Dear Dr. Hinson,

Editor’s comments:

I am enclosing the comments that the reviewers have made on your paper. I am afraid their opinion is that the data are not decisive enough for the paper to be a strong candidate for publication in Cell Host & Microbe, and therefore it seems that the paper should be published elsewhere. I want to emphasize that this decision is not intended to imply a lack of interest on our part in either your work in particular or this field in general. We hope that you will continue to consider Cell Host & Microbe for future submissions when appropriate.  
  
  
Reviewer #1   
  
Comment #1: “The paper attempts to fill a critical knowledge gap of understanding the abundance of oral microbes in fecal samples during states of ectopic expansion associated with gut dysbiosis. The hypothesis is that oral microbes out compete or are able to survive in the gut during gut dysbiosis and can perturb additional disease and or prevent the normal gut flora from reestablishing itself in order to return to the healthy state.

The authors create a mathematical model for the presence of oral microbes in the gut. The model makes a few predictions about the population. Namely, oral microbes are present in fecal samples in a bimodal, increased percent of oral microbes in fecal samples occurs because of reduced gut microbial levels, and that absolute values of oral microbes are very stable.  
To test these predictions, they created a program to estimate the quantity of oral microbes within human samples based on the composition of the sample's microbial population. As they point out, most studies do not have matched oral and fecal samples so a way to estimate oral microbes is required.

They use the program to investigate several studies of gastrointestinal disease. Combined with 16S and metagenomic data they found support for the model predictions as well as the Marker hypothesis, that the correlation of a higher percentage oral microbes in feces in gastrointestinal disease occurs due to reduced gut microbes rather than disease being driven by oral microbes.”

Comment #2: “The first major issue is the taxonomic identification which seemed to be restricted to the genus level. It is well established that multiple species within a genera can be detected across human body sites as is evident in Figure 2A as both oral and skin level taxa cluster together however these are not the same species or the same strains as oral that are found on the skin. No study to date has really conclusively proven at the strain level that oral bacteria are indeed found in the human gut in healthy subjects. The reference to finding evidence for this model prediction in the fecal microbiota of healthy Western adults, "which show bimodal distributions of Prevotella species commonly found in the oral cavity (Lahti et al.,88 2014; Tett et al., 2021)" is a major issue given that there is an abundance of non-oral Prevotella in the healthy human gut. This underlies the flaw of the manuscript in that it does not distinguish between oral Prevotella species like denticola, nigrescens and gut prevotella species copri, histicola as an example and assigns Prevotella to only oral. They wrongly conclude that even healthy people can have significant proportions of oral bacteria in their feces.  
See: The Prevotella copri Complex Comprises Four Distinct Clades Underrepresented in Westernized Populations. Adrian Tett et al. Cell Host Microbe. 2019  
Vinod Kumar Gupta; Narendrakumar M Chaudhari; Suchismitha Iskepalli; Chitra Dutta (March 5, 2015). "Divergences in gene repertoire among the reference Prevotella genomes derived from distinct human body sites of human”

Comment #3: “They also looked for co-occurrence of ASVs in the samples. However, again the assignment of ASVs to oral was based entirely on genus, rather than appearance of specific ASVs in oral and fecal samples. So, it is unclear what these results are really showing.”  
  
Comment #4: “HMP and additional studies listed for source data provided species level assignments. I am concerned that the lower taxonomic assignments only at the genus level affected the results of their model and maybe why they are not included. However, this is not addressed in the current manuscript.”  
  
Comment #5: “A related issue is with the estimation of oral microbes. the centerpiece of the entire paper. The authors develop the program using data from the Human Microbiome Project (HMP). Unfortunately, the program was not tested for effectiveness. Figure 2D shows predictions for HMP, but that was the training dataset and so isn't a test of the program. They also compare their results to a Random Forest classifier and a Bayesian approach using SourceTracker2. Their results had reasonable correlation with these quantification approaches. But again, these models were trained on HMP and the results compared from analyzing HMP. While a means of estimating oral microbes is required because most studies do not have matched oral and fecal data, some studies should. To show the efficacy of the program the authors would ideally use matched data to establish a ground truth for the presence of oral microbes within fecal samples and compare these results with their estimates. At the very least, the authors need to test the program on data not used in the training set.  
  
With that said the paper points out that there are few data sets that have paired gut and oral samples. However when trying to extrapolate to data without matched samples it would have been critical to include an evaluation of at least one matched dataset that could have been used to evaluate the models they developed and trained on unmatched sets.”  
  
