome shifts associated with lifestyle and disease in Great Britain
Corresponding author name(s): Laura S. Weyrich

Editorial Notes:

Transferred manuscripts This manuscript has been previously reviewed at another journal that is not operating a transparent peer review scheme. This document only contains reviewer comments, rebuttal and decision letters for versions considered at Nature Microbiology.

Reviewer Comments & Decisions:

Decision Letter, initial version:

Subject: Decision on Nature Microbiology manuscript NMICROBIOL-22051270-T
Message: 3rd August 2022

*Please ensure you delete the link to your author homepage in this e-mail if you wish to forward it to your co-authors.

Dear Laura,

Thank you for your patience while your manuscript "Ancient dental calculus reveals novel oral microbiome diversity in Great Britain" was under peer-review at Nature Microbiology. It has now been seen by 4 referees, whose expertise and comments you will find at the end of this email. Although they find your work of some potential interest, they have raised a number of concerns that will need to be addressed before we can consider publication of the work in Nature Microbiology.

In particular, all referees ask for additional methods details including sample processing, controls for contamination, and various bioinformatics analyses. Referee #2 asks that you use alternative methods for nucleotide-to-protein alignment and suggests using published oral microbiome datasets from the UK to support your conclusions. Referee #3 asks you to consider other factors that could be driving the observed associations and specifically how the black death impacted oral microbiotas, as well as several questions regarding

contamination and approaches taken to address this. This referee also asks that you are more specific when describing the results and conclusions and to specify in what direction the observed microbiota changes happened. Should further experimental data allow you to address these criticisms, we would be happy to look at a revised manuscript.

We are committed to providing a fair and constructive peer-review process. Please do not hesitate to contact us if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

We strongly support public availability of data. Please place the data used in your paper into a public data repository, if one exists, or alternatively, present the data as Source Data or Supplementary Information. If data can only be shared on request, please explain why in your Data Availability Statement, and also in the correspondence with your editor. For some data types, deposition in a public repository is mandatory - more information on our data deposition policies and available repositories can be found at https://www.nature.com/nature-research/editorial-policies/reporting-standards#availability-of-data.

Please include a data availability statement as a separate section after Methods but before references, under the heading "Data Availability". This section should inform readers about the availability of the data used to support the conclusions of your study. This information includes accession codes to public repositories (data banks for protein, DNA or RNA sequences, microarray, proteomics data etc…), references to source data published alongside the paper, unique identifiers such as URLs to data repository entries, or data set DOIs, and any other statement about data availability. At a minimum, you should include the following statement: "The data that support the findings of this study are available from the corresponding author upon request", mentioning any restrictions on availability. If DOIs are provided, we also strongly encourage including these in the Reference list (authors, title, publisher (repository name), identifier, year). For more guidance on how to write this section please see:
http://www.nature.com/authors/policies/data/data-availability-statements-data-citations.pdf

If revising your manuscript:

* Include a "Response to referees" document detailing, point-by-point, how you addressed each referee comment. If no action was taken to address a point, you must provide a compelling argument. This response will be sent back to the referees along with the revised manuscript.

* If you have not done so already we suggest that you begin to revise your manuscript so that it conforms to our Article format instructions at http://www.nature.com/nmicrobiol/info/final-submission. Refer also to any guidelines provided in this letter.

* Include a revised version of any required reporting checklist. It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review. A revised checklist is essential for re-review of the paper.

When submitting the revised version of your manuscript, please pay close attention to our href="https://www.nature.com/nature-portfolio/editorial-policies/image-integrity">Digital Image Integrity Guidelines. and to the following points below:

-- that unprocessed scans are clearly labelled and match the gels and western blots presented in figures.
-- that control panels for gels and western blots are appropriately described as loading on sample processing controls
-- all images in the paper are checked for duplication of panels and for splicing of gel lanes.

Finally, please ensure that you retain unprocessed data and metadata files after publication, ideally archiving data in perpetuity, as these may be requested during the peer review and production process or after publication if any issues arise.

Please use the link below to submit a revised paper: [REDACTED]

Note: This url links to your confidential homepage and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this e-mail to co-authors, please delete this link to your homepage first.

Nature Microbiology is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. This applies to primary research papers only. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit please visit www.springernature.com/orcid.

If you wish to submit a suitably revised manuscript we would hope to receive it within 6 months. If you cannot send it within this time, please let us know. We will be happy to consider your revision, even if a similar study has been accepted for publication at Nature Microbiology or published elsewhere (up to a maximum of 6 months).

In the meantime we hope that you find our referees' comments helpful.

Yours sincerely,

****
Reviewer Expertise:

Referee #1: metagenomics, ancient microbiomes
Referee #2: ancient microbiomes, anthropology

Referee #3: low biomass microbiomes
Referee #4: ancient DNA, anthropology

Reviewer Comments:

Reviewer #1 (Remarks to the Author):

The authors here analyze an expansive and potentially quite valuable new collection of human dental calculus from across Britain spanning ancient and modern times. They identify two clusters of microbiome composition, including one that has not been observed in modern humans. They identify potential correlations to diet that might underlie the clusters

There is not enough information to evaluate the sample authenticity and contamination levels of the samples studied here. For a paper describing a rich, new source of ancient human microbial genomes authenticity analysis should be central. But here it is described in only a half sentence in the main text and relegated to a rather cumbersome and unorganized supplementary text and just a few supplementary tables.

There is not nearly enough detail in the Methods on how the clusters were constructed, which is a pivotal analysis in the paper. It's not sufficiently clear how membership is defined in the 3 clusters in Fig3A. What are the co-occurance rate cut-offs used, how were they chosen, and how does the number of clusters and sample assignment change as those cut-offs are modified?

An important indication that the clusters are not robust is their lack of new biological insight. It is telling that both clusters appear to be randomnly distributed in FigS10 and there is no time-trend supporting the hypothesis that the Methanobrevibacter cluster was a pre-Industrialized cluster. Some earlier time periods do not show the Methanobrevibacter cluster. Also, the cluster is the dominant one in the most recent time period before modern, which includes Industrialized people/samples.

A version of Fig3A needs to be shown that labels samples by their origin (time period, geography, etc) rather than just assignment to the co-occurance clusters. It's not obvious that clusters are capturing biologically meaningful information.

In Fig4 I'm not sure exactly what I'm meant to be observing. I don't really see much of a difference between the Streptococcus, Methanobrevibacter, or Modern groups in this plot. If there are features that are significantly different between groups, can they be labelled with an effect size and P-val, and name of the relavent feature? The corresponding LefSE plots in the supplement are not sufficient to demonstrate a statistically significant difference between groups. LefSE is designed to be a data exploration tool. Appropropriate two group or ANOVA statistical tests must be applied to each feature. The authors can use a different BioBakery tool such as MaAsLin for this purpose or one of the many other packages for multivariate analysis.

There is a statement (line 119) that oral microbiome diversity was likely greater in pre-Industrialized oral microbiomes than modern Industrialzed today. Diversity in what sense? I don't see direct metrics of diversity. If the diversity is significantly different, does this analysis adequately account for the unavoidable contamination bias in ancient samples that could inflate diversity metrics?

Please do not display 3-dimensional plots as static figures, paper printouts only have 2 dimensions. Represent 2 axes at a time and use video files for 3D rendering if really necessary.

The manunscript has many references to timeperiodes (Civil War, Medieval/post-Medieval in Table 1 for example and also Fig1) that may not be easily recognizable to non-British people. It is better to use carbon-dating verified date ranges or numeric time periods.

Reviewer #2 (Remarks to the Author):

In this manuscript, Gancz et al. study the oral microbiome diversity of British populations through shotgun metagenomic sequencing of ancient dental calculus sample spanning the Bronze Age to the mid-18th century.

Overall, the manuscript is well-written. The laboratory work is exhaustive and has resulted in an excellent dataset of >200 ancient oral metagenomes from Britain with associated meta-data. The data analyses are well-thought out, but there a few issues to consider (see below comments). I appreciated the public outreach that has been conducted as part of this study.

Major Comments:

One of my major concerns is the use of a MALT-x (nucleotide-to-protein alignment) approach for taxonomic profiling in this study. This approach is not used in most ancient microbiome studies, and in fact, previous research from this group (Eisenhofer and Weyrich, 2019) has shown that nucleotide-to-nucleotide alignment approaches, such as MALT-n, outperform MALT-x for profiling ancient microbiome datasets. Given this, it is unclear why the authors use a MALT-x approach in this study. If computing power necessary for running MALT-n is a limitation, perhaps an alternative approach would be to use Kraken and/or a customized database.

The analysis of these new ancient British dental calculus samples provides interesting findings, such as oral communities dominated by either Methanobrevibacter or Streptococcus. But these findings can be better contextualized using previously published ancient dental calculus samples from the UK. However, the authors include samples from Neanderthals and medieval Germany as comparative samples, which do not seem relevant to this analysis. Instead, it would be good to include the data from Velsko et al. 2019 (40 ancient dental calculus samples from a single cemetery in Oxford). Also, is there any overlap between the Jewbury samples included in this study and the ones from this group's previous study (Farrer et al. 2021)? The latter seems to have considerably more samples and it is unclear if/why these were not included.

In general, the SourceTracker analysis could have been better described. In Table S20, it is unclear where the modern dental calculus and soil samples used as sources came from – the table seems to suggest they were generated as part of this study? If so, what is their context and how were these data generated? If these were from previously published datasets, they should be cited. Second, in addition to modern dental calculus and EBCs, the number of other source samples is very low. The authors could include other oral sources such as modern dental plaque, skin, and more soil samples in their analysis to

make it more rigorous. That may help reduce the high % of Unknown in Figure S2B. Third, I do not understand how Figure S2A was generated – what Ancient Dental Calculus samples were used as a source here? Regardless, I do not think it is appropriate to use other ancient dental calculus samples as sources in this type of analysis, since one can never be certain that their taxonomic profile is free from non-endogenous taxa.

The presence of a Methanobrevibacter-dominated oral community was found to be associated with size of calculus deposit. Given modern dental practices and frequency of calculus removal, it is likely that modern dental calculus deposits do not grow sufficiently large to provide the anaerobic environments that support the growth of Methanobrevibacter. This point is mentioned rather off-handedly in the Main Text, but in my opinion, it seems to be the most likely explanation for the "disappearance" of this Methanobrevibacter-dominated ecology from modern dental calculus. The dearth of metagenomic sequencing data for modern dental calculus samples from Britain, and in general, from different dietary and geographical contexts, limits understanding whether Methanobrevibacter-dominated communities continue to exist even today. This point receives some attention in the Supplementary Material, but I think it needs to be discussed in the Main Text.

Minor Comments:

Supplementary Methods: In addition to removing reads greater than 300 bp, was a minimum length and quality threshold (e.g. Q30 and length 30?) applied to reads after merging?

In my opinion, the term "industrialized" should not be capitalized when referring to industrialized countries, industrialized/pre-industrialized/non-industrialized microbiomes, or industrialization.

Line 38-39 and 85-86: In Lines 38-39, the authors state that they are the first "to directly reconstruct the microbial history of an Industrialized population for the first time". In Lines 85-86, they state they "… directly describe the history of a pre-Industrialized microbiome for the first time." These sentences are overstating the findings of this study. Both sets of lines should be rephrased to say something along the lines of "this is the first study to directly reconstruct the oral microbial history of a population spanning the transition to industrialization".

Lines 48: Recommend changing "unappreciated" to something else, due to repetitive use of appreciated in this and the previous line.

Line 58: Obregon-Tito et al. 2015 could be added as a citation as well.

Line 91: It would be good to discuss findings from Fagernäs et a. 2021 about oral geography affecting oral microbiome composition, since that paper also studies this using ancient dental calculus samples.

Lines 116-119: It is unclear how the authors "confirmed that species within these co-occurring genera can cohabitate with Methanobrevibacter oral species, as all are anaerobic and can produce metabolic by-products that support methanogenesis". Was this done using just literature review, or was any kind of actual experimentation involved?

Line 138: Correct typo in "populations".

Lines 142 and elsewhere – I recommend referring to the communities consistently as "Streptococcus-dominated" and "Methanobrevibacter-dominated" rather than just with the genera names.

Table S13 is difficult to interpret – specifically with regards to Line 143.

Line 163: Specify which modern, industrialized population was studied.

Line 282: Correct typo in "AdapterRemoval".

Figure 4: I recommend switching the color scheme, since readers will instinctively associate the color red with higher abundance.

Figure S5: Some X-axis labels are cut off.

Supplementary Information should be proof-read for typos.

Reviewer #3 (Remarks to the Author):

It is of great interest to investigate the evolution of the oral microbiome. The question that poses itself is whether we are dealing with an improvement or a devolution, like is seen in the gut which is associated with an increase in the incidence of various diseases of affluence, or whether it is matter of perspective. Whilst oral health has improved in modern times the Methanobrevibacter-associated community was not found to be associated with periodontal disease (whilst bugs associated with Streptococcus & co. were). Perhaps even more specifically, is the increase in oral hygiene linked with the decrease of gut health? Or is there perhaps a case to be made to adjust modern oral hygiene so that the Methanobrevibacter-associated community still has a chance? In Western populations Methanobrevibacter (associated with low BMI) is slowly also disappearing from faeces (see paper on Irish travellers, for example) being replaced by a more Bacteroides dominated Microbiome (https://www.nature.com/articles/s41591-020-0963-8/figures/8). Is oral seeding of the gut no longer possible or much more difficult due to oral extinction of Methanobrevibacter (& friends) as a result of supposedly improved oral hygiene? Please feel free to incorporate or discuss aspects of the above in the discussion.

Major comments:
1a: The authors are of course clearly aware of the issue of contamination and they clearly did something about it (Table S4) yet I can't really ascertain how they checked for laboratory and environmental contaminant DNA. The main thing I found was the abundant use of negative controls. One step that was perhaps even too strict was that species identified within negative controls samples were removed from all samples. Sometimes genuine signals end up in negative controls due to a variety of reasons (barcode bleedthrough, machine contamination, "splashome", etc). Was it verified that these signals found in the negative controls (and subsequently removed from real samples) were indeed not of ancient origin? Did you compare/assess DNA damage levels for each of these signals? Which leads me to my next question.

1b: It is unclear how the authors used DNA damage levels of reads as a tool to ascertain which reads are genuine (damaged due to old age) and which ones are contaminants. I do see mention of: "Reads longer than 300 bp were discarded, as they likely represent modern contamination." and that "Shotgun metagenomic libraries were generated without enzymatic damage repair" but I do not see mention of making use of tools such as MapDamage2. For example, in the articles of Wiboho, Granehäll and Rampelli (references 16, 24 and PMC7864912) they for example use MapDamage2. I assume that DNA damage profiles were assessed and used for screening reads, and if so, this needs to be written down more clearly. If not, then this simply needs to be done.

1c: In short, I need a lot more information/detail on how contaminants were separated from genuine signals; not just that negative controls were used. Furthermore, there are many more tools than just the use of negative controls or DNA damage (though both are very important) to identify contaminants. Commonly, contaminants from the same source typically correlate strongly with one another allowing one to use contaminants to identify other contaminants. Also for example, you could perform a correlation analysis of all signals with axis 1 and axis of Figure S1. All signals strongly negatively correlated with both axes are contaminants. Furthermore, were particular signals perhaps part of an artificial batch effect (associated with the use of a particular lot-number)? It is hence good to show per signal that they are 1) ancient/damaged (yes/no), 2 present in negative controls (yes/no), 3) associated with other contaminants (yes/no) 4) part of a batch effect (yes/no), something else you can think of (strain-tracker(?)) (yes/no). A multi-layered approach is recommended and the separation of genuine/contaminant signals should not be a mere side-note.

2: Possibly related to contamination: large differences between different cemeteries were found ("We found that 64.83% of the compositional variation explained by time could equally well be explained by cemetery"). How did DNA-damage profiles differ between cemeteries? How do the DNA damage profiles of the different bacteria that are mostly representative of this "64.83% compositional variation" compare with the DNA-damage profiles of species of which one is 100% certain they are genuine? Do older cemeteries/remains btw in general indeed show more DNA damage?

3: Abstract: "While many factors likely contribute to this observation, microbial functional analysis suggests that common oral microbes today may reflect past differences in diet, specifically dairy and carbohydrate consumption. In London, the less common microbiome is linked with skeletal pathologies, and its disappearance is consistent with temporal shifts, including the arrival of the Second Plague Pandemic. These findings suggest that pre- Industrial microbiomes may have been more diverse than previously appreciated and linked to health in previously unappreciated ways, shifting our understanding of chronic, noncommunicable disease origins in Industrial populations.

The dietary aspects indeed seem credible (have changed significantly in modern times and was likely associated with social status in ancient samples) but the others factors mentioned don't yet seem that convincing. When looking at Figure S10 it is quite clear that the Methanobrevibacter-rich-composition was still live and kicking between 1600-1900. The main elephant in the room, is improved dental hygiene. Specifically, it has been found (a main finding that should be mentioned in the abstract) that the Methanobrevibacter cluster was associated with calculus fragment size. This if anything most likely explains why the Methanobrevibacter cluster has gone the way of the Dodo. The Methanobrevibacter cluster & this relevant association are not even mentioned in the

abstract. I think the abstract can be rewritten/refocused a bit. The summary does not really mirror the last paragraph of the main text.

4: In the text it is said that ”the Methanobrevibacter-associated communities were associated with the presence of several skeletal markers of systemic disease, including non-specific periostitis, joint porosity, osteophytic lipping, and overall scores for joint pathologies”. Is this association perhaps not indirect? Were these markers of systemic disease perhaps associated with age? Is age perhaps associated with larger (“more mature”) dental calculus size? Alternatively, is this perhaps due to the possible association with lower economic class? This is alluded to somewhat in the Supplemental section but that this association is not unlikely indirect should probably be mentioned in the main text more clearly.

5a: Lines. 225-239. A lot of lines and complicated analyses and some numbers are given in regards to the possible effect of the black death on the oral microbiome but I do not see any sort of direction being given. What kind of (small but significant?) shift did this (temporarily) induce? Did food become more affordable as there were less mouths to feed causing a shift to the supposedly “higher status” Streptococccus composition? Or because people took oral hygiene more seriously for a while leading to more Streptococcus? Or did Methanobrevibacter (or Prevotella) become more dominant?

