In this work, Sze and Schloss perform a meta-analysis of previously published 16S rRNA sequencing studies that focused on individuals with colon adenoma pr colorectal cancer. The main finding was that they were able to generate a model using these data that might be useful as a non-invasive predictive marker of cancer, though the models underperformed with adenomas. In my personal opinion, the most important and interesting finding in this work is what the authors did not find: that is, several of the previously identified "CRC-associated taxa" including Fusobacterium and several others, did not appear to be generalizable markers, at least across the studies included in this meta-analysis. I am enthusiastic about this sort of replicability, as there is indeed quite a lot of hype surrounding the microbiome and its role in human health and disease. This is pervasive in pop science and unfortunately also among some funding study section members who have the mistaken assumption that we should find single, causative taxa in disease states.

One point I am curious about, that perhaps the authors can address: would you expect to see the same "patchiness" in the microbial data were you to collapse the taxa down to functions, al a PICRUSt or Tax4Fun? How much of the patchiness would you predict is real biological signal as opposed to noise generated by heterogenous sampling, etc? In other words, it would interesting to see a short paragraph in the discussion related to function. If you wanted to run a small analysis on predicted function, that would be incredible, but not required for this particular manuscript. Nice work.

We appreciate the reviewer's support for our work. Like the reviewer and Editor, we think these types of replication analyses are important to move the microbiome research field forward. We are reluctant to say too much more about functions than we already have in the manuscript because the correlation between 16S rRNA gene sequences and function is rather tenuous. The manuscript currently states:

"In contrast to the patchiness of the taxa that were positively associated with carcinomas, potentially beneficial taxa had a more consistent association [Figure 6]. This result was particularly interesting because members of these taxa (i.e. Ruminococcus and Clostridium XI in fecal samples and Dorea and Blautia in tissue) are thought to be beneficial due to their involvement in production of anti-inflammatory short chain fatty acids (32–34)."

Previous work by Turnbaugh et al. (DOI: 10.1038/nature07540) when looking at their Missouri twins cohort showed that when a sequenced metagenome is analyzed using KEGG categories, there is little inter-personal variation. We suspect that this is due to the fact that KEGG categories largely focus on house keeping functions that all bacteria need. PICRUSt and Tax4Fun are largely based on this framework as well. Although

they add other sources for mapping 16S rRNA genes to annotations, the level of annotation is somewhat limited. Furthermore, the inferences provided by such tools are difficult to place a lot of confidence in. This is because the analysis only suggests whether a gene is present, not that it is being transcribed and translated. Also, our public database are biased towards bacteria that are associated with pathogenesis and so those annotations dominate even if the bacterium that is found in the 16S rRNA gene survey is not pathogenic. We look forward to the possibility that some day more metagenomic shotgun sequencing data will be published allowing us to perform a similar meta-analysis.