Comment #6: “Additionally, Random Forest classifier and SourceTracker 2 are referred to as well-established microbial source tracking methods. There is little justification to develop a third method. Their estimates might be better than these methods, but without any demonstration of superior accuracy or explanation this alternative approach seems to lack support.”  
  
Comment #7: “More generally, the writing of the paper seems somewhat unfocused. For example. It is established that the percent of oral microbes in fecal samples is higher in the case of gastrointestinal disease. Thus, the main results of the paper are showing evidence for their mathematical model and for the Marker hypothesis. The primary support for these is shown in Figure 2B and C, out of A-I. They are almost lost in the clutter of plots. The paper needs a more organized approach to presenting the results. Overall the idea is interesting, but the data is a long way from supporting either hypothesis in the current form. Given these issues I cannot recommend publication of the paper in its current form.  
  
Comment #8: “The issue of depth is a very important one. If depth of sequencing is low, and not capturing the diversity of a sample, it will appear to have lower diversity. The correlation of sequence depth with the presence of "oral" bacteria reads should be determined.”

Comment #9: “Line 146. Shotgun sequencing does not produce ASV's, one can get strain information however this is not the same as an ASV. And the two cannot be linked unless there is a complete 100% overlap at the sequence level for the 16S rRNA gene to prove they are the same ASV.”

Comment #10: “It is unclear if the differences across studies maybe be distorting the authors interpretation of their results. For example, it is unclear if the studies included in this meta-analysis were comparable at all. It is more likely that the different 16 studies were not performed using the same sequencing platform (Roche, Illumina, etc), similar primers, similar amplified regions of interest (V1-9), sequence depth, single or paired end, number of bp overlap, extraction/sequencing kit, etc. This is evident when the authors suggest they used the OTU or ASV designation provided by the original study. Thus it is likely clustering and divisive noise reducing strategies were used differently in the studies used to generate their model. This is not addressed in the current manuscript.”

The taxonomic assignments were also likely assessed with various versions of different reference databases, this is not addressed in the current manuscript.”

Comment #11: “The authors include a random forest model with a maximum of 1000 trees. This is a vast under sampling of the data, whereas common 16S microbiome papers that apply this strategy have used a >5000 trees.

<https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-018-0568-3>  
a microbiome paper applying a similar technique using 10,000 trees”  
  
  
  
Reviewer #2

Comment #1: “In their manuscript, "Oral bacteria in feces reflect loss of intestinal bacteria", Liao and colleagues develop a theoretical model and examine data from several studies, most notably a study involving 10,000 HCT patients, in order to answer the question of whether oral microbes bloom in the gut or whether other organisms are depleted. I commend the authors in looking at a wide range of datasets. This is an important question and the authors note that it would have implications for treatment of microbiome-associated disorders. The authors also note that the predominance of oral bacteria in patient cohorts, specifically those involving supragingival and subgingival plaque, have reduced survival rates.  
  
The authors' argument rests on models, and the observations that oral species are common to all individuals' guts and that the 'gut-only' species are depleted under antibiotic treatment, as evidenced by relative abundance data and absolute abundance data (measured by qPCR and the weight of feces). I have several major concerns about these claims, which affect all subsequent observations.”

Comment #2: “The authors provide limited data that oral species are in fact in all individuals' guts. The lack of oral data on the HCT or any antibiotic treated cohort is a major weakness in the paper, but the authors could have propped up their modeling efforts by showing identical ASVs in paired HMP samples, or genome identities in paired HMP metagenomic samples. This is critical for their claims and important to show for both healthy individuals and individuals with dysbiotic microbiomes.”  
  
Comment #3: “The authors focus heavily on genera common to both oral and gut sites, yet grouping species at the genera-level is likely to lead to spurious inferences. The most obvious example is the Prevotella genus (see figure 1E, lower section and Figure S1). P. copri and to some extent P. ruminocola are the only Prevotella species commonly found in the gut, whereas the oral sites are dominated by species that diverge from the gut species, namely P. dentalis, P. melaninogenica, P. denticola and P. intermedia. These species have starkly distinct habitats and are rarely found in the other site, yet they would be grouped in this analysis. P. copri is also more prevalent in non-US cohorts, and may skew results from other cohorts imputed from the HMP data. Prevotella is the obvious case, but this issue likely affects additional genera.”  
  
Comment #4: “Given that another major claim rests on abundance measurements that allow the researchers to calibrate their relative abundance data, it would have been nice to see more than one quantification of the stool amounts (another method besides qPCR).”  
  