5b: Similarly, it is stated that “Our results suggest that the composition of oral microbiota in ancient London changed over the 800-year time- period from 1100 to 1900.” yet I’m still in the dark in regards to the direction of this change or the reasons for this supposed change. It is also stated that “Our study reveals the existence of a now rare or even extinct oral microbial ecosystem that remained relatively consistent in British populations over at least 2,200 years. This oral microbial community persisted through major bio-cultural transitions and historically significant socio-political events, only to diminish in recent history – a phenomenon associated with the rise of Industrialized diseases.” How do you rhyme these two statements? The latter one seems logical yet I’m not convinced about the former one.

6: Figures 2 and 3: Figure 2A is supposedly of all teeth and I reckon Figure 3 is also of all teeth. Why then are these plots not identical? Figure 3 is stated to be of the genus level while I’m unsure about Figure 2. Why not use the coordinates of Figure 3 for all selections for Figure 2B, 2C (and 2A?)? I think this would be more informative in regards to ascertaining which species are associated with the differences shown in Figure 2. Would it perhaps be an idea to first show Figure 3 and then Figure 2 (so switch them, and the text, around)? One of the most interesting findings of biological relevance is the association between calculus fragment size and Methanobrevibacter. You can see this somewhat unclearly currently in figure 2B. This would be clearer if the coordinates of figure 3 are instead used.

7: Supplemental Tables: These contain a lot of stuff but not simple tables showing the abundance of each bacterial signal pre- and post-contamination removal, both on the species and genus level in each individual sample (and in each of the negative controls; also show the order in which samples were processed for identifying possible batch effects). People could of course spend a lot of time processing the reads themselves but a simple summary like this would be valuable, especially for a reviewer (with limited time) who wants to do some checks on the reagent contamination removal process. At least Figure S3 gives a lot of per sample information (visually) but a simple excel file. (with a

few tabs) would do a better job for purposes of being able to check the numbers.

8: Project Outreach. Why is this in the Methods section? It has more in common with something like an acknowledgement section if anything. Its position in the Supplementary information section is just fine on the other hand (and currently kind of double with the main text). There is enough I'd like to see in the Methods section which I'm currently not yet seeing (contamination issue).

9: Table 1: I mainly see a lot of numbers of stuff linked to oral microbiome composition without even telling me the direction when a significant different is found. Might be more of interest if an indication of such a direction is added (like superscript "M" or "S", association with Methanobrevibacter- or Streptococcus-rich composition. Also, I cannot understand the Table without reading the main text; the legend might need some work. What does "dominant" furthermore stand for?

Minor comments:
M1: It might perhaps be worth it to draw a parallel between the two distinct oral microbial communities found and the Prevotella vs the Bacteroides enterotype in the gut. The Prevotella community has largely disappeared or is disappearing (depending on which country you're looking at). Have you perhaps tried to do some kind of enterotyping approach to see what rolls out of that?
M2: Figures 3B & 3C: A co-occurrence network can also be visualised using a heatmap using hierarchically clustered spearman rho coefficients. Core members of the 2 clusters should be easily visualized with such an intuitive and simplistic approach. A heatmap might even visualize sub-clusters more clearly (maybe shed some light on Prevotella-rich compositions shown in the middle of 3A?).
M3: Figure 4: Table 15 is stated in the legend. This should be Table S15.
M4: Figure S4A: Where is the Streptococcus arrow?
M5: Page 13 Supplementary information. Streptococci should be Streptococcus (italics) or streptococci (not in italics).
M6: Line 67:.. sites has not been full examined …
-fully

Reviewer #4 (Remarks to the Author):

The manuscript by Gancz et al presents the molecular analysis of 201 calculus samples from Britain, representing a time period from 2200BCE to the 19th century. Through this analysis the authors aim to investigate the origin of the Industrial oral microbiome through correlation with time, health as determined by detectable skeletal pathology, and broad cultural classification.

There seems to be a lack of accord between the research focus established in the introduction and the study the authors present. The introduction is crafted to explore the origin of the modern Industrial, or Western, gut. With a dataset that derives only from calculus and hence oral flora, the lengthy discussion of modern guts seems misplaced. I would recommend a focus in the introduction that is bears closer relationship to the actual study presented, namely exploration of oral microflora.

The manuscript very rapidly moves from a lengthy introduction to a discussion of results. I suggest the authors better prepare their audience for digestion of their analyses by describing their sample dataset in a few sentences. Also, Figure 1A should ideally have the location circles color-coded to reflect their chronological age.

The novelty in this study comes from the restriction of their analysis to the geographical confines of Britain (namely England and Scotland). Their central conclusion is that two broad microbial communities are detectable in their datasets, and a shift in the relative proportions of each of these communities (at least in the city of London) is temporally correlated with the mid-14th century. The authors in turn speculate that the mortality event of the Black Death and unspecified cultural, genetic, or population health changes related to this event may have provided the substrate upon which shifts in oral microbial content occurred. While this is supported through statistical analysis, I suggest the authors display this trend visually in some way. It would be useful to see a graph that plots the population proportion of the Methanobrevibacter component over time to observe the purported effect of the mid-14th century.

What seems missing is a comparison of the dominant post-14th century Streptoccocus oral community with modern oral community structure. That would be needed to justify their research focus on origins of Industrial microbiomes outlined in their introduction. It would also be needed to justify their final statement in the abstract that the analysis they present changes our perspectives on the origins of non-communicable diseases, which are assumed to be linked to Western gut microflora (and perhaps indirectly to modern oral communities).

Could the temporal correlation be influenced by the fact that the dominant time period of the collection is from the Medieval period? Figure 1B should also provide the time ranges in years the authors used to define their time periods.

Their statement in lines 97 and 98 that tooth location in the dental arcade influences oral microbial structure is an important one, and their conclusion that this phenomenon could have compromised interpretations made in previous works is indeed provocative. For such a strong statement, however, quantitative data should be presented and thoroughly discussed in the main manuscript.

The authors make quite a huge jump in their correlation of oral microbial content and systemic disease. It's important to recognize that not all systemic diseases contribute to detectable morphological changes in the skeleton. Also, not all modern chronic and non-communicable diseases, to which they refer in their introduction as being related to modern microbiomes, will contribute to physical changes in the skeleton. In short, their statement here (lines 213 and 214) is too strong and carries too many caveats.

I'm somewhat concerned about their observation that presence of the Methanobrevibacter cluster correlates with size of the calculus deposit used in the laboratory processing. Could this be related to a complexity issue? Is it clear that small calculus samples harvested from larger calculus deposits would still yield a dominant Methanobrevibacter signal? Did

the authors normalize the amount of calcified deposit that went into the laboratory processing? That is not stated in the methods.

Minor comments:

Line 33 – grammatical error with "rising"

Line 37 – they should state their dataset derives from 201 dental calculus samples, as only 201 passed their specified thresholds for inclusion in their analyses

Line 40 – I suggest "identified" be changed to "identify" to agree with tense used in the preceding sentence

Line 67 – grammatical error with "full"

Line 78 – I think it more fitting to replace "Industrialized population" with "communities that pre-date the Industrialized population", since this more accurately reflects the temporal transect they consider

Line 81 – "a proven way" is a bit too strong and does not transparently accommodate biases with the method.

Line 83 – again, this should be 201, to reflect the number of individuals included in their final dataset.

Lines 86 and 87 – there is a grammatical problem here with the tenses

Line 90 – at first read of the manuscript I was not sure what "oral geography" meant. A quick definition of this term, such as position of a tooth in the dental arcade, would be useful for a non-specialist audience.

Lines 99 – 100 – it is not clear if these 954 taxa were found consistently across all datasets, or if 954 species were identified in total.

Lines 120 – 121 – The statement here implies a temporal relationship in microbial community structure, but it is made before the temporal correlation is discussed. Perhaps it could be moved to later section of the discussion? Otherwise your subsequent investigation of temporal correlation seems superfluous

Line 139 – the authors state that oral microflora composition is not linked to periodontal disease or other oral pathology, though I wonder if the authors considered only pathology on the tooth from which they derived their calculus sample. I would assume that a large carious lesion or dental abscess on an adjacent tooth could influence microbial structure globally in the mouth

Line 163 – the authors should be clear that they investigated the presence of genes associated with these microbial functions

Lines 193 – 194 – It should be stated what time periods are represented in their London dataset

196 – the authors should be clear that the physiological or cultural factors *they considered* did not correlate with oral flora content

Line 207 – some of the skeletal pathological features they mention can be age-related. Did the authors consider age of the individual in their analyses?

Line 222 – East Smithfield is a plague cemetery, not a post-plague cemetery

Lines 231 – 232 – It's confusing for the authors to use the language of "arrival of the Second Plague Pandemic", when they're specifically referring to the distinct interval of the Black Death.

Line 240 – This part focuses on the Methanobrevibacter community exclusively, but I suggest the authors reword this to describe the existence of two community structures that are detectable as far back as 2200 BCE, where one of the two declined as of the mid-14th century (if I've properly interpreted their manuscript)

Line 241 – here the authors state that the Methanobrevibacter component diminished only in recent history. How does this accord with their statements implicating the mid-14th century as a period of change in microbial flora? A figure that plots the proportion of this community over time in the population would be really helpful.

Author Rebuttal to Initial comments

We thank the reviewers for each of their thoughtful and detailed comments on our submission. Detailed responses and references to the requested addendums are included below.

Reviewer Expertise:

Referee #1: metagenomics, ancient microbiomes
Referee #2: ancient microbiomes, anthropology
Referee #3: low biomass microbiomes
Referee #4: ancient DNA, anthropology

Reviewer #1 (Remarks to the Author):

The authors here analyze an expansive and potentially quite valuable new collection of human dental calculus from across Britain spanning ancient and modern times. They identify two clusters of microbiome composition, including one that has not been observed in modern humans. They identify potential correlations to diet that might underlie the clusters.

We thank this reviewer for their insightful summary of our manuscript.

There is not enough information to evaluate the sample authenticity and contamination levels of the samples studied here. For a paper describing a rich, new source of ancient human microbial genomes authenticity analysis should be central. But here it is described in only a half sentence in the main text and relegated to a rather cumbersome and unorganized supplementary text and just a few supplementary tables.

We agree that verification of authenticity and contamination is of utmost importance to the validity of this manuscript. To provide further insights into our approach, we now provide a detailed description of our contaminant analysis and authentication in the main text (L99-L114). Further, we now include several additional descriptions of analyses in the SI that demonstrate the authentication and results from contamination analysis workflow. We also now provide a summary of the entire process, alongside the number of samples retained at each step (Figure S3). Overall, we took these steps to ensure authenticity of these samples using the MALTx pipelines:

1.) Compositional assessment against laboratory controls to demonstrate significantly different composition between samples and controls;
2.) Filtered contaminant species identified in laboratory controls and provide access to pre- and post-filtering data for assessment;
3.) Authenticated DNA damage for all sequences in each sample using ChangePoint, a new program that uses a non-reference-based approach to assesses damage across all DNA sequences;
4.) Authenticated damage of known oral and contaminant species detected in dental calculus using MapDamage2.0;
5.) Applied SourceTracker2.0 analysis against modern oral samples, ancient oral samples, environmental controls, and laboratory controls to demonstrate the majority of species are also found in both modern and ancient oral microbiome samples; and
6.) Conducted an analysis of individual genera in ancient calculus samples to identify the proportion of genera that are also found in the Human Oral Microbiome Database (HOMD).

In addition, we also examined the samples using an updated pipeline (as suggested by Reviewer 2) with MALTn, which identifies sequences based on nucleotide alignment instead of protein alignment. Using this pipeline, we cross-validated samples with an increased array of reference sources using SourceTracker2.0 and found concomitant authentication results and overall findings with MALTx (Figure S13 and S14).

There is not nearly enough detail in the Methods on how the clusters were constructed, which is a pivotal analysis in the paper. It's not sufficiently clear how membership is defined in the 3 clusters in Fig3A. What are the co-occurrence rate cut-offs used, how were they chosen, and how does the number of clusters and sample assignment change as those cut-offs are modified?

We thank the reviewer for this comment, and we provide more clarity on the selection of the different groups examined in this publication in the main text (L135) and SI. We observed in Figure 3A that *Methanobrevibacter, Streptococcus,* and *Actinomyces* were the highest ranked genera identified using Biplots and thus were driving most of the 'u' shaped variation observed within the PCoA plot (Figure 3A). Therefore, we assigned each sample to a group according to whether or not *Methanobrevibacter, Streptococcus,* or *Actinomyces* were more abundant in each sample. For example, if a sample had higher levels of *Streptococcus* compared to *Methanobrevibacter*, the sample was placed in the *Streptococcus* group, and so on for each sample. We then colored the PCoA plot according to those groups. No co-occurrence analysis was involved to determine these group assignments. We have now improved the Figure descriptions of Figure 3 and Figure 4 (as the same groups are used in both figures), and we have clarified this process in the main text and SI.

Our full scripts for the co-occurrence analysis are available on GitHub. To elaborate on settings, we used a correlation threshold of 0.3 and a conservative p-value threshold of 0.0002. We also tested p-value thresholds up to 0.1, and the same results were obtained. No specific numbers of clusters were selected using the igraph R package, and we were repeatedly able to visualize two major sets of co-associated taxa across both the genus and species levels. To improve the visualization of the negative correlation results in the text, we limited the number of genera to the most abundant 70 genera. For both visualizations, we used the Fruchterman-Reingold layout algorithm for maximal clarity. This methodologies is now more thoroughly described in the SI methods (pg 17).

An important indication that the clusters are not robust is their lack of new biological insight. It is telling that both clusters appear to be randomly distributed in FigS10 and there is no time-trend supporting the hypothesis that the Methanobrevibacter cluster was a pre-Industrialized cluster. Some earlier time periods do not show the Methanobrevibacter cluster. Also, the cluster is the dominant one in the most recent time period before modern, which includes Industrialized people/samples.

We thank the reviewer for their comment. We have attempted to clarify this point by recreating the previous Figure S10 using absolute sample counts and discrete time periods (now Figure S12). We hope that these clarifications better show the differences in samples available across time. We have also been more intentional with our wording in the text to assess the presence of the *Methanobrevibacter* community over time, specifically investigating whether or not it had shifted at distinct points in time or whether or not it is routinely present today in modern populations.

A version of Fig3A needs to be shown that labels samples by their origin (time period, geography, etc) rather than just assignment to the co-occurrence clusters. It's not obvious that clusters are capturing biologically meaningful information.

Thank you for this feedback; we have added two additional supplementary figures (Figure S7A and S7B) that show samples based on their regions and time periods.

In Fig4 I'm not sure exactly what I'm meant to be observing. I don't really see much of a difference between the Streptococcus, Methanobrevibacter, or Modern groups in this plot. If there are features that are significantly different between groups, can they be labelled with an effect size and P-val, and name of the relevant feature? The corresponding LefSE plots in the supplement are not sufficient to demonstrate a statistically significant difference between groups. LefSE is designed to be a data exploration tool. Appropropriate two group or ANOVA statistical tests must be applied to each feature. The authors can use a different BioBakery tool such as MaAsLin for this purpose or one of the many other packages for multivariate analysis.

We thank the reviewer for their critique. This figure demonstrates the abundance-based differences of selected microbial functions linked to diet. We now provide more clarity by reformatting Figure 4, only including functions that are significantly Streptococcus-associated or Methanobrevibacter-associated based on a newly completed ALDEx2 Benjamini-Hochberg corrected p value of Welch's t test (<.05; Figure S15 and Table S21). ALDEx2 analysis was recently shown to be the most robust for abundance-based detection (Nearing, et al. *Nature Communications*, 2022), so we chose that tool in addition to the LefSE tests. Both tests resulted in similar findings highlighting the dietary differences between groups. In Figure 4, the names of each function (Table S14) were physically too big to include in a main text figure; as a result, we provided a full, detailed table (identical to Figure 4) as SI Table S15. We now provide clarity on this in the Figure 4 legend.

There is a statement (line 119) that oral microbiome diversity was likely greater in pre-Industrialized oral microbiomes than modern Industrialized today. Diversity in what sense? I don't see direct metrics of diversity. If the diversity is significantly different, does this analysis adequately account for the unavoidable contamination bias in ancient samples that could inflate diversity metrics?

Thank you for indicating that this should be reworded. Our intention was to suggest that the *Methanobrevibacter*-associated oral community we identified in ancient Britain is largely absent from modern individuals, and that, as a result, modern industrialized communities may be lacking this diversity. We have reworded this section to better reflect our conclusion.

Please do not display 3-dimensional plots as static figures, paper printouts only have 2 dimensions. Represent 2 axes at a time and use video files for 3D rendering if really necessary.

We thank the reviewer for sharing this perspective, although we respectfully disagree, as the size of the spheres in the plot are consistent with the depth of field.

The manuscript has many references to time periods (Civil War, Medieval/post-Medieval in Table 1 for example and also Fig1) that may not be easily recognizable to non-British people. It is better to use carbon-dating verified date ranges or numeric time periods.

We agree with this suggestion, which is a complex issue due to the lack of carbon dates for all of the samples. To address this, we conservatively selected the oldest possible date for each individual and provide the number for each sample under the 'Date' variable in Table S1. The time periods we utilize in this paper are associated with the following time periods: Pre-Roman Britain (Colonization – 43 CE), Roman Britain (43 CE – 410 CE), Anglo-Saxon Britain (410 – 1066 CE), Norman Britain and the Middle Ages (1066 CE – 1547 CE), Reformation (1547 CE – 1750 CE), and Industrial (1750 CE – 1900 CE). Where possible, we have now added these specific date ranges in the main text wherever period was mentioned (i.e., Figure 1 visual and legend) and added these specific date ranges to the supplementary text under the "British Time Periods" Section.

Reviewer #2 (Remarks to the Author):

In this manuscript, Gancz et al. study the oral microbiome diversity of British populations through shotgun metagenomic sequencing of ancient dental calculus sample spanning the Bronze Age to the mid-18th century.

Overall, the manuscript is well-written. The laboratory work is exhaustive and has resulted in an excellent dataset of >200 ancient oral metagenomes from Britain with associated meta-data. The data analyses are well-thought out, but there a few issues to consider (see below comments). I appreciated the public outreach that has been conducted as part of this study.

Major Comments:

One of my major concerns is the use of a MALT-x (nucleotide-to-protein alignment) approach for taxonomic profiling in this study. This approach is not used in most ancient microbiome studies, and in fact, previous research from this group (Eisenhofer and Weyrich, 2019) has shown that nucleotide-to-nucleotide alignment approaches, such as MALT-n, outperform MALT-x for profiling ancient microbiome datasets. Given this, it is unclear why the authors use a MALT-x approach in this study. If computing power necessary for running MALT-n is a limitation, perhaps an alternative approach would be to use Kraken and/or a customized database.

We thank the review for this suggestion, and the authorship team is well aware of these limitations for taxonomic identifications (i.e., Eisenhofer and Weyrich, 2019). However, MALTx was selected to provide better information on the *functional* content of these microbes, as MALTx leverages protein sequences for mapping. This was specifically done to complete functional analysis that is critical for the dietary assessment work included in this paper. Functional analysis in the ancient microbial community is currently limited, as other functional tools can provide limited assessments for ancient species. Our analysis was also initially done prior to the inclusion of the AADDER program to MALTn, which now allows for some limited functional information to be obtained while using MALTn (Yates, et al. *PNAS*, 2021). However, the code for AADDER has not yet been published as a stand-alone program, which is critical to assess its efficacy against MALTx. Further, AADDER examines the functions of microbes using well annotated reference genomes. This is problematic for some of the critical taxa within this paper; for example, there is only one oral *Methanobrevibacter* reference genome (*Methanobrevibacter oralis*) in the NCBI database, compared to 59 for one oral *Streptococcus* species alone (*Streptococcus sobrinus*), which could bias our analysis significantly. The MALTx approach does not limit the functional analysis to well annotated reference genomes and leverages a variety of identified proteins to perform its current functional analysis.

Nevertheless, we sincerely appreciate this comment and now include a comparative analysis between MALTn and MALTx (Figure S14); we also provide an additional authentication assessment using SourceTracker2.0 on MALTn results (Figure S13). We then tested these data sets to see if the taxonomic communities are the same using a mantel test with a Spearman Rho; this test confirmed that taxonomic assignments using either method were correlated with one another (Spearman Rho=0.665231; p-value= 0.001). We also confirm the same U-shaped PCoA plot of MALTn identified taxa (Figure S14) and ensure that our conclusions are tempered to the genus level to avoid overinterpretations of the MALTx data.

The analysis of these new ancient British dental calculus samples provides interesting findings, such as oral communities dominated by either Methanobrevibacter or Streptococcus. But these findings can be better contextualized using previously published ancient dental calculus samples from the UK. However, the authors include samples from Neanderthals and medieval Germany as comparative samples, which do not seem relevant to this analysis. Instead, it would be good to include the data from Velsko et al. 2019 (40 ancient dental calculus samples from a single cemetery in Oxford). Also, is there any overlap between the Jewbury samples included in this study and the ones from this group's previous study (Farrer et al. 2021)? The latter seems to have considerably more samples and it is unclear if/why these were not included.

We assume that the reviewer is referring to the SourceTracker analysis in this comment, as Neandertals and medieval German samples were *only* used as ancient calculus sources in the SourceTracker analysis for this paper. We thank the reviewer for suggesting a comparison to the Velsko paper. As such, we completed our new SourceTracker2.0 analysis with the Velkso et al. 2019 data, alongside a further improved list of sources from Cadiff et al 2017, Larson et al 2022, and Perez et al 2019. This new SourceTracker analysis confirms the authentication results that we observed previously. We were unfortunately unable to include the Velkso et al. 2019 data in other analyses in this paper, as the samples come from a mixed dentition (i.e., calculus was subsampled from multiple teeth and pooled before extraction; Figure S18). As such, we are unable to account for the oral geography signatures in these individuals and were not able to include them here, where the impact of oral geography plays a significant role in our analysis.

In general, the SourceTracker analysis could have been better described. In Table S20, it is unclear where the modern dental calculus and soil samples used as sources came from – the table seems to suggest they were generated as part of this study? If so, what is their context and how were these data generated? If these were from previously published datasets, they should be cited. Second, in addition to modern dental calculus and EBCs, the number of other source samples is very low. The authors could include other oral sources such as modern dental plaque, skin, and more soil samples in their analysis to make it more rigorous. That may help reduce the high % of Unknown in Figure S2B. Third, I do not understand how Figure S2A was generated – what Ancient Dental Calculus samples were used as a source here? Regardless, I do not think it is appropriate to use other ancient dental calculus samples as sources in this type of analysis, since one can never be certain that their
taxonomic profile is free from non-endogenous taxa.

We now provide a more detailed description of the SourceTracker1.0 analysis used in this paper and also include a new analysis using SourceTracker2.0. We maintained the SourceTracker1.0 analysis in the paper for better comparisons to other published approaches (Weyrich et al, 2017), and we now include clarity on the approach, including references for the sources (SI pg. 11; Table S20).

For the new SourceTracker2.0 analysis, we included a detailed description of the approach (SI Text pg. 11-12), used an increased number of sources (including the Velsko, et al 2019 data), and new prepared additional figures (Figure S13A-B). Similar results were observed between the two approaches, signifying the robusticity of the approach. We also compared dominant taxa to those found in the Human Oral Microbiome Database (HOMD) as a further analysis to support the robustness of oral taxa found in this data.

The presence of a Methanobrevibacter-dominated oral community was found to be associated with size of calculus deposit. Given modern dental practices and frequency of calculus removal, it is likely that

modern dental calculus deposits do not grow sufficiently large to provide the anaerobic environments that support the growth of Methanobrevibacter. This point is mentioned rather off-handedly in the Main Text, but in my opinion, it seems to be the most likely explanation for the "disappearance" of this Methanobrevibacter-dominated ecology from modern dental calculus. The dearth of metagenomic sequencing data for modern dental calculus samples from Britain, and in general, from different dietary and geographical contexts, limits understanding whether Methanobrevibacter-dominated communities continue to exist even today. This point receives some attention in the Supplementary Material, but I think it needs to be discussed in the Main Text.

We thank the reviewer for this insightful comment and have included additional discussions of this topic in the Main Text (L181-191). We additionally ran another adonis test to further identify the species driving the calculus fragment sizes and to see if the 'Dominant grouping' that each sample fell into interacted with the calculus fragment bin. ALDeX2 results suggest that *Gemella* and *Lautropia* genera are associated with the differences in calculus fragment size (Figure S17B); both genera are linked to the Streptococcus group. However, the adonis test also failed to identify any interaction with the 'dominant grouping' and calculus fragment size. This suggests that the fragment size is being drive by signatures that are not likely within the *Methanobrevibacter*-associated group (p=0.101), as previously suggested. Nevertheless, we agree that calculus size is a contributing factor in the description of the communities we observe, although we do believe it is also important to factor in additional information on dietary and disease associated in this study. We also add this into the abstract and clarify that hygiene is a major driver in the concluding paragraph (L51and L300).

Minor Comments:

Supplementary Methods: In addition to removing reads greater than 300 bp, was a minimum length and quality threshold (e.g. Q30 and length 30?) applied to reads after merging?

We thank the reviewer for requesting this clarification. The quality threshold utilized in this study was Q30, and the minimum length was 25 base pairs. This information has now been added into the SI text (pg. 9).

In my opinion, the term "industrialized" should not be capitalized when referring to industrialized countries, industrialized/pre-industrialized/non-industrialized microbiomes, or industrialization.

We thank the reviewer for their comment and have implemented these changes in the Main Text.

Line 38-39 and 85-86: In Lines 38-39, the authors state that they are the first "to directly reconstruct the

microbial history of an Industrialized population for the first time". In Lines 85-86, they state they "… directly describe the history of a pre-Industrialized microbiome for the first time." These sentences are overstating the findings of this study. Both sets of lines should be rephrased to say something along the lines of "this is the first study to directly reconstruct the oral microbial history of a population spanning the transition to industrialization".

We thank the reviewer for this comment and have rephrased both segments of the Main Text accordingly.

Lines 48: Recommend changing "unappreciated" to something else, due to repetitive use of appreciated in this and the previous line.

We thank the reviewer for this comment and have rephrased the Main Text as indicated.

Line 58: Obregon-Tito et al. 2015 could be added as a citation as well.

We thank the reviewer for their comment and have added this publication as a citation.

Line 91: It would be good to discuss findings from Fagernäs et a. 2021 about oral geography affecting oral microbiome composition, since that paper also studies this using ancient dental calculus samples.

We have included discussion and citations of ancient DNA oral geography studies (Farrer 2018, Fagernäs 2022) into the SI.

Lines 116-119: It is unclear how the authors "confirmed that species within these co-occurring genera can cohabitate with Methanobrevibacter oral species, as all are anaerobic and can produce metabolic by-products that support methanogenesis". Was this done using just literature review, or was any kind of actual experimentation involved?

We thank the reviewer for this comment and have adjusted the text to indicate that this was ascertained via a literature review (L139).

Line 138: Correct typo in "populations".

We have made this adjustment and thank the reviewer for pointing it out.

Lines 142 and elsewhere – I recommend referring to the communities consistently as ”Streptococcus-dominated” and “Methanobrevibacter-dominated” rather than just with the genera names.

We have implemented this clarification throughout the Main Text, and we have shifted the name to now better reflect our group categorization, as suggested by Reviewer 1 to now be: *Streptococcus*-associated and *Methanobrevibacter*-associated communities.

Table S13 is difficult to interpret – specifically with regards to Line 143.

We thank the reviewer for their comment and have improved the design of the table. In addition, we have altered the description of the table in the main text in an effort to make it easier to interpret.

Line 163: Specify which modern, industrialized population was studied.

We thank the reviewer for this comment and have clarified that this was a modern European, Spanish population (L190) and have additionally added a reference to Table S1, which describes these samples in richer detail.

Line 282: Correct typo in “AdapterRemoval”.

We thank the reviewer for this comment and have adjusted accordingly.

Figure 4: I recommend switching the color scheme, since readers will instinctively associate the color red with higher abundance.

We thank the reviewer for their comment and have adjusted the plot accordingly.

Figure S5: Some X-axis labels are cut off.

We thank the reviewer for this comment and have adjusted the figure accordingly.

Supplementary Information should be proof-read for typos.

We thank the reviewer for this comment and have endeavored to improve the readability of all documents.

Reviewer #3 (Remarks to the Author):

It is of great interest to investigate the evolution of the oral microbiome. The question that poses itself is whether we are dealing with an improvement or a devolution, like is seen in the gut which is associated with an increase in the incidence of various diseases of affluence, or whether it is matter of perspective. Whilst oral health has improved in modern times the Methanobrevibacter-associated community was not found to be associated with periodontal disease (whilst bugs associated with Streptococcus & co. were). Perhaps even more specifically, is the increase in oral hygiene linked with the decrease of gut health? Or is there perhaps a case to be made to adjust modern oral hygiene so that the Methanobrevibacter-associated community still has a chance? In Western populations Methanobrevibacter (associated with low BMI) is slowly also disappearing from faeces (see paper on Irish travelers, for example) being replaced by a more Bacteroides dominated Microbiome (https://www.nature.com/articles/s41591-020-0963-8/figures/8). Is oral seeding of the gut no longer possible or much more difficult due to oral extinction of Methanobrevibacter (& friends) as a result of supposedly improved oral hygiene? Please feel free to incorporate or discuss aspects of the above in the discussion.

This is an incredibly insightful comment, and we thank the reviewer for it. It is difficult to recovery ancient gut microbiomes from a technical perspective (i.e., see Weyrich, et al. JHE, 2013 for more detailed discussion). Nevertheless, it is interesting to posit additional studies that could be done to test this. For example, ongoing research in our team is exploring the strain diversity of *Methanobrevibacter* species within these individuals and assessing their origins and the evolutionary pressures they may be under over time. The same analysis could be comparatively done alongside the few modern oral *Methanobrevibacter oralis* isolates that have been isolated from living people today (Lepp, et al. PNAS, 2004). We do note that the gut *Methanobrevibacter* isolates (typically *M. smithii*) are quite distinct from *M. oralis*; for example, the gut-associated *M. smithii* has large gene regions for adhesion to epithelial tissue that are lost in oral *M. oralis* genomes (Weyrich, et al. Nature, 2017). Once these microbes are lost from the mouth or gut, we may not have the ability to 'reseed' quickly from another body site, which may explain by methanogens are becoming less common in certain body sites at different times.

Major comments:
1a: The authors are of course clearly aware of the issue of contamination and they clearly did something about it (Table S4) yet I can't really ascertain how they checked for laboratory and environmental contaminant DNA. The main thing I found was the abundant use of negative controls. One step that was perhaps even too strict was that species identified within negative controls samples were removed from

all samples. Sometimes genuine signals end up in negative controls due to a variety of reasons (barcode bleedthrough, machine contamination, "splashome", etc). Was it verified that these signals found in the negative controls (and subsequently removed from real samples) were indeed not of ancient origin? Did you compare/assess DNA damage levels for each of these signals? Which leads me to my next question.

We thank the reviewer for their comment and appreciate the attention they paid to our authentication protocols. In order to improve the clarity of our procedures, we now elaborate on this topic in the main (L98-116) and supplementary text (pgs. 10-13) and provide a new overview figure outlining our sample retention and authentication steps (Figure S3). To elaborate on our conservative filtering approach, we examine the data using both strict filtering and no-filtering to make sure our interpretations are sound. For example, we did filter out species found in EBCs from species-level analyses, but no higher taxonomic levels (as this can remove whole genera that likely have biological and contaminant taxa, e.g., *Psuedomonas* (Eisenhofer, et al. Trends in Micro, 2019)). We then also conduct our statistical analyses on both genus-level analysis where no filtering occurred and species-level where filtering occurred. The findings were similar for both filtered (species-level) and unfiltered (genus-level) data sets, except in association to cemetery (Main Text, L250), suggesting that contamination may add to the signatures across cemeteries, but little else. Secondly, we extensively characterized the EBCs to ensure that ancient oral signal was not present in them. To help demonstrate this, we include a full list of taxa found in EBCs (Table S4) and now include MapDamage2.0 analyses on four major taxa found in the EBCs that were filtered from the data (Table S13; SI text pg. 13). These species all contained damage consistent with the geologic age.

1b: It is unclear how the authors used DNA damage levels of reads as a tool to ascertain which reads are genuine (damaged due to old age) and which ones are contaminants. I do see mention of: "Reads longer than 300 bp were discarded, as they likely represent modern contamination." and that "Shotgun metagenomic libraries were generated without enzymatic damage repair" but I do not see mention of making use of tools such as MapDamage2. For example, in the articles of Wiboho, Granehäll and Rampelli (references 16, 24 and PMC7864912) they for example use MapDamage2. I assume that DNA damage profiles were assessed and used for screening reads, and if so, this needs to be written down more clearly. If not, then this simply needs to be done.

We thank the reviewer for pointing out the lack of clarity in the text. We were using a new reference-free damage tool called 'ChangePoint' to authenticate DNA damage in the previous version. However, we have now added the gold-standard MapDamage2.0 analysis to this paper and discuss the results in the main text and SI (SI pg. 13; SI Table 13).

1c: In short, I need a lot more information/detail on how contaminants were separated from genuine signals; not just that negative controls were used. Furthermore, there are many more tools than just the

use of negative controls or DNA damage (though both are very important) to identify contaminants. Commonly, contaminants from the same source typically correlate strongly with one another allowing one to use contaminants to identify other contaminants. Also for example, you could perform a correlation analysis of all signals with axis 1 and axis of Figure S1. All signals strongly negatively correlated with both axes are contaminants. Furthermore, were particular signals perhaps part of an artificial batch effect (associated with the use of a particular lot-number)? It is hence good to show per signal that they are 1) ancient/damaged (yes/no), 2 present in negative controls (yes/no), 3) associated with other contaminants (yes/no) 4) part of a batch effect (yes/no), something else you can think of (strain-tracker(?)) (yes/no). A multi-layered approach is recommended and the separation of genuine/contaminant signals should not be a mere side-note.

We thank the reviewer for their attention to authentication, as this is extremely critical in ancient DNA studies. We now provide an overview of our filtering and authentication approaches in the SI (Figure S3). With respect to the reviewer's comments that information from Figure S1 could be used to identify contaminated signals, we did remove samples that were more proximal to EBCs than to other samples (Table S2). We also remove any species within the EBCs at a species-level, which is a conservative approach that is examined carefully by conducting analyses with both filtered (species-level) and non-filtered (genus level) results for all compositional analyses (Table 1). While we appreciate tools to assess contamination from biological signal (i.e., decontam), we did not observe significant oral signatures in our extraction blank controls (Table S4), so a more conservative approach was used. We also specifically addressed each reviewer's comment, as below:

1.) We now include damage profiles of four major contaminant taxa in the EBCs and demonstrate that the taxa are likely modern in origin (Table S13)
2.) We provide a plot of all EBC taxa in Figure S4D. We also provide the full list of species that were filtered from the samples (Table S4), which is overwhelming not oral. Some minor potential oral species are found in the EBCs (including *Aggregatibacter actinomycetemcomitans*, four species of *Corynebacterium*, one species of *Neisseria*, one species of *Streptococcus* (*S. pneumoniae)*, and 10 species of *Sphingomonas*). However, the species diversity within these genera is largely maintained post filtering, as many of the major oral species within these genera are maintained. The strict filtering approach in this study does not appear to be too strict, although we do agree that this approach may not be possible if the EBCs were more contaminated. We have added rational and discussion around this topic in the SI text (pg. 10-11) and now also show how total read counts were minimally influenced by our filtering strategy (Table S4).
3.) We also assessed the EBCs to ensure they were similar to other published extraction controls (i.e., summary of common contaminant genera from several contaminant studies is present in Eisenhofer, et al. Trends in Micro, 2019). A new table (Table S22) now shows that many of genera detected in the EBCs of this study are indeed shared as contaminants in other studies. However, it is critical to note than many oral taxa are routinely oral in many studies that are conducted outside of a clean lab, due to the shedding of oral taxa from technicians (Weyrich, et al. Mol Ecol Res, 2019), and that nearly all of these genera are absent in our EBCs.