Comment #5: “Although the writing is clear for the most part, there are places where the authors either assume knowledge of the reader or provide too little information for the author to make out what they did. For example: Line 148, the researchers mention peak-to-trough ratio. The authors assume the reader is familiar with this approach. Additionally, metagenomic sequencing is necessary for this analysis, but there was no preface of what metagenomic data they had (it was all 16S up until this point). It is also misleading as it seems to imply that this analysis was performed on the same samples from which the 16S data came from, but it is not clear if that was the case.”  
  
Comment #6: “Overall, the researchers should be more transparent when talking about predicted or imputed oral abundances, since the way it reads now, with so many datasets, it obscures the places in which the authors used actual data (HMP) versus predicted data (most everything else). This also goes for the validation of their predictions—it was difficult to determine what was the actual measured ASV amounts versus predicted via their method or SourceTracker etc.”  
  
Comment #7: “The authors' model and the text overall implies that migration of oral organisms to the gut microbiome is mainly through the alimentary canal. There is evidence that translocation happens through the blood stream (doi: 10.3389/fcimb.2020.00400).”  
  
Comment #8: “The github link was non-functional.”  
  
Reviewer #3

Comment #1: “This manuscript by Liao et al. tackles the open question of why increases of typically oral microbes are seen in gut microbiomes of diseased individuals. There are two open hypotheses, with analysis and data in this work supporting the "marker hypothesis". This hypothesis posits that these increases in oral bacteria are simply a marker of a diseased state, resulting from the loss gut taxa (vs. the oral taxa being the drivers of disease). An ecological model that supports this hypothesis is presented that predicts how the abundance of oral taxa in the gut would be affected by flow rates through the system, carrying capacity, and perturbations. Actual gut microbiome data in the form of 16S are then examined across multiple datasets to demonstrate predictions from the model. This manuscript addresses a pertinent question in the field using the mix of theory with data validation. There are many really well done, convincing analyses. The use of a variety of datasets and disease types and the inclusion of total microbial density were especially effective. That being said, there are several areas of concern I have that limit my enthusiasm for the conclusions as presented.”  
  
Comment #1 “My largest concern is over the accurate calling of "oral" taxa in the fecal dataset. Most of the taxa identified as "oral" are gut residents as well, including Prevotella, Veillonella, Actinomyces, etc. Prevotella for example are commonly found in the gut, and often make up large percentages of gut bacteria in non-Western populations. Furthermore, Prevotella collected from different body sites contain different genomic content, suggesting these aren't just translocations from one body site to another, but of evolution while residing in different body sites (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4359502/>). It's not clear what amount of each of these taxa are oral transplants vs. resident taxa in the gut. Classifying them purely as oral taxa will likely lead to misleading estimates of oral transfer, depending on the baseline occupancy of the gut strains of these taxa. Without a calibrated sense of how many of these "oral" bacteria are actually oral in origin, it's difficult to interpret the results of this manuscript.”  
  
Comment #2: “What would clear this issue up would be having a set of paired oral and gut microbiome samples from the same individuals where the transfer of strains of oral origin could actually be quantified down to the strain level. The methods in this paper may very accurately correlate with the actual rate of transfer, but without the underlying experimental data showing that's true (vs. different strains in the oral and gut habitats), it's very hard to assess this. This would require further data collection, but is essential to demonstrate that "oral" taxa are being classified accurately.”

Comment #3: “There are several concerns with the methodology used to identify oral taxa as presented (NMF).  
a. There are large differences between the methodology proposed here (NMF) and RF and SourceTracker as indicated by the r > 0.6 between methods. The taxa defined as "oral" do not correlate well between methods (as demonstrated in Figure S1). It's not clear which method is potentially most accurate though, including the newly proposed method. A simulation analysis that demonstrates precision and recall of each method to distinguish source environments on known data would improve interpretability for the reader, providing evidence of effectiveness of each method.  
b. For analyses like those presented in Fig. 3B, how are the abundances of gut and oral taxa being calculated? Is it from estimated proportions of gut and oral overall from each method, or was it by pulling the abundances of individual indicator taxa? Further clarification on how outputs are being used would be beneficial for the reader.”  
  
Comment #4: “The github page with sourcecode is not accessible (perhaps set as private repo?)”

Comment #5: “Page 19: Rarefaction to 100 sequences/sample doesn't seem like enough sequences for decent accuracy with SourceTracker.”