4.) Our batch effects from DNA extractions and library preparations were observed even after conservatively filtering all species in EBCs from the samples, suggesting that the batch effects are driven by something beyond the contaminant taxa (Table S8). This potentially suggests that efficiencies in our homemade extraction protocols may vary from batch to batch, as is common with DNA extractions in the microbiome field (Shaffer et al, 2022), especially when different contaminant species may be present at higher or lower levels (Marotz et al, 2017).
5.) We now added an additional SourceTracker2.0 approach to better investigate this (SI text pg. 11-12). This analysis cross validated the strong oral signal and limited contributions of contaminants in our study.

2: Possibly related to contamination: large differences between different cemeteries were found ("We found that 64.83% of the compositional variation explained by time could equally well be explained by cemetery"). How did DNA-damage profiles differ between cemeteries? How do the DNA damage profiles of the different bacteria that are mostly representative of this "64.83% compositional variation" compare with the DNA-damage profiles of species of which one is 100% certain they are genuine? Do older cemeteries/remains btw in general indeed show more DNA damage?

We thank the reviewer for this comment and have added an addition analysis to investigate how damage linked to cemetery may contribute to our observed compositions shifts. A description of this approach is now included in the SI text (pg. 19-20) and the results into the main text (L282-286). We incorporated a marker of overall DNA damage acquired from oral taxa analyzed via Mapdamage2.0, deltaD (Table S13), into our Bayesian multinomial logistic-normal linear model. When we include damage seen in representative oral species, only 1.5% of the variation is explainable by damage values (95% CI: [0.57%, 3.08%]). We interpret this to mean that damage is a poor predictor of the signal that is associated with differences across cemeteries and had limited impacts on our interpretation of change over time.

3: Abstract: "While many factors likely contribute to this observation, microbial functional analysis suggests that common oral microbes today may reflect past differences in diet, specifically dairy and carbohydrate consumption. In London, the less common microbiome is linked with skeletal pathologies, and its disappearance is consistent with temporal shifts, including the arrival of the Second Plague Pandemic. These findings suggest that pre- Industrial microbiomes may have been more diverse than previously appreciated and linked to health in previously unappreciated ways, shifting our understanding of chronic, noncommunicable disease origins in Industrial populations.

The dietary aspects indeed seem credible (have changed significantly in modern times and was likely associated with social status in ancient samples) but the others factors mentioned don't yet seem that convincing. When looking at Figure S10 it is quite clear that the Methanobrevibacter-rich-composition was still live and kicking between 1600-1900. The main elephant in the room, is improved dental hygiene. Specifically, it has been found (a main finding that should be mentioned in the abstract) that the

Methanobrevibacter cluster was associated with calculus fragment size. This if anything most likely explains why the Methanobrevibacter cluster has gone the way of the Dodo. The Methanobrevibacter cluster & this relevant association are not even mentioned in the abstract. I think the abstract can be rewritten/refocused a bit. The summary does not really mirror the last paragraph of the main text.

We thank the reviewer for this comment and have worked to make the SI and abstract more coordinated and to highlight the potential impact of shifts in hygiene in both the abstract and the discussion (L53and L300). We agree that the potential impact of dental hygiene should be more prominent in the paper.

4: In the text it is said that "the Methanobrevibacter-associated communities were associated with the presence of several skeletal markers of systemic disease, including non-specific periostitis, joint porosity, osteophytic lipping, and overall scores for joint pathologies". Is this association perhaps not indirect? Were these markers of systemic disease perhaps associated with age? Is age perhaps associated with larger ("more mature") dental calculus size? Alternatively, is this perhaps due to the possible association with lower economic class? This is alluded to somewhat in the Supplemental section but that this association is not unlikely indirect should probably be mentioned in the main text more clearly.

We thank the reviewer for this comment, and this is indeed a key finding that our team is working to explore in detail. We agree that the effect could be associative or indirect, so we hesitate to provide explanations until further analysis is done (i.e., mechanistic studies into these microbes and the chronic, systemic diseases they may cause, as well as age-associated analyses of frailty). Accurate age estimates were not available, and we are currently working on employing Transition analysis to obtain more accurate age estimates for each individual. Given the current relocation of the individuals from the Museum of London, we could not include these in the current publication. We also examined socioeconomic contributions to these findings, but the archaeological evidence in London, where our data is heavily biased, suggests that past assumptions about socioeconomic status made from the individuals were done in a way that did not accurately represent the population. As such, we have also avoided sociodemographic conclusions in the main text but included it in the SI discussion (pg. 21-24) as a possibility. Nevertheless, we now briefly modified the wording of the main text (L260). We also add in further testing to try and understanding the factors driving calculus fragment size, including an adonis test (L450-456) and ALDEx2 analysis (Table S17), which suggest it may be Streptococcus-associated taxa are driving these observations. Clearly more work is needed in this area to link calculus size and the species living within.

5a: Lines. 225-239. A lot of lines and complicated analyses and some numbers are given in regards to the possible effect of the black death on the oral microbiome but I do not see any sort of direction being given. What kind of (small but significant?) shift did this (temporarily) induce? Did food become more affordable as there were less mouths to feed causing a shift to the supposedly "higher status"

Streptococcus composition? Or because people took oral hygiene more seriously for a while leading to more Streptococcus? Or did Methanobrevibacter (or Prevotella) become more dominant?

This is indeed an interesting and complex problem that requires further research. We now include a detailed paragraph of discussion of these post-plague impacts on England and what this could have meant for the oral microbiome (SI text pg. 21).

5b: Similarly, it is stated that "Our results suggest that the composition of oral microbiota in ancient London changed over the 800-year time- period from 1100 to 1900." yet I'm still in the dark in regards to the direction of this change or the reasons for this supposed change. It is also stated that "Our study reveals the existence of a now rare or even extinct oral microbial ecosystem that remained relatively consistent in British populations over at least 2,200 years. This oral microbial community persisted through major bio-cultural transitions and historically significant socio-political events, only to diminish in recent history – a phenomenon associated with the rise of Industrialized diseases." How do you rhyme these two statements? The latter one seems logical yet I'm not convinced about the former one.

We thank the reviewer for this comment. In an attempt to clarify this point, we have updated Figure S12 and amended the wording of the main text (L292). Our statistical modeling of composition suggests a change in the overall composition of the microbiome from 1100-1900, but this does not mean that the *Methanobrevibacter*-associated community is missing during that time.

6: Figures 2 and 3: Figure 2A is supposedly of all teeth and I reckon Figure 3 is also of all teeth. Why then are these plots not identical? Figure 3 is stated to be of the genus level while I'm unsure about Figure 2. Why not use the coordinates of Figure 3 for all selections for Figure 2B, 2C (and 2A?)? I think this would be more informative in regards to ascertaining which species are associated with the differences shown in Figure 2. Would it perhaps be an idea to first show Figure 3 and then Figure 2 (so switch them, and the text, around)? One of the most interesting findings of biological relevance is the association between calculus fragment size and Methanobrevibacter. You can see this somewhat unclearly currently in figure 2B. This would be clearer if the coordinates of figure 3 are instead used.

We thank the reviewer for this point. Figure 2A and 3A are not identical due to the differences in metadata available about respective samples. That is, samples that did not have information regarding tooth-of-origin were excluded from the analyses in Figure 2. Due to these limitations as well as oral geography being critical for downstream analyses in our paper, we have chosen to present the figures in this order. Nevertheless, we now provide Figure 2A-B with the coordinates from Figure 3 (Figure 3). Figure 2C is still limited to the tooth type where this observation as made. In addition, we also ran an adonis test to see whether fragment size and the dominant category interacted, and we found that they do

not (p = 0.101), suggesting other taxa are driving the size of calculus. This information is now added into the main text (184-191) in the context of hygiene.

7: Supplemental Tables: These contain a lot of stuff but not simple tables showing the abundance of each bacterial signal pre- and post-contamination removal, both on the species and genus level in each individual sample (and in each of the negative controls; also show the order in which samples were processed for identifying possible batch effects). People could of course spend a lot of time processing the reads themselves but a simple summary like this would be valuable, especially for a reviewer (with limited time) who wants to do some checks on the reagent contamination removal process. At least Figure S3 gives a lot of per sample information (visually) but a simple excel file. (with a few tabs) would do a better job for purposes of being able to check the numbers.

We thank the reviewer for this comment and have addressed this point by doing the following. First, we now provide a new supplemental figure (Figure S4) that show the genera and species taxa summaries before and after filtering, as well as how filtering directly impacts the read account in each sample (Table S4). We also added a new summary figure (Figure S3) that provides an overview of the contaminant assessment and removal process, as well as indicates the number of samples retained and discarded at each step. Moreover, by looking at the 'Used' column in Table S1, it is possible to see exactly which samples were discarded, and for which reasons. We also now provide the full pre- and post-species filtering files on the github depository for this project.

8: Project Outreach. Why is this in the Methods section? It has more in common with something like an acknowledgement section if anything. Its position in the Supplementary information section is just fine on the other hand (and currently kind of double with the main text). There is enough I'd like to see in the Methods section which I'm currently not yet seeing (contamination issue).

We thank the reviewer for this comment and have added more information about the contamination and authentication approaches utilized in this paper, as well as an additional supplementary figure (Figure S3). The project outreach section was included because there are strong debates in the ancient DNA field about good ethical practice (i.e., ensuring community input). We strongly believe that acknowledging the local community and the work we put towards engaging with them is extremely important for this, and all, ancient DNA research. We would also like to highlight that Reviewer 2, with experience in ancient microbiome work, appreciated this part of the study.

9: Table 1: I mainly see a lot of numbers of stuff linked to oral microbiome composition without even telling me the direction when a significant different is found. Might be more of interest if an indication of such a direction is added (like superscript "M" or "S", association with Methanobrevibacter- or

Streptococcus-rich composition. Also, I cannot understand the Table without reading the main text; the legend might need some work. What does "dominant" furthermore stand for?

We thank the reviewer for their inquiry. Table 1 was conducted on the microbiome as a whole, rather than specifically on whether or not *Streptococcus*- or *Methanobrevibacter*-associated communities were driving these effects. To help aid in this inquiry, we performed chi-square or Fisher's Exact tests (as appropriate) to investigate whether the dominant category classification was significantly associated with the significant variables from Table 1 (Table S23). We also conducted ALDEx2 analyses on each of the significant variables contributing to compositional shifts (Figure S17). In some cases, other taxa are driving these relationships and not the presence of either just *Streptococcus*- or just *Methanobrevibacter*-associated communities.

Minor comments:
M1: It might perhaps be worth it to draw a parallel between the two distinct oral microbial communities found and the Prevotella vs the Bacteroides enterotype in the gut. The Prevotella community has largely disappeared or is disappearing (depending on which country you're looking at). Have you perhaps tried to do some kind of enterotyping approach to see what rolls out of that?

We thank the reviewer for this comment. While enterotyping has been used in association with the gut microbiome, there remain questions about the existence and robustness of enterotypes described (Cheng, et al. 2019). Due to the limited number of samples in this study as well as its exploratory nature, we have chosen not to pursue enterotyping at this time.

M2: Figures 3B & 3C: A co-occurrence network can also be visualized using a heatmap using hierarchically clustered spearman rho coefficients. Core members of the 2 clusters should be easily visualized with such an intuitive and simplistic approach. A heatmap might even visualize sub-clusters more clearly (maybe shed some light on Prevotella-rich compositions shown in the middle of 3A?).

We thank the reviewer for their comment and have provided a heatmap as requested. We have now included this in the SI (Figure S16) and a description of this approach in the supplemental text.

M3: Figure 4: Table 15 is stated in the legend. This should be Table S15.

We thank the reviewer for noticing this error and have corrected it accordingly.

M4: Figure S4A: Where is the Streptococcus arrow?

We thank the reviewer for this question. No single Streptococcus species emerges as a top contributor to the variation when a reasonable number of arrows (<40) are included, which is not to be unexpected given the high level of *Streptococcus* species diversity which is commonly found in a single individual's mouth (Abranches et al, 2018). However, as noted in 3A, the streptococcus genus, as opposed to a singular species, is a top 5 driver of variation.

M5: Page 13 Supplementary information. Streptococci should be Streptococcus (italics) or streptococci (not in italics).

Thank you for this comment. We have modified the text accordingly.

M6: Line 67:.. sites has not been full examined …
-fully

Thank you for this comment. We have modified the text accordingly.

Reviewer #4 (Remarks to the Author):

The manuscript by Gancz et al presents the molecular analysis of 201 calculus samples from Britain, representing a time period from 2200BCE to the 19th century. Through this analysis the authors aim to investigate the origin of the Industrial oral microbiome through correlation with time, health as determined by detectable skeletal pathology, and broad cultural classification.

There seems to be a lack of accord between the research focus established in the introduction and the study the authors present. The introduction is crafted to explore the origin of the modern Industrial, or Western, gut. With a dataset that derives only from calculus and hence oral flora, the lengthy discussion of modern guts seems misplaced. I would recommend a focus in the introduction that is bears closer relationship to the actual study presented, namely exploration of oral microflora.

We thank the reviewer for this comment. We have now reworded the introduction to be more holistic and incorporate gut and oral microbiome data (main text L71) including a new reference on the impacts of industrialization on the oral microbiome. We have left some discussion on the gut for context, as this dominated the literature on industrialization and the microbiome.

The manuscript very rapidly moves from a lengthy introduction to a discussion of results. I suggest the authors better prepare their audience for digestion of their analyses by describing their sample dataset in a few sentences. Also, Figure 1A should ideally have the location circles color-coded to reflect their chronological age.

We thank the reviewer for their suggestion and have added more information about samples to the main text (Lines 93-96). We have also added Figure 1B that now includes chronological information on the samples, as requested by another reviewer.

The novelty in this study comes from the restriction of their analysis to the geographical confines of Britain (namely England and Scotland). Their central conclusion is that two broad microbial communities are detectable in their datasets, and a shift in the relative proportions of each of these communities (at least in the city of London) is temporally correlated with the mid-14th century. The authors in turn speculate that the mortality event of the Black Death and unspecified cultural, genetic, or population health changes related to this event may have provided the substrate upon which shifts in oral microbial content occurred. While this is supported through statistical analysis, I suggest the authors display this trend visually in some way. It would be useful to see a graph that plots the population proportion of the Methanobrevibacter component over time to observe the purported effect of the mid-14th century.

We thank the reviewer for this comment. As a response, we have improved Figure S10 (now S12) by separating the samples by discrete time periods. With this figure, it is possible to see the number of samples in the two communities as they change over time. We have also altered our language around changing diversity over time, as suggested by another reviewer.

What seems missing is a comparison of the dominant post-14th century Streptococcus oral community with modern oral community structure. That would be needed to justify their research focus on origins of Industrial microbiomes outlined in their introduction. It would also be needed to justify their final statement in the abstract that the analysis they present changes our perspectives on the origins of non-communicable diseases, which are assumed to be linked to Western gut microflora (and perhaps indirectly to modern oral communities).

We thank the reviewer for this comment. Showcasing the proximity between the historic *Streptococcus*-dominant oral communities and modern oral microbiomes is indeed a focus of this paper. As such, we

provide Figure S6 (PCoA of ancient *Methanobrevibacter*-dominated and *Streptococcus*-dominated samples alongside modern comparisons) and Table S11 (description of the top ten genera in the correlation matrix for all three groups, including modern). In addition, we also added modern comparisons to the dietary functional analysis diagram in Figure 4. We hope that these comparisons to modern individuals may provide more insights into the similarities between *Streptococcus*-associated communities and those found today.

Could the temporal correlation be influenced by the fact that the dominant time period of the collection is from the Medieval period? Figure 1B should also provide the time ranges in years the authors used to define their time periods.

We thank the reviewer for this comment and have included time ranges to Figure 1 (now Figure 1B. In addition, the number of samples is accounted for in the statistical model.

Their statement in lines 97 and 98 that tooth location in the dental arcade influences oral microbial structure is an important one, and their conclusion that this phenomenon could have compromised interpretations made in previous works is indeed provocative. For such a strong statement, however, quantitative data should be presented and thoroughly discussed in the main manuscript.

We thank the reviewer for this comment and describe this finding more in detail in the supplementary text (pg. 13-14). To better address this issue, we downloaded the British dental calculus data from Velsko et al. 2019 and included an analysis of composition according to tooth type, as well as single or mixed dentitions (Figure S15). As demonstrated in the PCoA of Bray Curtis distance, and as is supported by our analyses of oral geography biases (Tables S16-S19), mixed dentitions blur this signal. We have made the text more specific to directly address to this issue.

The authors make quite a huge jump in their correlation of oral microbial content and systemic disease. It's important to recognize that not all systemic diseases contribute to detectable morphological changes in the skeleton. Also, not all modern chronic and non-communicable diseases, to which they refer in their introduction as being related to modern microbiomes, will contribute to physical changes in the skeleton. In short, their statement here (lines 213 and 214) is too strong and carries too many caveats.

We thank the reviewer for this comment and have adjusted the wording of the relevant sentence. While not all systemic diseases contribute to morphological changes in the skeleton, the association between oral microbiomes and systemic diseases has been clearly demonstrated in modern individuals (Lee 2021, Peng 2022, etc.), often through inflammatory pathways. We have now added this to the main text to help illustrate our point.

I'm somewhat concerned about their observation that presence of the Methanobrevibacter cluster correlates with size of the calculus deposit used in the laboratory processing. Could this be related to a complexity issue? Is it clear that small calculus samples harvested from larger calculus deposits would still yield a dominant Methanobrevibacter signal? Did the authors normalize the amount of calcified deposit that went into the laboratory processing? That is not stated in the methods.

We thank the reviewer for this comment and have adjusted the sample information and collection methods accordingly. First, the amount of dental calculus that went into the laboratory processing was not normalized because intact pieces of material are critical for the decontamination process. As stated in the response to Reviewer 2, we agree that calculus size is likely to be a contributing factor in the description of the Methanobrevibacter-associated communities we observe, as this may be emblematic of poor hygiene. As such, we have added or altered text around this topic in the abstract (L46-47), in the text (L172-175), and discussion (L283-287).

Minor comments:

Line 33 – grammatical error with "rising"

Thank you for indicating this error; we have fixed it.

Line 37 – they should state their dataset derives from 201 dental calculus samples, as only 201 passed their specified thresholds for inclusion in their analyses

Thank you for pointing this out. We have adjusted the phrasing to discuss sampling and, in addition, have created Figure S3 to indicate the numbers of samples retained throughout each step of the filtering process.

Line 40 – I suggest "identified" be changed to "identify" to agree with tense used in the preceding sentence

We thank the reviewer for this suggestion and have changed the text accordingly.

Line 67 – grammatical error with "full"

We thank the reviewer for this comment and have changed the text accordingly.

Line 78 – I think it more fitting to replace “Industrialized population” with “communities that pre-date the Industrialized population”, since this more accurately reflects the temporal transect they consider

Thank you for this comment. We have incorporated this phrasing into our text.

Line 81 – “a proven way” is a bit too strong and does not transparently accommodate biases with the method.

We thank the reviewer for this comment and have changed the wording form ‘proven’ to ‘established’.

Line 83 – again, this should be 201, to reflect the number of individuals included in their final dataset.

We thank the reviewer for this comment and have adjusted the number to 183, the number of samples that passed through all of the filtering and authentication steps included in this study.

Lines 86 and 87 – there is a grammatical problem here with the tenses

Thank you for this comment- we have adjusted the tenses.

Line 90 – at first read of the manuscript I was not sure what “oral geography” meant. A quick definition of this term, such as position of a tooth in the dental arcade, would be useful for a non-specialist audience.

We thank the reviewer for this comment and have added a short definition.

Lines 99 – 100 – it is not clear if these 954 taxa were found consistently across all datasets, or if 954 species were identified in total.

We hope we have clarified this process within the main text to indicate that 954 species were identified in total and filtered out of remaining samples for species-level analyses.

Lines 120 – 121 – The statement here implies a temporal relationship in microbial community structure,

but it is made before the temporal correlation is discussed. Perhaps it could be moved to later section of the discussion? Otherwise your subsequent investigation of temporal correlation seems superfluous

We thank the reviewer for their comment. In this section, we describe the *Methanobrevibacter*-associated and the *Streptococcus*-associated communities we observed and discuss how they relate to commonly observed oral microbiome communities in industrialized individuals. While we appreciate that this section and the temporal correlation may pair well, we believe that it is vital to establish these observations at this point in the paper.

Line 139 – the authors state that oral microflora composition is not linked to periodontal disease or other oral pathology, though I wonder if the authors considered only pathology on the tooth from which they derived their calculus sample. I would assume that a large carious lesion or dental abscess on an adjacent tooth could influence microbial structure globally in the mouth

We thank the reviewer for their inquiry and are happy to clarify. Oral pathology was recorded for the entire mouth and not for specific teeth. In addition, we did not intend to imply that the oral microbiota is not linked to oral pathologies; indeed, we found that periodontal disease was significantly linked within the oral-geography controlled London-subset. This section pertains to the associations between the *Methanobrevibacter*-associated and *Streptococcus*-associated communities and the oral pathologies.

Line 163 – the authors should be clear that they investigated the presence of genes associated with these microbial functions

We thank the reviewer for this comment and have clarified the relevant sentence.

Lines 193 – 194 – It should be stated what time periods are represented in their London dataset

We thank the reviewer for their suggestion and have incorporated it into the text.

196 – the authors should be clear that the physiological or cultural factors *they considered* did not correlate with oral flora content

We thank the reviewer for this suggestion and have modified the text accordingly.

Line 207 – some of the skeletal pathological features they mention can be age-related. Did the authors

consider age of the individual in their analyses?

We thank the reviewer for this inquiry and, indeed, this is a topic our group is addressing. Unfortunately, age-recordings for the populations included in this study are extremely broad, thereby constraining paleoepidemiological inquiries. Due to the broad age categorizations and differing skeletal-age estimation methods, we were not able to evaluate age as a confounder for disease-related analyses. This is a current research area in our team to collect better age estimations in this population, but we were unable to do it for this publication due to the physical relocation of the Museum of London collections.

Line 222 – East Smithfield is a plague cemetery, not a post-plague cemetery

We thank the reviewer for this comment and have adjusted the text accordingly.

Lines 231 – 232 – It's confusing for the authors to use the language of "arrival of the Second Plague Pandemic", when they're specifically referring to the distinct interval of the Black Death.

We thank the reviewer for their comment and have added a description of the second plague pandemic as the black death.

Line 240 – This part focuses on the Methanobrevibacter community exclusively, but I suggest the authors reword this to describe the existence of two community structures that are detectable as far back as 2200 BCE, where one of the two declined as of the mid-14th century (if I've properly interpreted their manuscript)

We thank the reviewer for their comment and have included the *Streptococcus*-associated community in this section.

Line 241 – here the authors state that the Methanobrevibacter component diminished only in recent history. How does this accord with their statements implicating the mid-14th century as a period of change in microbial flora? A figure that plots the proportion of this community over time in the population would be really helpful.

We thank the reviewer for their insightful comment and suggestion for an additional figure. We have now added a new supplemental figure (Figure S12) that plots Streptococcus- and *Methanobrevibacter*-associated community composition (as seen in Figure 3) against a fixed axis of time. This figure helps

demonstrate the gradual shift in *Methanobrevibacter* to *streptococcus* community ratios over time. While limited modern samples were included in this study, the *Methanobrevibacter*-dominated community described in this study is not one that is abundant in modern populations.

Decision Letter, first revision:

Subject: Decision on Nature Microbiology manuscript NMICROBIOL-22051270A

Message: 22nd May 2023

*Please ensure you delete the link to your author homepage in this e-mail if you wish to forward it to your co-authors.

Dear Laura,

Thank you for your patience while your manuscript "Ancient dental calculus reveals novel oral microbiome diversity in Great Britain" was under peer-review at Nature Microbiology. It has now been seen by 4 referees, whose expertise and comments you will find at the end of this email. Although they find your work of some potential interest, they have raised a number of concerns that will need to be addressed before we can consider publication of the work in Nature Microbiology.

Thank you for sending us a revision plan - this was very helpful to inform our decision. Please go ahead and address the main concerns as outlined in your plan. Once these additional figures and discussion have been added, we would be happy to look at a revised manuscript.

We are committed to providing a fair and constructive peer-review process. Please do not hesitate to contact us if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

We strongly support public availability of data. Please place the data used in your paper into a public data repository, if one exists, or alternatively, present the data as Source Data or Supplementary Information. If data can only be shared on request, please explain why in your Data Availability Statement, and also in the correspondence with your editor. For some data types, deposition in a public repository is mandatory - more information on our data deposition policies and available repositories can be found at https://www.nature.com/nature-research/editorial-policies/reporting-standards#availability-of-data.

Please include a data availability statement as a separate section after Methods but before references, under the heading "Data Availability". This section should inform readers about the availability of the data used to support the conclusions of your study. This information includes accession codes to public repositories (data banks for protein, DNA or RNA sequences, microarray, proteomics data etc…), references to source data published alongside the paper, unique identifiers such as URLs to data repository entries, or data set DOIs, and any other statement about data availability. At a minimum, you should include the following statement: "The data that support the findings of this study are available from the corresponding author upon request", mentioning any restrictions on availability. If DOIs are provided, we also strongly encourage including these in the Reference list (authors, title, publisher (repository name), identifier, year). For more guidance on how to write this section please see:

http://www.nature.com/authors/policies/data/data-availability-statements-data-citations.pdf

If revising your manuscript:

* Include a "Response to referees" document detailing, point-by-point, how you addressed each referee comment. If no action was taken to address a point, you must provide a compelling argument. This response will be sent back to the referees along with the revised manuscript.

* If you have not done so already we suggest that you begin to revise your manuscript so that it conforms to our Article format instructions at http://www.nature.com/nmicrobiol/info/final-submission. Refer also to any guidelines provided in this letter.

* Include a revised version of any required reporting checklist. It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review. A revised checklist is essential for re-review of the paper.

When submitting the revised version of your manuscript, please pay close attention to our href="https://www.nature.com/nature-portfolio/editorial-policies/image-integrity">Digital Image Integrity Guidelines. and to the following points below:

-- that unprocessed scans are clearly labelled and match the gels and western blots presented in figures.
-- that control panels for gels and western blots are appropriately described as loading on sample processing controls
-- all images in the paper are checked for duplication of panels and for splicing of gel lanes.

Finally, please ensure that you retain unprocessed data and metadata files after publication, ideally archiving data in perpetuity, as these may be requested during the peer review and production process or after publication if any issues arise.

Please use the link below to submit a revised paper: [REDACTED]

Note: This url links to your confidential homepage and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this e-mail to co-authors, please delete this link to your homepage first.

Nature Microbiology is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. This applies to primary research papers only. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit please visit www.springernature.com/orcid.

If you wish to submit a suitably revised manuscript we would hope to receive it within 6 months. If you cannot send it within this time, please let us know. We will be happy to consider your revision, even if a similar study has been accepted for publication at Nature Microbiology or published elsewhere (up to a maximum of 6 months).

In the meantime we hope that you find our referees' comments helpful.

Yours sincerely,

****

Reviewer Comments:

Reviewer #1 (Remarks to the Author):

Overall, the authors have satisfactorily addressed my main concerns and the manuscript is much improved. However, I still have a concern about the lack of explanation on why some samples were removed: The authors provided a new figure (Figure S3) to demonstrate the filtering process for sample selection. However, they did not provide a clear explanation for the criteria used in determining which samples to remove at each step or the rationale behind these decisions. There is some explanation in the Supplement, but what about "DNA extraction", "aDNA signal"?

Reviewer #2 (Remarks to the Author):

The revised version of the manuscript by Gancz et al. is much improved, and I am happy to recommend it for publication. I am satisfied with the responses given by the authors to my comments on the previous version of the manuscript. I think the authors have done a thorough job of bioinformatically authenticating their ancient dental calculus data and removing contaminant sequences. I only have some minor comments that I would like the authors to address prior to this manuscript being accepted:

One of the main takeaways of this study for other ancient oral microbiome researchers is that depending on what analyses they are doing, they need to include oral geography as a variable. This is discussed briefly in the last section of the SI but I think it needs to be more prominently featured in the Main Text.
In the SI, the authors cite the Fagernäs et al. 2022 study as an example of another ancient DNA study that suggests that oral geography affects oral microbiome composition. Actually, the takeaway message of Fagernäs et al. was that a single sample of ancient dental calculus from anywhere in the mouth can be used to represent an individual's oral microbiome, as the main source of variation was the sampled individual. Since a lot of researchers look to the Warinner Group for their dental calculus sampling procedures, they likely do not consider oral geography while sampling and/or analyzing ancient dental calculus data. I think the authors need to address why their findings, and therefore their recommendations, differ from the Fagernäs et al. 2022 study.

Throughout the Main Manuscript: Please run a check for instances of "industrial microbiome". I think it would be best to be consistent in calling them industrialized microbiomes, or non-

industrialized microbiomes, as the case may be.

Line 90: This sentence would read better as – "Although this approach involves a number of complex challenges in recovering gut microbiomes, reconstructing ancient oral microbiomes preserved within calcified dental plaque (i.e., calculus) is an established way of tracing past oral microbial histories."

Line 110: I do not think that the word "bacterium" in Flavobacteriaceae bacterium should be italicized. It is not the species name.

Line 112: I suggest rephrasing this to "… by removing samples whose microbial composition was similar to that of laboratory controls …"

Line 173: Typo in the word "ADONIS". Currently it reads "ADNOIS".

Line 263: The phrasing " …this finding is the first time that an ancient microbiome has been linked to the presence of systemic diseases… " is awkward. Please rephrase.

Line 273: Change to "caused by Yersinia pestis"

Line 305: Remove "as" and change to "… including migration and nutrition"

Line 333: I think the you might have meant "high output kit", not "high sensitivity" here.

Line 370: Remove unnecessary instance of the word "LefSe" occurring before the word "conducted".

Figure 4: Methanobrevibacter and Streptococcus should be italicized.

Supplementary Information:
The Library Preparation section could use more details of the procedure. It is unclear if the libraries were single or dual barcoded/indexed.

My suggestion for future studies would be to include library controls as well as extraction controls.

Specify what % identity thresholds were used in the MALTx and MALTn analyses.

Figure S13-C: Please be consistent and write out both genus and species of all bacteria in the legend.

Figure S18: Title is wrong.

Data Availability:
The Data Deposition Section in the Supplemental Materials mentions that raw FASTQ data are available on the SRA, whereas the Data Availability Section in the Main Manuscript mentions that trimmed and merged reads are available. Which one is it?

Also, some readers who wish to use these data in their own analyses might like to have access to un-merged reads, depending on what analytical program they are using. Personally, I am fine with the authors making only trimmed and merged reads available. However,

perhaps they should consider including a line in the Data Availability section that says unmerged reads might be available upon author request (or something to that effect).

Reviewer #3 (Remarks to the Author):

The manuscript has been much improved, especially in regards to the materials and methods section (what has actually been done). I still have a number of comments/questions.

Consider rewriting lines 49-52 as follows: "After controlling for oral geography and technical biases, we identified multiple distinct oral microbial communities that co-existed in ancient British populations the past millennia, including a composition dominated by i.a. strictly anaerobic Archaea which is no longer found in industrialized societies.

Line 54-56: Name in the following section WHICH community specifically disappeared (just like in the previous section). Don't just say stuff changed.

Line 58: … recognized and are linked …
Line 59 … industrialized populations.

Second, this research places unnecessary responsibilities and obligations on 85 Indigenous communities to participate in microbiome research, where the benefits of these 86 studies may not directly serve Indigenous peoples17.

Truly, this should not be an argument. They never have any responsibility or obligation to participate in the first place. Anybody who does try to place such responsibility or obligations on them should be jailed. Simply, offer them some sort of minor compensation (in whatever form) for something which represents a minor effort for them to provide (if they want to). Typically, if approached in a proper and respectful way (taking local culture/traditions/etc into account), they'll probably be interested and willing to participate (not considering it a burden in any way). The responsibilities and obligations are with the researchers. Forcing people to do things supposedly in the name or because of "science" is what has been giving science a very bad name lately…

First, we only included high quality samples with >100,000 identified sequences22 and
103 more than five phyla.

When selecting which samples to include, I'd first remove contaminant reads (using a multi-layer approach) and then see how many reads are left and then set a minimum (arbitrary) threshold for the number of reads in order to determine which samples to keep. The current approach seems wrong (you'd through away a sample with say 90,000 reads which after screening is left with say 60,000 reads whilst a sample with say 110,000 reads is kept even though it might only be left with say 10,000 reads after contamination removal).

We were conservative and limited the effects of
111 potential laboratory and environmental contaminant DNA by removing samples that were
112 similar to laboratory controls (Figure S2; Table S2) and conservatively filtering
contaminant 113 species identified in environmental and laboratory controls (Figure S4; Table S4).

Again, this does not appear to be the right approach things in my mind. First, samples that

are more similar to laboratory controls should have a much lower number of reads left after screening for contaminant reads after a multi-layer approach. Subsequently, most samples that look like controls will indeed not have enough reads left to meet the (arbitrary) read threshold number (see previous point) and will still be discarded, but some may be left with sufficient genuine reads. Discarding these latter samples is a waste (unless your multi-layer approach is indeed still not very good).

Second, conservatively filtering contaminant species found in environmental and laboratory controls seems like overly conservative (or like a shortcut) as you actually have one more tool at your disposal than most do; DNA damage levels. Sure, most of these signals will have lower DNA damage levels and can be safely removed but some may be genuine, or partially genuine. If you want additional stringency add more layers to your multi-layer approach. Look for batch effects (associate signals with particular lot-number of your DNA isolation kit, PCR kit, plastics used, take note of the sequencing run, etc). You could sequence samples twice in order to figure out sequencing machine contamination (but this is only possibly of value if other ancient DNA studies have been done on/in the same machine/room). Importantly, use known contaminants (based on your own analyses; not literature) to fish out other contaminants (use unfiltered data and cluster them hierarchically to figure out which samples and negative controls (X-axis) cluster together and are associated with which microbial signals (Y-axis). Then, remove some of the most clear culprits (contaminant cluster) and run it again if you still see such another clustering of other (less abundant) contaminants (perhaps from a another source or batch effect). With Excel and some simplistic statistical tools (and/or R-scripts) most contaminants can be visualized and identified. The other advanced tools are all absolutely fine, just also do the basics. In the end, it will be indeed hard to say something about some microbial groups of (very) low abundance and then it is indeed better to err on the side of cautiousness and to indeed remove them if you have any indication that they might be contaminants. Due to their low abundance their removal will not affect results in any significant way. In short, all somewhat abundant (biologically relevant) signals deserve a good look; don't waste too much effort on the long tail of low abundant signals.

Overall, 954 microbial species were identified 128 across all ancient British calculus samples, predominantly spanning Firmicutes, 129 Actinobacteria, Proteobacteria, and Bacteroidetes phyla (Figure S4).

What about Archaea? Maybe the number of Archaea may be limited but the abundance of groups is also very important when talking about compositions.

Figures:
Figure 2: The 3rd dimension is not really adding anything as far as I can see. If it is, plot it separately (against the 1st, 2nd or 4th).

Figure 2B: The color scheme, going from light brown to dark brown, is chosen very poorly. Go from green to red or something; I'd appreciate more visual contrast. Same is true for Figure S7 (maybe sort these form south to north?).

Figure 2C: Use the same coordinates as calculated for 2A and 2B; otherwise 2C has little relatively little us, showing only that things are different without showing how. Even if samples are missing metadata, they can often still provide information in regards to the composition of a microbiome. For example, if you have fecal samples of a cohort of people with IBD and some household matched controls, a lot of value can be obtained by adding a

large cohort of unrelated people from that same country (sick, healthy and everything in between), if you happen to have that data. By doing so, you'll be better be able to show ecological gradients.

Reviewer #4 (Remarks to the Author):

Overall comments:

I thank the authors for resubmission of their manuscript and the care they have taken in composing their response to the four reviewers. The central conclusion the authors draw in this manuscript is the Methanobrevibacter community, which they identify in samples as old as 1000 BCE, and which dominated oral communities in the Medieval and post-Medieval periods through to the beginning of the Industrial revolution, is potentially no longer represented in modern populations. As requested, the authors have produced a time-delineated bar chart to demonstrate the frequency of this community over time (Figure S12). What becomes clear from this figure, however, is the shallow depth of their modern sample. How confident are the authors that this community is indeed unique to pre-Industrial periods? From the period 0 – 500CE it seems they have only one individual out of 9 that has this community: it is, however, absent in the modern cohort, which is smaller in overall sample size (n=6 according to description in the SI). Could stochasticity in the data play a role in its detection in the modern population? This is especially important given that their models show stability in oral microbial community structures in the 2,200-year period considered. This is confirmed by their PCoA in Figure S7A, where I see no genus-level clustering based on time period. Why is the modern population not included in Figure S7A for comparison? While Figure S6 compares ancient samples to modern, this is done only at the species level, and I suspect higher representation of genus-level taxa in the ancient datasets owing to biases in DNA preservation (shorter fragments that will assign to a genus rather than species). Also, why does Figure S6 show 9 datapoints for the modern population when there are only 6 datasets described in the SI (5 plaque and 1 calculus)?

Ideally a representation similar to Figure S12 should be included with just the London population to better visually demonstrate the purported effect of the Black Death, with finer time intervals (if dates for the collections permit). Further to this, demographic change as a result of the mortality of the Black Death is controversial, and this is best explored in the context of their statements regarding the impact of this event on the microbiome. Aside from disease susceptibility, migration patterns to the city of London in the decades following the pandemic, and the growth in the London population as a result of the Industrial revolution, are insufficiently explored as possible reasons for the trends they observe in their data. The authors should cite and discuss the following publications that have surfaced since their original submission:
Gopalarkishnan et al www.sciencedirect.com/science/article/pii/S0960982222014671
Klunk et al doi.org/10.1002/ajpa.23820
Klunk et al www.nature.com/articles/s41586-022-05349-x
Barton et al https://www.biorxiv.org/content/10.1101/2023.03.14.532615v1

There is also no discussion of sequencing depth or library complexity anywhere in the paper or in the supplementary methods or tables that I have been able to find. While the authors are considering proportions of Genus-level read assignments in their identification of microbial communities, I would like to see that their identifications of Methanobrevibacter or Streptococcus communities is not influenced by number of sequences that went into the

analysis (for both ancient and modern samples). While the authors state that they limit their analyses to those datasets with greater than 100,000 reads "identified", clarity here would be appreciated (see my specific comment below).

Greater care is needed in presentation of information in the supplementary tables. Please see my comments below.

Specific comments:

Line 46 and 47- I suggest the authors specify that aDNA was the method of analysis

Line 50 – please change to "after controlling for some technical biases" or equivalent

Line 52 and throughout – "calculus fragment size" can be misinterpreted as average DNA fragment length (at least I was confused by what was meant by fragment size). Please change this to something more intuitive, such as "weight of calculus extracted" or equivalent.

Line 55 – "linked to skeletal pathologies" is far too broad a summary. Linked with what type of skeletal pathologies? Age-related? Infectious disease? Metabolic? Non-specific indicators of physiological stress?

Line 56 – arrival of second pandemic is a strange term to use. Do the authors simply mean the Black Death? If so, just state that.

Line 76 – adjust the misplaced modifier, as the sentence could be misinterpreted as meaning Indigenous people are on the path to extinction.

Line 102-103 – explain what is meant by ">100,000 identified sequences". Do you mean a minimum of 100,000 sequences taxonomically identified via the applied methods? Or taxonomically assigned to a known oral microbe? Are these non-duplicate reads?

Line 104 – is "ChangePoint" a novel program if it is published elsewhere? Perhaps the authors here use it for a novel application?

Line 160 – As the authors here refer to age and sex, "demographic variables" would be more accurate than "physiological variables", which is too broad a term and includes infectious disease.

Line 181 – Figure 2C should be Figure 2B

Line 188 – Here the authors make the observation that greater amounts of calculus in the extraction were correlated with communities dominated by Streptococcus. Their argument is that dental hygiene plays a role in community structure, which may be true. Could it be that Streptococcus creates an environment conducive to the formation of more calculus?

Line 199 – regarding animal origin, I assume human DNA was present. Perhaps "non-host animal origin"

Line 225 – I assume the authors mean "gluconeogenesis"?

Line 242 – "previous Britain-wide analysis" – this could be interpreted as referring to a

different study, so I suggest the authors use "above Britain-wide analysis" or equivalent

248 – I suggest qualifying language here about the connection between skeletal indicators of health and oral microbiome composition

Line 261 – Could modern antibiotic treatment confound this relationship by causing shifts in the microbiome?

Line 264 – not all diseases identified macroscopically in their table 1 are obviously systemic.

Line 270 – the sample number of this reduced London dataset should be stated

Line 290 – the authors' use of the term "Second Plague Pandemic" is confusing. It should be made clear that the Black Death refers to one pandemic event in the mid-14th century, whereas the "second pandemic" refers to the Black Death and subsequent plague outbreaks until the mid-18th century

Supplemental materials:
S2 – Colors for "soil" and "unknown" are flipped in panels A and B, which is confusing. Same phenomenon in Figure S13 between panels A and B

S4 – the samples (along the X axes) seem to be in random order. Specifically Medieval and Post-Medieval are mixed and presented together. Please separate these into intuitive categories and include some time-informative labels along the x-axis.

In general the supplementary tables are extremely difficult to interpret, as they contain columns with cryptic labels and rarely include units. Clarity needs to be improved here. I was also unable to find any legends for the tables, only titles in the table of contents for the SI.

Table S1 is extremely difficult to follow, with too much information and too little explanation. I have looked for the very simple field of "sequencing depth" and I cannot find it. It also includes a column entitled "extraction notes" that can likely only be understood by the authors. Remove passages such as the following, or better explain them:
Tim_in_Env._Rm._Start_of_2nd_day_(Handling:_14785,_14786,_14787,_14788,_14789._Both_he_and_I_cleaned_before_I_started_work)

Table S4: what is the significance of the red highlighting? There is highlighting in other tables that is also left unexplained (at least that I can see).

Author Rebuttal, first revision:

Reviewer Comments:

Reviewer #1 (Remarks to the Author):

Overall, the authors have satisfactorily addressed my main concerns and the manuscript is much improved. However, I still have a concern about the lack of explanation on why some samples were removed: The authors provided a new figure (Figure S3) to demonstrate the filtering process for sample

selection. However, they did not provide a clear explanation for the criteria used in determining which samples to remove at each step or the rationale behind these decisions. There is some explanation in the Supplement, but what about "DNA extraction", "aDNA signal"?

To improve clarity on the filtering issue, we did two things. First, we expanded the Figure S3 caption to add definitions to all the steps. The caption now reads: "**Figure S3:** Ancient DNA and Authentication workflow for this study including the relevant steps and the numbers of samples retained and discarded at each. The procedures were conducted left to right in order. Following sampling, 15 samples were discarded due to a lack of any DNA recovery after DNA extraction (i.e., amplification during the first step of library preparation did not yield measurable DNA quantities). 14 samples were compositionally similar to the extraction blank controls (EBCs) and were therefore discarded (see SI text for details). Another four samples had fewer than 100,000 assigned reads and were therefore discarded to accurately reconstruct diversity. For species-based analyses, we also filtered out species found in the EBCs. Finally, we used MapDamage2.0 and ChangePoint to determine whether samples had DNA-damage patterns as expected in ancient samples; samples without meaningful DNA damage were excluded from downstream analyses." Second, we also added a Key for Table S1, which clearly defines why each sample was included/excluded in each column, which represents different steps in the filtering process (Table S1 Key in SI Tables).

Reviewer #2 (Remarks to the Author):

The revised version of the manuscript by Gancz et al. is much improved, and I am happy to recommend it for publication. I am satisfied with the responses given by the authors to my comments on the previous version of the manuscript. I think the authors have done a thorough job of bioinformatically authenticating their ancient dental calculus data and removing contaminant sequences. I only have some minor comments that I would like the authors to address prior to this manuscript being accepted:

One of the main takeaways of this study for other ancient oral microbiome researchers is that depending on what analyses they are doing, they need to include oral geography as a variable. This is discussed briefly in the last section of the SI but I think it needs to be more prominently featured in the Main Text.

In the SI, the authors cite the Fagernäs et al. 2022 study as an example of another ancient DNA study that suggests that oral geography affects oral microbiome composition. Actually, the takeaway message of Fagernäs et al. was that a single sample of ancient dental calculus from anywhere in the mouth can be

used to represent an individual's oral microbiome, as the main source of variation was the sampled individual. Since a lot of researchers look to the Warinner Group for their dental calculus sampling procedures, they likely do not consider oral geography while sampling and/or analyzing ancient dental calculus data. I think the authors need to address why their findings, and therefore their recommendations, differ from the Fagernäs et al. 2022 study.

We thank the reviewer for this point, and indeed, the ethical and analytical benefits of controlling for oral geography are a major facet of this paper. We have now expanded upon this topic in the main text (L126-130) "As these biases likely reflect biological and ecological differences in the mouth[25], these differences attributed to oral geography raise questions about the interpretations of previous paleomicrobiome studies that used calculus samples collected from a mixed dentition[29–31] and suggest that future studies should control for oral geography during sampling and analysis (Figure S18)." We now cite several of the Warinner papers that use this approach, drawing specific attention to this issue.

We also think it's important to directly discuss these implications, as we do in the SI on pages 26-27. The published Warinner Group's dental calculus sampling approach (e.g., sampling all the calculus material from an individual) raises sustainability and ethical questions, so we include that in the SI to provide an alternative perspective.

Throughout the Main Manuscript: Please run a check for instances of "industrial microbiome". I think it would be best to be consistent in calling them industrialized microbiomes, or non-industrialized microbiomes, as the case may be.

We thank the reviewer for this comment and have modified the main text accordingly.

Line 90: This sentence would read better as – "Although this approach involves a number of complex challenges in recovering gut microbiomes, reconstructing ancient oral microbiomes preserved within calcified dental plaque (i.e., calculus) is an established way of tracing past oral microbial histories."

This has been implemented.

Line 110: I do not think that the word "bacterium" in Flavobacteriaceae bacterium should be italicized. It is not the species name.

Thank you for catching this oversight.

Line 112: I suggest rephrasing this to "… by removing samples whose microbial composition was similar to that of laboratory controls …"

Fixed, as suggested.

Line 173: Typo in the word "ADONIS". Currently it reads "ADNOIS".

Thank you for catching this typo; this has been fixed.

Line 263: The phrasing " …this finding is the first time that an ancient microbiome has been linked to the presence of systemic diseases… " is awkward. Please rephrase.

This has been rephrased to "this is the first time that ancient microbiomes have been linked to systemic diseases that manifest in the skeleton".

Line 273: Change to "caused by Yersinia pestis"

We modified the main text accordingly.

Line 305: Remove "as" and change to "… including migration and nutrition"

This has been done. Thank you!

Line 333: I think the you might have meant "high output kit", not "high sensitivity" here.

Correct. Thank you!

Line 370: Remove unnecessary instance of the word "LefSe" occurring before the word "conducted".

Done.

Figure 4: Methanobrevibacter and Streptococcus should be italicized.

Thank you! Done.

Supplementary Information:
The Library Preparation section could use more details of the procedure. It is unclear if the libraries were single or dual barcoded/indexed.

We thank the reviewer for this comment and have clarified that the libraries were dual barcoded (i.e., internal barcodes) and single indexed (i.e., Illumina indexes).

My suggestion for future studies would be to include library controls as well as extraction controls.

This was done but the data was poor and not included, due to low sequence recovery of these controls, as expected. We have now included this and made this clearer in the SI text (pg. 10).

Specify what % identity thresholds were used in the MALTx and MALTn analyses.

0.8 was used. This has now been included in the SI text.

Figure S13c: Please be consistent and write out both genus and species of all bacteria in the legend.

We thank the reviewer for this comment and have made the proper revisions.

Figure S18: Title is wrong.

This is now corrected.

Data Availability:

The Data Deposition Section in the Supplemental Materials mentions that raw FASTQ data are available on the SRA, whereas the Data Availability Section in the Main Manuscript mentions that trimmed and merged reads are available. Which one is it?

We thank the reviewer for this comment! The trimmed and merged fastq files are the ones that have been uploaded. We have modified the text accordingly.

Also, some readers who wish to use these data in their own analyses might like to have access to un-merged reads, depending on what analytical program they are using. Personally, I am fine with the authors making only trimmed and merged reads available. However, perhaps they should consider including a line in the Data Availability section that says unmerged reads might be available upon author request (or something to that effect).

We thank the reviewer and have added this option in both the SI and main text.

Reviewer #3 (Remarks to the Author):

The manuscript has been much improved, especially in regards to the materials and methods section (what has actually been done). I still have a number of comments/questions.

Consider rewriting lines 49-52 as follows: “After controlling for oral geography and technical biases, we identified multiple distinct oral microbial communities that co-existed in ancient British populations the past millennia, including a composition dominated by i.a. strictly anaerobic Archaea which is no longer found in industrialized societies.

Thank you for this suggestion! We have reworded this sentence as follows: “After controlling for oral geography and technical biases, we identified multiple distinct oral microbial communities that co-existed in ancient British populations for millennia, including a community associated with Methanobrevibacter – a strictly anaerobic Archaea – that is not well-described in current industrialized societies.”

Line 54-56: Name in the following section WHICH community specifically disappeared (just like in the previous section). Don’t just say stuff changed.

Thank you for this suggestion. We have now modified the article summary to clearly state which communities have shifted (e.g., *Methanobrevibacter*-associated).

Line 58: … recognized and are linked …

Thank you; we have implemented these changes!

Line 59 … industrialized populations.

Thank you, done!

Second, this research places unnecessary responsibilities and obligations on Indigenous communities to participate in microbiome research, where the benefits of these studies may not directly serve Indigenous peoples1.

Truly, this should not be an argument. They never have any responsibility or obligation to participate in the first place. Anybody who does try to place such responsibility or obligations on them should be jailed. Simply, offer them some sort of minor compensation (in whatever form) for something which represents a minor effort for them to provide (if they want to). Typically, if approached in a proper and respectful way (taking local culture/traditions/etc into account), they'll probably be interested and willing to participate (not considering it a burden in any way). The responsibilities and obligations are with the researchers. Forcing people to do things supposedly in the name or because of "science" is what has been giving science a very bad name lately…

We thank the reviewer for their comment, and we agree that we wish comments like this were not needed in the literature. Indeed, our research team is strongly against taking advantage of any population in the name of science, and so we feel this is important to make our stance clear.

First, we only included high quality samples with >100,000 identified sequences and
103 more than five phyla.

When selecting which samples to include, I'd first remove contaminant reads (using a multi-layer approach) and then see how many reads are left and then set a minimum (arbitrary) threshold for the number of reads in order to determine which samples to keep. The current approach seems wrong (you'd through away a sample with say 90,000 reads which after screening is left with say 60,000 reads

whilst a sample with say 110,000 reads is kept even though it might only be left with say 10,000 reads after contamination removal).

We thank the reviewer for their suggestion and appreciate the strengths associated with the reviewer's proposed approach. The reason we filter samples as an initial step is to ensure that we have an adequate signal to accurately assess contamination within the sample. As described in our supplementary materials, the samples removed via this filtering procedure typically had a low number of reads, suggesting that true biological signatures were not obtained. However, we now include the final number of sequences (i.e., the number of assigned sequences after any contaminant filtering) in Supplementary Table 3 (row D). As shown, the lowest number of assigned sequences for any biological sample post-filtering was 126,912 (Kirkhiill15924). As such, the retained samples are likely to represent a more accurate taxonomic reconstruction than samples with lower numbers of reads (Hillmann, 2016). While our approach is on the more conservative side, caution is critical when reconstructing ancient microbial signals.

We were conservative and limited the effects of
111 potential laboratory and environmental contaminant DNA by removing samples that were 112 similar to laboratory controls (Figure S2; Table S2) and conservatively filtering contaminant 113 species identified in environmental and laboratory controls (Figure S4; Table S4).

Again, this does not appear to be the right approach things in my mind. First, samples that are more similar to laboratory controls should have a much lower number of reads left after screening for contaminant reads after a multi-layer approach. Subsequently, most samples that look like controls will indeed not have enough reads left to meet the (arbitrary) read threshold number (see previous point) and will still be discarded, but some may be left with sufficient genuine reads. Discarding these latter samples is a waste (unless your multi-layer approach is indeed still not very good).

While we appreciate the suggested approach, we take a cautious, multi-tiered approach to ensure that our sample retention is robust. Our primary goal is to reconstruct **oral** microbiome signals to compare across individuals – the goal is not to examine the total (including environmental) microbial signatures in all samples that contribute to variable preservation across a site, although these are very interesting and worthy hypotheses. Our approach (i.e., assessing similarities to EBCs before contaminant taxa filtration) allows us to assess if signatures in the sample are more environmental or oral in nature as a first step before proceeding with filtration. Even if a sample is sequenced deeply (i.e., >100K sequences), these could be environmental. If a sample only has a handful of oral microbes and is otherwise more dominated by contaminant signatures (i.e., its signal makes it more similar to controls than other

biological samples), we infer that the sample is likely to be more degraded is therefore unlikely to have maintained all of oral microbiome signature. In other words, we cannot in good conscious analyze a sample as oral if it is dominated by contaminants, regardless of sequencing depth. It is possible that a sample that contains a robust oral signature only has a handful of sequences, and therefore, would need to scrutinized differently; this method allow us to assess that possibility. Further research on this issue is warranted though, which could include detailed assessments of each individual per site and multi-species DNA damage profiles against species diversity and environmental conditions that dictate why some individuals preserve more or less oral microbial signatures than others. However, this is currently out of the scope of the study.

Nevertheless, this filtering step only resulted in removing only 13 of the 235 samples we assessed, and only two of our 23 sites had more than one sample removed because of this process (Breedon on the Hill and Spital Square each had two samples removed; 'EBC_Prox' in Column 'Used' in Table S1). While we think this filtering is critical to accurately compare robust oral microbiomes across individuals, it truly had a minimal impact on this data set. Further, we are transparent about the species present in our control samples (Table S4), and the dominant taxa and associated oral species (*Methanobrevibacter* and *Streptococcus* and their related taxa) that serve as the primary basis of this paper's conclusions were not in the controls, again suggesting that clustering with controls is representative of poor sample preservation and should result in the exclusion of those samples.

Second, conservatively filtering contaminant species found in environmental and laboratory controls seems like overly conservative (or like a shortcut) as you actually have one more tool at your disposal than most do; DNA damage levels. Sure, most of these signals will have lower DNA damage levels and can be safely removed but some may be genuine, or partially genuine. If you want additional stringency add more layers to your multi-layer approach. Look for batch effects (associate signals with particular lot-number of your DNA isolation kit, PCR kit, plastics used, take note of the sequencing run, etc). You could sequence samples twice in order to figure out sequencing machine contamination (but this is only possibly of value if other ancient DNA studies have been done on/in the same machine/room). Importantly, use known contaminants (based on your own analyses; not literature) to fish out other contaminants (use unfiltered data and cluster them hierarchically to figure out which samples and negative controls (X-axis) cluster together and are associated with which microbial signals (Y-axis). Then, remove some of the most clear culprits (contaminant cluster) and run it again if you still see such another clustering of other (less abundant) contaminants (perhaps from a another source or batch effect). With Excel and some simplistic statistical tools (and/or R-scripts) most contaminants can be visualized and identified. The other advanced tools are all absolutely fine, just also do the basics. In the end, it will be indeed hard to say something about some microbial groups of (very) low abundance and then it is indeed better to err on the side of cautiousness and to indeed remove them if you have any

indication that they might be contaminants. Due to their low abundance their removal will not affect results in any significant way. In short, all somewhat abundant (biologically relevant) signals deserve a good look; don't waste too much effort on the long tail of low abundant signals.

We appreciate this reviewer's knowledge and expertise in this area. Indeed, we have considered and implemented several of these items in the original analyses. For example, we already tested for the impacts of batch effects on this data, as suggested (i.e., sampling date, DNA extraction ID (which incorporated date and kit number), technician, library prep ID, and sequencing run; Table S7-S9) and found no significant contributions. We also think it is important to highlight the fact that we used our own controls (not published contaminant profiles) that were created alongside our samples in every preparation to assess contamination during the wet lab work to assess contributions of contaminant profiles for this publication, again as the reviewer suggests. Lastly, we appreciate that other methods, such as hierarchical ranking and abundance-based methods of assessing contaminant exist (i.e., decontam), and we regularly employ these in much of our research. However, as our primary goal in this work was to reconstruct **oral** microbiomes as accurately as possible and include as many rare taxa not simply described in the modern populations, we chose a conservative process that was likely to remove as many possible contaminant taxa as possible. In the end, we chose this conservative approach after a very stringent analysis and examination of the species and abundances within our controls and after drastic improvements in laboratory strategies to reduce contaminant DNA in our wet lab work. We completely agree that assessments of all parts of the data are useful.

To further describe our contaminant signals, we now include an assessment using decontam to characterize and filter contamination. Our decontam results, run on the species level, indicated a total of 321 contaminant species (Table S4, Column C); approximately 52.7% of those labeled as contaminants using our more conservative EBC-species filtering method were also identified using decontam. The decontam method clearly misses species that are likely to be contaminants; for example, decontam identifies only 22 species of *Acinetobacter*, while the conservative approach identified 25 – a genera that is very well known to be a contaminant in oral studies. No one approach is perfect.

While we proceed with the original filtration method due to its more conservative approach, a comparison of these two different approaches is now included in the supplementary materials (Table S4) and text (SI page 11).

Overall, 954 microbial species were identified 128 across all ancient British calculus samples, predominantly spanning Firmicutes, Actinobacteria, Proteobacteria, and Bacteroidetes phyla (Figure S4).

What about Archaea? Maybe the number of Archaea may be limited but the abundance of groups is also very important when talking about compositions.

We are unsure of the reviewer's suggestion here, as archaea (e.g., Euryarchaeota) are included in Figure S4. This line lists the four dominant phyla, instead of trying to describe all phyla present.

Figures:
Figure 2: The 3rd dimension is not really adding anything as far as I can see. If it is, plot it separately (against the 1st, 2nd or 4th).

The third axis shows less variation than 1 or 2, but still shows more dimensionality than axis 4 or 5. While we thank the reviewer for their expertise, we have chosen to leave the plot as it is as we believe it to be more informative.

Figure 2B: The color scheme, going from light brown to dark brown, is chosen very poorly. Go from green to red or something; I'd appreciate more visual contrast. Same is true for Figure S7 (maybe sort these from south to north?).

These colors were chosen to be friendly to color blind individuals. We have chosen to leave them in that color scheme.

Figure 2C: Use the same coordinates as calculated for 2A and 2B; otherwise 2C has little relatively little us, showing only that things are different without showing how. Even if samples are missing metadata, they can often still provide information in regards to the composition of a microbiome. For example, if you have fecal samples of a cohort of people with IBD and some household matched controls, a lot of value can be obtained by adding a large cohort of unrelated people from that same country (sick, healthy and everything in between), if you happen to have that data. By doing so, you'll be better be able to show ecological gradients.

In the revisions of Figures 2A and 2B, we apologize that information clarifying this was accidently lost. In this plot, we have used a different set of samples (i.e., only molars in 2C, while its all teeth in 2A and 2B), so the PCoA was regenerated based on those samples included and therefore cannot use the same coordinates. This information is now included back into the Figure 2 legend.

Reviewer #4 (Remarks to the Author):

Overall comments:

I thank the authors for resubmission of their manuscript and the care they have taken in composing their response to the four reviewers. The central conclusion the authors draw in this manuscript is the Methanobrevibacter community, which they identify in samples as old as 1000 BCE, and which dominated oral communities in the Medieval and post-Medieval periods through to the beginning of the Industrial revolution, is potentially no longer represented in modern populations. As requested, the authors have produced a time-delineated bar chart to demonstrate the frequency of this community over time (Figure S12). What becomes clear from this figure, however, is the shallow depth of their modern sample. How confident are the authors that this community is indeed unique to pre-Industrial periods? From the period 0 – 500CE it seems they have only one individual out of 9 that has this community: it is, however, absent in the modern cohort, which is smaller in overall sample size (n=6 according to description in the SI). Could stochasticity in the data play a role in its detection in the modern population? This is especially important given that their models show stability in oral microbial community structures in the 2,200-year period considered. This is confirmed by their PCoA in Figure S7A, where I see no genus-level clustering based on time period. Why is the modern population not included in Figure S7A for comparison? While Figure S6 compares ancient samples to modern, this is done only at the species level, and I suspect higher representation of genus-level taxa in the ancient datasets owing to biases in DNA preservation (shorter fragments that will assign to a genus rather than species). Also, why does Figure S6 show 9 datapoints for the modern population when there are only 6 datasets described in the SI (5 plaque and 1 calculus)?

We appreciate this comment. We do not want to rule out the possibility that *Methanobrevibacter* communities never occur in modern populations, as these oral archaea are reported in the mouth of people with severe periodontal disease (as we report in the text, pg. 7) as well as smokers (Grine, et al. Scientific Reports, 2018). Nevertheless, there is a significant lack of current literature examining modern dental calculus, especially from Indigenous communities, and most of the modern calculus microbiome work has been done using 16S rRNA approaches, which are inappropriate to compare to shotgun methods used for ancient samples (i.e., see Zeismer, et al. 2015). A very recent study published by our team on modern calculus (i.e., 16S rRNA of 52 individuals) did not detect any *Methanobrevibacter* species from either Indigenous and non-Indigenous modern Australian dental calculus from people with periodontal disease or without (Handsley-Davis, et al., EMPH, 2022), using the 16S rRNA V4 primers that we know will detect this archaea. This finding supports our conclusions in this paper and further suggests that the *Methanobrevibacter*-associated community is less likely to be found in modern populations.

Further, associations between disease/smoking and the presence of *Methanobrevibacter* in modern populations is in stark contrast to how we observed this microbe in pre-Industrial populations. First, we observe this community in people without periodontal disease; for example, *Methanobrevibacter* was observed in 39 people (21%) in our study and was not associated with periodontal disease, although it is difficult to rule out any effects of smoking in skeletal material. Methanogens have also been identified in nearly every ancient calculus microbiome study published to date across Europe and Asia, suggesting that this genus and its associated species are common in the calculus of past individuals (i.e., Weyrich et al, Nature, 2017; Mann, et al. Scientific Reports, 2020; Granehall, et al., Microbiome 2021, etc.). This certainly is not the case in modern oral microbiome studies on dental plaque.

While several studies are underway to better characterize modern calculus samples and better explore if methanogens are widely present in the modern mouth, we do not believe we should ignore the mounting evidence that there is a clear change in archaeal biodiversity from pre-Industrial to modern populations with the data we have in hand. However, we do want to ensure that our claims are clear and not misleading. As a result, we further clarified this point in the main text (P6L157; P10L214) and provided a more detailed perspective on the current literature in the SI discussion section for clarity (pg 26).

As for Figure S6 and S7, we only included ancient samples from Table S7, because we wanted to be able to illustrate the breadth of ancient composition over time, even as exists within *Methanobrevibacter* dominated communities (avoiding the clustering shown in Figure S6, for example). We included 9 published modern calculus samples throughout the whole study, and we apologize for the oversight of missing one sample in Table S1 as it was being prepared. This has now been fixed. We also now include a genus-level PCoA plot similar to Figure S6A that includes modern samples (Figure S6B); the modern samples still clearly cluster with the *Streptococcus*-associated communities.

Ideally a representation similar to Figure S12 should be included with just the London population to better visually demonstrate the purported effect of the Black Death, with finer time intervals (if dates for the collections permit). Further to this, demographic change as a result of the mortality of the Black Death is controversial, and this is best explored in the context of their statements regarding the impact of this event on the microbiome. Aside from disease susceptibility, migration patterns to the city of London in the decades following the pandemic, and the growth in the London population as a result of the Industrial revolution, are insufficiently explored as possible reasons for the trends they observe in their data. The authors should cite and discuss the following publications that have surfaced since their

original submission:
Gopalarkishnan et al www.sciencedirect.com/science/article/pii/S0960982222014671
Klunk et al doi.org/10.1002/ajpa.23820
Klunk et al www.nature.com/articles/s41586-022-05349-x
Barton et al https://www.biorxiv.org/content/10.1101/2023.03.14.532615v1

This is a great suggestion, and we now include a Figure S12B with only the London samples at the finest scale possible. We also added more discussion on this topic to the SI (pg. 25) and now include all of the suggested references (SI refs. 60-63).

There is also no discussion of sequencing depth or library complexity anywhere in the paper or in the supplementary methods or tables that I have been able to find. While the authors are considering proportions of Genus-level read assignments in their identification of microbial communities, I would like to see that their identifications of Methanobrevibacter or Streptococcus communities is not influenced by number of sequences that went into the analysis (for both ancient and modern samples). While the authors state that they limit their analyses to those datasets with greater than 100,000 reads "identified", clarity here would be appreciated (see my specific comment below).

We thank the reviewer for this suggestion. First, we define 'identified' as 'assigned', as up to 50% of ancient DNA sequences are not identifiable to a particular bacterial taxa (Eisenhofer and Weyrich, PeerJ, 2019). We now define this in the SI (pg. 11). Next, we generated a density plot that displays the number of total post-filtering, genus-assigned reads for each of the *Methanobrevibacter* and *Streptococcus* dominated samples (Figure S19; SI pg. 25). As shown in Figure S19, the *Methanobrevibacter* and *Streptococcus*-associated samples exhibited significantly different distributions of assigned sequences (t = -4.9244, df = 60.609, p-value = 6.85e-06), with shorter DNA fragments observed in the *Methanobrevibacter*-associated communities. While more research is required to fully understand these patterns, we hypothesize that the reason for this is the differences between the GC content of the dominant microbes. For example, *Methanobrevibacter* genomes have much lower GC content than *Streptococcus* species, which impacts the recovery of their DNA (Ziesemer, et al. Sci Rep, 2015). We would therefore expect the read lengths of *Methanobrevibacter*-dominated communities to be shorter, as observed. Nevertheless, we now include a description of this observation in the SI (pg. 25).

Greater care is needed in presentation of information in the supplementary tables. Please see my comments below.

We appreciate this comment. We read over the SI in great detail and have made several minor corrections in the SI.

Specific comments:

Line 46 and 47- I suggest the authors specify that aDNA was the method of analysis

This is now amended.

Line 50 – please change to "after controlling for some technical biases" or equivalent

'Select technical biases' is now used in this line.

Line 52 and throughout – "calculus fragment size" can be misinterpreted as average DNA fragment length (at least I was confused by what was meant by fragment size). Please change this to something more intuitive, such as "weight of calculus extracted" or equivalent.

It is now described as 'calculus sample size' through the main text, SI, and SI tables.

Line 55 – "linked to skeletal pathologies" is far too broad a summary. Linked with what type of skeletal pathologies? Age-related? Infectious disease? Metabolic? Non-specific indicators of physiological stress?

We thank the reviewer for this comment, and we have now added some of the information from the main body of the text into the abstract, specifically defining the skeletal pathologies as skeletal markers of systemic diseases including periostitis and joint pathologies. This point is elaborated upon in the main text (pg 9).

Line 56 – arrival of second pandemic is a strange term to use. Do the authors simply mean the Black Death? If so, just state that.

The Black Death has now been described as a derogatory term to describe this pandemic. As such, *Y. pestis* pandemics have now been referred to as their historical event names.

Line 76 – adjust the misplaced modifier, as the sentence could be misinterpreted as meaning Indigenous people are on the path to extinction.

Thank you for this comment; we have adjusted the sentence.

Line 102-103 – explain what is meant by ">100,000 identified sequences". Do you mean a minimum of 100,000 sequences taxonomically identified via the applied methods? Or taxonomically assigned to a known oral microbe? Are these non-duplicate reads?

We thank the reviewer for their inquiry and have elaborated upon this point in the SI. To clarify, we mean a minimum of 100,000 taxonomically identified sequences.

Line 104 – is "ChangePoint" a novel program if it is published elsewhere? Perhaps the authors here use it for a novel application?

We thank the reviewer for this comment! ChangePoint is a program originally created by Yichen Liu in her dissertation and is set to be published shortly. We hope to be able to cite that publication by the time this paper is published; for now, we cite her dissertation, which describes the algorithm and approach used in ChangePoint.

Line 160 – As the authors here refer to age and sex, "demographic variables" would be more accurate than "physiological variables", which is too broad a term and includes infectious disease.

We have adjusted the text accordingly.

Line 181 – Figure 2C should be Figure 2B

We thank the reviewer for catching this oversight.

Line 188 – Here the authors make the observation that greater amounts of calculus in the extraction were correlated with communities dominated by Streptococcus. Their argument is that dental hygiene plays a role in community structure, which may be true. Could it be that Streptococcus creates an

environment conducive to the formation of more calculus?

We thank the reviewer for this comment. This is absolutely a possibility. The role of different microbes in calculus deposits and their size is indeed something our team is an active area of research in our team.

Line 199 – regarding animal origin, I assume human DNA was present. Perhaps “non-host animal origin”

We have changed the text accordingly.

Line 225 – I assume the authors mean “gluconeogenesis”?

That is indeed what was meant; thank you for catching the typo!

Line 242 – “previous Britain-wide analysis” – this could be interpreted as referring to a different study, so I suggest the authors use “above Britain-wide analysis” or equivalent

That is an excellent point, and we have changed the text accordingly.

248 – I suggest qualifying language here about the connection between skeletal indicators of health and oral microbiome composition

We thank the reviewer for this comment and have ensured qualifying language is present.

Line 261 – Could modern antibiotic treatment confound this relationship by causing shifts in the microbiome?

Published oral microbiome research has shown that plaque microbiomes (the precursors to calculus) are minimally impacted by antibiotics treatment (Zaura, et al. mBio, 2015). This is likely due to the biofilm matrix that encapsulates these microbes. However, it is not possible to rule this out over longer time periods, so we now include this discussion point in the SI (pg 26).

Line 264 – not all diseases identified macroscopically in their table 1 are obviously systemic.

We thank the reviewer for this comment and have adjusted the wording to more closely match our intended meaning.

Line 270 – the sample number of this reduced London dataset should be stated

We thank the reviewer for this comment; this number is given above in line 234 ("we further narrowed our analysis to 127 medieval and post-medieval individuals from London").

Line 290 – the authors' use of the term "Second Plague Pandemic" is confusing. It should be made clear that the Black Death refers to one pandemic event in the mid-14th century, whereas the "second pandemic" refers to the Black Death and subsequent plague outbreaks until the mid-18th century

The Black Death has now been described as a derogatory term to describe this pandemic. As such, *Y. pestis* pandemics have now been referred to as their historical event names.

Supplemental materials:
S2 – Colors for "soil" and "unknown" are flipped in panels A and B, which is confusing. Same phenomenon in Figure S13 between panels A and B

We have altered these figures accordingly.

S4 – the samples (along the X axes) seem to be in random order. Specifically Medieval and Post-Medieval are mixed and presented together. Please separate these into intuitive categories and include some time-informative labels along the x-axis.

We thank the reviewer for this suggestion and have ordered the samples by the earliest date when possible.

In general the supplementary tables are extremely difficult to interpret, as they contain columns with cryptic labels and rarely include units. Clarity needs to be improved here. I was also unable to find any legends for the tables, only titles in the table of contents for the SI.

Table S1 is extremely difficult to follow, with too much information and too little explanation. I have looked for the very simple field of “sequencing depth” and I cannot find it. It also includes a column entitled “extraction notes” that can likely only be understood by the authors. Remove passages such as the following, or better explain them:
Tim_in_Env._Rm._Start_of_2nd_day_(Handling:_14785,_14786,_14787,_14788,_14789._Both_he_and_I_cleaned_before_I_started_work)

We thank the reviewer for this suggestion. We now provide a key to Table S1, where each header is defined and more clearly explained. We also include sequencing depth both before and after filtering in Table S3.

Table S4: what is the significance of the red highlighting? There is highlighting in other tables that is also left unexplained (at least that I can see).

We also now describe any highlighting in SI tables at the top of each table.

Decision Letter, second revision:

Subject: Your manuscript, NMICROBIOL-22051270B
Message: Our ref: NMICROBIOL-22051270B

24th August 2023

Dear Laura,

Thank you for your patience as we’ve prepared the guidelines for final submission of your Nature Microbiology manuscript, "Ancient dental calculus reveals novel oral microbiome diversity in Great Britain" (NMICROBIOL-22051270B). Please carefully follow the step-by-step instructions provided in the attached file, and add a response in each row of the table to indicate the changes that you have made. Ensuring that each point is addressed will help to ensure that your revised manuscript can be swiftly handed over to our production team.

We would like to start working on your revised paper, with all of the requested files and forms, as soon as possible (preferably within two weeks). Please get in contact with us if you anticipate delays.

When you upload your final materials, please include a point-by-point response to any remaining reviewer comments.

If you have not done so already, please alert us to any related manuscripts from your

group that are under consideration or in press at other journals, or are being written up for submission to other journals (see: https://www.nature.com/nature-research/editorial-policies/plagiarism#policy-on-duplicate-publication for details).

In recognition of the time and expertise our reviewers provide to Nature Microbiology's editorial process, we would like to formally acknowledge their contribution to the external peer review of your manuscript entitled "Ancient dental calculus reveals novel oral microbiome diversity in Great Britain". For those reviewers who give their assent, we will be publishing their names alongside the published article.

Nature Microbiology offers a Transparent Peer Review option for new original research manuscripts submitted after December 1st, 2019. As part of this initiative, we encourage our authors to support increased transparency into the peer review process by agreeing to have the reviewer comments, author rebuttal letters, and editorial decision letters published as a Supplementary item. When you submit your final files please clearly state in your cover letter whether or not you would like to participate in this initiative. Please note that failure to state your preference will result in delays in accepting your manuscript for publication.

Cover suggestions

COVER ARTWORK: We welcome submissions of artwork for consideration for our cover. For more information, please see our guide for cover artwork.

Nature Microbiology has now transitioned to a unified Rights Collection system which will allow our Author Services team to quickly and easily collect the rights and permissions required to publish your work. Approximately 10 days after your paper is formally accepted, you will receive an email in providing you with a link to complete the grant of rights. If your paper is eligible for Open Access, our Author Services team will also be in touch regarding any additional information that may be required to arrange payment for your article.

Please note that you will not receive your proofs until the publishing agreement has been received through our system.

Please note that Nature Microbiology is a Transformative Journal (TJ). Authors may publish their research with us through the traditional subscription access route or make their paper immediately open access through payment of an article-processing charge (APC). Authors will not be required to make a final decision about access to their article until it has been accepted. Find out more about Transformative Journals

Authors may need to take specific actions to achieve compliance with funder and institutional open access mandates. If your research is supported by a funder that requires immediate open access (e.g. according to Plan S principles) then you should select the gold OA route, and we will

direct you to the compliant route where possible. For authors selecting the subscription publication route, the journal's standard licensing terms will need to be accepted, including self-archiving policies. Those licensing terms will supersede any other terms that the author or any third party may assert apply to any version of the manuscript.

For information regarding our different publishing models please see our Transformative Journals page. If you have any questions about costs, Open Access requirements, or our legal forms, please contact ASJournals@springernature.com.

Please use the following link for uploading these materials: [REDACTED]

If you have any further questions, please feel free to contact me.

Best regards,

Reviewer #2:
Remarks to the Author:
I am satisfied with the responses given by the authors to my comments, as well as those of the other reviewers, on the previous version of the manuscript. I think the current version is much more stream-lined, and the over-interpretation of some results in previous versions has also been addressed. This study will add to the growing compendium of ancient dental calculus metagenomes, but more importantly, it will highlight the importance of considering oral geography while sampling dental calculus as well as a novel method of estimating diet from functional analysis of ancient oral metagenomes, both of which will be of interest to ancient microbiome researchers. I am happy to recommend this manuscript for publication and look forward to seeing it published soon.

My only comments on this version are related to the following edits:
Line 61 - Change "its disappeareance is" to "their disappearance is".
Line 172 - Capitalize "Second Plague Pandemic" as you have done in the rest of the manuscript. Also, it would be better to say here that this is also known as the Black Death, since this is the first mention of it in the main text, rather than in Line 282.
Line 269 - Typo. It should be "Peptoniphilus".
Line 269 - Change to "while not all systemic diseases..."
Line 322 - Typo. Change "sampled" to "samples".

Reviewer #3:
Remarks to the Author:
The authors have satisfactorily addressed my main concerns and I am happy to

recommend it for publication. I just added some thoughts below in regards to some of the answers given to my previous questions.

In regards to the comment about the number of reads:
While I might be able to think of some reasons why my approach might be more correct the discussion has been made rather irrelevant by providing the number of reads per sample. The chosen threshold seems rather sensible (I'd pick the same) and the order of operations would not have changed the selection of samples. In regards to contamination removal strategies involving multiple filtering steps; I typically keep all contaminant reads until the final step and then clean up the data. In any case, as you stated below, no approach with these kinds of samples is probably perfect, but I'm satisfied with this reply/outcome.

In regards to the contamination fraction comment:
In several of my own medium/low biomass studies (non-ancient) contamination removal was quite straightforward. Even if a significant fraction of reads was contamination the remaining fraction in these samples clearly represented a microbial profile (after contamination removal) as could be expected (similar to samples with a low contamination fraction). However, a very good point is made in regards to ancient DNA samples that a high amount of contamination reads might indeed be indicative of increased degradation and hence of a lower quality oral composition representation after contamination removal (in addition, degradation might not affect all DNA from all species equally hence skewing results even more). It might be worth it to mention this line of reasoning in the contamination removal section.

In regards to my comment regarding various contamination removal strategies:
I personally don't even use decontam myself and I'm very aware of its limitations. Commonly, people use decontam and then claim that they "solved" the contamination issue. This is especially problematic in very-low to no biomass samples, as the remaining signals are then presented as the "microbiome" of those samples. In any case, perfection is not needed for these samples (as it is clear that they had a microbiome) and it is now clear to me that sufficient effort has been put into getting the results to properly represent the ancient oral signatures (and importantly, that these efforts have become part of the manuscript instead).

Reviewer #4:
Remarks to the Author:
I thank the authors for the latest revised version of their work. While I appreciate the lengthy responses to my queries from the last round of review, I recommend more complete discussion of some concepts in the main manuscript. This most importantly applies to their disclosure that Methanobrevibacter is found in modern smokers and those with advanced periodontal disease. As the novelty of the ancient Methanobrevibacter community is a central conclusion of their manuscript, the authors are best advised to more transparently describe its novelty. While they do report on lines 181-183 that Methanobrevibacter is associated with severe periodontal disease in modern populations, they do not mention its association with smoking (as disclosed in their reply to my query). These modern contexts for Methanobrevibacter questions their core conclusion that "this Methanobrevibacter-associated community has not yet been described in studies of modern dental calculus" (lines 162-164). The distinction may boil down to a compositional

difference between the Methanobrevibacter community they characterize in the ancient calculus deposits and those that exist in rare modern contexts. A good way to demonstrate the difference would be to include a modern smoker's oral community in the PCoA plots, if possible, but from their response to my previous query, perhaps such data do not exist in a quantitatively compatible form. If that's not available, greater transparency about the frequency in which Methanobrevibacter is found in mouths of modern populations is warranted. They should for certain draw attention to the fact that this archaeal community is found in contexts without periodontal disease in the ancient cohort, as they have mentioned in their response to me.

Additional points:

1. I still find the introduction to be weighted on the importance of modern gut microbiomes, where the transition to and justification for the study of ancient oral microbiomes could be improved.

2. In line 292, I believe the authors are referring specifically to the mortality event in London between 1348 and 1350. Referring to this as "the arrival of the Second Pandemic" to me implies a continuous influence of the second pandemic throughout its duration. If the authors opt not to use the term "Black Death" (and there is justification for that), simply referring to the intense and temporally narrow mortality event might improve clarity.

3. Line 107 – "…we only included high quality samples with >100,000 sequences…" Again, I assume they mean >100,000 sequences that could be taxonomically assigned? This should be clarified in the main text.

4. Table S1 – SamplingNotes – what does "landed on jacket" mean? Or "landed on foam"? I suggest the authors go through this and make sure only accessible details are retained.

5. Table S2 has prolific highlighting outside the cells with text

6. The SI table file includes the MainTextTable, which it need not.

Author Rebuttal, second revision:

Replies to Reviewers:

Reviewer #2:

Remarks to the Author:

I am satisfied with the responses given by the authors to my comments, as well as those of the other reviewers, on the previous version of the manuscript. I think the current version is much more streamlined, and the over-interpretation of some results in previous versions has also been addressed. This study will add to the growing compendium of ancient dental calculus metagenomes, but more importantly, it will highlight the importance of considering oral geography while sampling dental calculus as well as a novel method of estimating diet from functional analysis of ancient oral metagenomes, both of which will

be of interest to ancient microbiome researchers. I am happy to recommend this manuscript for publication and look forward to seeing it published soon.
We thank the reviewer for their support of this publication and are glad that we were able to address their feedback!

My only comments on this version are related to the following edits:
Line 61 - Change "its disappearance is" to "their disappearance is".
We have made this modification to the manuscript.

Line 172 - Capitalize "Second Plague Pandemic" as you have done in the rest of the manuscript. Also, it would be better to say here that this is also known as the Black Death, since this is the first mention of it in the main text, rather than in Line 282.
We agree and have made these changes.

Line 269 - Typo. It should be "Peptoniphilus".
Thank you for catching this typo.

Line 269 - Change to "while not all systemic diseases..."
We agree and have made this change in the manuscript.

Line 322 - Typo. Change "sampled" to "samples".
Thank you. Done.

Reviewer #3:
Remarks to the Author:
The authors have satisfactorily addressed my main concerns and I am happy to recommend it for publication. I just added some thoughts below in regards to some of the answers given to my previous questions.
We thank the reviewer and are deeply appreciative of the insightful feedback they have provided throughout this process.

In regards to the comment about the number of reads:
While I might be able to think of some reasons why my approach might be more correct the discussion has been made rather irrelevant by providing the number of reads per sample. The chosen threshold seems rather sensible (I'd pick the same) and the order of operations would not have changed the selection of samples. In regards to contamination removal strategies involving multiple filtering steps; I typically keep all contaminant reads until the final step and then clean up the data. In any case, as you stated below, no approach with these kinds of samples is probably perfect, but I'm satisfied with this reply/outcome.
As the reviewer mentions, each approach has its strengths and drawbacks, and we are pleased that the reviewer finds ours to be well-described and appropriate.

In regards to the contamination fraction comment:
In several of my own medium/low biomass studies (non-ancient) contamination removal was quite straightforward. Even if a significant fraction of reads was contamination the remaining fraction in these samples clearly represented a microbial profile (after contamination removal) as could be expected (similar to samples with a low contamination fraction). However, a very good point is made in regards to ancient DNA samples that a high amount of contamination reads might indeed be indicative of increased degradation and hence of a lower quality oral composition representation after contamination removal (in addition, degradation might not affect all DNA from all species equally hence skewing results even more). It might be worth it to mention this line of reasoning in the contamination removal section.
We thank the reviewer for this suggestion and have ensured that this line of reasoning is present in the SI filtration and contamination-assessment section (pg. 11).

In regards to my comment regarding various contamination removal strategies:
I personally don't even use decontam myself and I'm very aware of its limitations. Commonly, people use decontam and then claim that they "solved" the contamination issue. This is especially problematic in very-low to no biomass samples, as the remaining signals are then presented as the "microbiome" of those samples. In any case, perfection is not needed for these samples (as it is clear that they had a microbiome) and it is now clear to me that sufficient effort has been put into getting the results to properly represent the ancient oral signatures (and importantly, that these efforts have become part of the manuscript instead).
We thank the reviewer for pushing for clarity and completeness in terms of ensuring that these efforts are represented in the manuscript.

Reviewer #4:
Remarks to the Author:
I thank the authors for the latest revised version of their work. While I appreciate the lengthy responses to my queries from the last round of review, I recommend more complete discussion of some concepts in the main manuscript. This most importantly applies to their disclosure that Methanobrevibacter is found in modern smokers and those with advanced periodontal disease. As the novelty of the ancient Methanobrevibacter community is a central conclusion of their manuscript, the authors are best advised to more transparently describe its novelty. While they do report on lines 181-183 that Methanobrevibacter is associated with severe periodontal disease in modern populations, they do not mention its association with smoking (as disclosed in their reply to my query). These modern contexts for Methanobrevibacter questions their core conclusion that "this Methanobrevibacter-associated community has not yet been described in studies of modern dental calculus" (lines 162-164). The distinction may boil down to a compositional difference between the Methanobrevibacter community they characterize in the ancient calculus deposits and those that exist in rare modern contexts. A good way to demonstrate the difference would be to include a modern smoker's oral community in the PCoA plots, if possible, but from their response to my previous query, perhaps such data do not exist in a quantitatively compatible form. If that's not available, greater transparency about the frequency in which Methanobrevibacter is found in mouths of modern populations is warranted. They should for certain draw attention to the fact that this

archaeal community is found in contexts without periodontal disease in the ancient cohort, as they have mentioned in their response to me.
We thank the reviewer for their comments on this manuscript. To address these points, we now include a line in the main text (L164 onward), which now also includes a direction to the SI for a more detailed discussion on the association between *Methanobrevibacter* in modern populations (pg. 26). We also agree with the reviewer that an in-depth examination of *Methanobrevibacter* across modern populations would be a fantastic project, but it outside of the scope of this project.
Additional points:

1. I still find the introduction to be weighted on the importance of modern gut microbiomes, where the transition to and justification for the study of ancient oral microbiomes could be improved.
We thank the reviewer for sharing their opinion regarding this point, but have elected to maintain the current structure of the introduction for the following reasons: First, only lines 77 to 82 explicitly focus on just the gut microbiome, while the rest of the first two paragraphs apply to both the gut and oral microbiomes. Second, we believe that the studies we mention in those lines are critical for setting up the goals of this paper.

2. In line 292, I believe the authors are referring specifically to the mortality event in London between 1348 and 1350. Referring to this as "the arrival of the Second Pandemic" to me implies a continuous influence of the second pandemic throughout its duration. If the authors opt not to use the term "Black Death" (and there is justification for that), simply referring to the intense and temporally narrow mortality event might improve clarity.
We thank the reviewer for their comment. As advised by another reviewer, we moved the phrase "also known as the Black Death" to lines 172-3 (where the Second Plague Pandemic is first mentioned) and now include the specific dates for clarity. We also more in depth describe this event in the SI text (pg. 20).

3. Line 107 – "…we only included high quality samples with >100,000 sequences…" Again, I assume they mean >100,000 sequences that could be taxonomically assigned? This should be clarified in the main text.
As advised by the reviewer, we have amended the phrasing in this section.

4. Table S1 – SamplingNotes – what does "landed on jacket" mean? Or "landed on foam"? I suggest the authors go through this and make sure only accessible details are retained.
We thank the reviewer for their detailed examination of this table and for providing these suggestions. We have elected to retain the sampling notes in their original format so as to provide as much information as possible for those who may want to include or reanalyze these samples in the future.

5. Table S2 has prolific highlighting outside the cells with text

We thank the reviewer for their attention and have addressed this formatting error.

6. The SI table file includes the MainTextTable, which it need not.
We thank the reviewer for their note.

Final Decision Letter:

Subject: Decision on Nature Microbiology manuscript NMICROBIOL-22051270C

Message: 16th October 2023

Dear Laura,

I am pleased to accept your Article "Ancient dental calculus reveals oral microbiome shifts associated with lifestyle and disease in Great Britain" for publication in Nature Microbiology. Thank you for having chosen to submit your work to us and many congratulations.

Over the next few weeks, your paper will be copyedited to ensure that it conforms to Nature Microbiology style. We look particularly carefully at the titles of all papers to ensure that they are relatively brief and understandable.

Once your paper is typeset, you will receive an email with a link to choose the appropriate publishing options for your paper and our Author Services team will be in touch regarding any additional information that may be required. Once your paper has been scheduled for online publication, the Nature press office will be in touch to confirm the details.

After the grant of rights is completed, you will receive a link to your electronic proof via email with a request to make any corrections within 48 hours. If, when you receive your proof, you cannot meet this deadline, please inform us at rjsproduction@springernature.com immediately. You will not receive your proofs until the publishing agreement has been received through our system

Due to the importance of these deadlines, we ask you please us know now whether you will be difficult to contact over the next month. If this is the case, we ask you provide us with the contact information (email, phone and fax) of someone who will be able to check the proofs on your behalf, and who will be available to address any last-minute problems.

Acceptance of your manuscript is conditional on all authors' agreement with our publication policies (see https://www.nature.com/nmicrobiol/editorial-policies). In particular your manuscript must not be published elsewhere and there must be no announcement of the work to any media outlet until the publication date (the day on which it is uploaded onto our website).

Please note that Nature Microbiology is a Transformative Journal (TJ). Authors may publish their research with us through the traditional subscription access route or make their paper immediately open access through payment of an article-processing charge (APC). Authors will not be required to make a final decision about access to their article until it has been accepted. Find out more about Transformative Journals

Authors may need to take specific actions to achieve compliance with funder and institutional open access mandates. If your research is supported by a funder that requires immediate open access (e.g. according to Plan S principles) then you should select the gold OA route, and we will direct you to the compliant route where possible. For authors selecting the subscription publication route, the journal's standard licensing terms will need to be accepted, including self-archiving policies. Those licensing terms will supersede any other terms that the author or any third party may assert apply to any version of the manuscript.

If you have any questions about our publishing options, costs, Open Access requirements, or our legal forms, please contact ASJournals@springernature.com

An online order form for reprints of your paper is available at https://www.nature.com/reprints/author-reprints.html. All co-authors, authors' institutions and authors' funding agencies can order reprints using the form appropriate to their geographical region.

You can now use a single sign-on for all your accounts, view the status of all your manuscript submissions and reviews, access usage statistics for your published articles and download a record of your refereeing activity for the Nature journals.

To assist our authors in disseminating their research to the broader community, our SharedIt initiative provides you with a unique shareable link that will allow anyone (with or without a subscription) to read the published article. Recipients of the link with a subscription will also be able to download and print the PDF.

As soon as your article is published, you will receive an automated email with your shareable link.

With kind regards,