**Finding microbiome diversity and composition difference of Hmong and Karen ethnicity groups based on public housing condition**

Guangxin Chen, Ben Greene, Laine Ciaramitaro, Laurel Schuck, Nicole Koziol

University of Minnesota

05/05/2021

**Abstract**

Public housing is a form of housing tenure that often located near highways or busy roads and would cost the least to live. Residents of public housing usually have relatively poor health condition that can be caused by living environment or their low income, which would potentially affect their microbiome pattern. Thus, we assumed that residents of public housing have a different diversity and composition of human intestinal microbiome than individuals in private housing. Our hypothesis was that individuals living in public housing would have lower microbiome alpha and beta diversity than people in private housing, regardless of their ethnicity. The raw sequencing data and file with metadata information provided in the paper of Vangay were analyzed in our project, and we used QIIME2 to convert raw DNA reads to text files related to alpha and beta diversity metric. Three plots related to alpha diversity, beta diversity and taxa summary were made using R based on those text files. As a result, no significant correlation between microbiome diversity and public housing was found. Beta-diversity of individuals for group with different public housing condition did not have a significant difference which is contradict to our hypothesis. While further measurements like recollecting data and using different statistical techniques can be made to make this result more complete.

**Introduction**

Human microbiome refers to trillions of microbes which inhabit in our bodies and can interact with host cells, leading to an adaptive and body-part-specific ecosystem that can affect host both genetically and physiologically. It is estimated that around 500-1000 species of bacteria can exist in certain human body part at a time, and variations usually exist between people and within a person over time (Turnbaugh et al. 2007). Changes in human microbiome, and their interaction with nervous, immune system are related to many diseases and disorders. Thus, understanding how the human microbiome changes with age is important both for scientific discovery and human health. Human gut microbiome is one of the well-studied system due to its simplicity to acquire sample and its direct relation with diet. It is found that dietary patterns and environmental factors have a significant effect on the abundance and composition of human gut microbiome. For example, researchers found that migration from a non-Western country like Laos and Burma to the United States tends to have a loss of microbiome function and composition. In this process, native strains and functions were displaced by US-associated strains and functions (Vangay et al. 2018). Besides immigration, human gut microbiome is also found to be related with other environmental factors like living conditions, neighborhood environment situations and host behaviors (Pearson et al. 2020). But the specific relation and mechanism associated with these factors are not clear. In this paper, we try to find how the diversity and abundance of human gut microbiome is related to living conditions by analyzing provided dataset. And try to answer the question that whether people living in public house have lower alpha and beta-diversity than people living in common apartment and houses.

Public housing refers to the housing tenure whose property is usually owned by the government authority and most public housing is located in the metropolitan areas where land is cheapest. This includes locations near highways, busy roads, or even airports; as that is where it would cost the least to live (Krieger & Higgins, 2002). Researches also find that public housing residents have extremely poor health condition, while it is debatable that whether the poor health is caused by public housing condition or the low income of the residents (Ruel et al. 2010). Based on these observations, we assume that residents of public housing have a different abundance or component of human gut microbiome comparing to non-public housing residents. Specifically, we hypothesis that individuals living in public housing will have lower microbiome alpha and beta-diversity than those that don’t regardless of their ethnicity. That is, people who live in public house could have less diverse gut microbiomes than individuals in private housing, and the similarity between public housing individuals are more prevalent than that of private housing individuals. This topic is novel because little conducted researches, it also has large societal impacts to address the problem that caused poor health condition of public housing residents.

To test our hypothesis, data set collected from “US Immigration Westerizes the Human Gut Microbiome” will be used. In this paper, researchers collected samples from migrations of non-western nation like Hmong, Karen and Caucasian to U.S and found loss of gut microbiome function and diversity. In our project, we choose samples of Karen and Hmong and divided them into two groups based on whether they are living in public housing. Our hypothesis would be supported if substantial change will be found in both alpha and beta-diversity based on public housing condition. In that case, we can conclude that living condition is a main driving force for human gut microbiome composition change and can be stronger than the effect caused by ethnicity.

**Method**

*The collection of datasets*

The dataset provided in the paper of Vangay et al. was analyzed in our project. Samples included adult females who were Hmong or Karen and born in Southeast Asia. They were either currently living in Thailand or moved to U.S. Stool samples of them were collected and were submitted to University of Minnesota Genomics Center (UMGC) for sequencing. In this process, DNA samples from 663 subjects were extracted and their 16S ribosomal rRNAs were sequenced, along with performing shotgun metagenomics sequencing on the Illumina HiSeq platform of fecal sample. The 16S maker sequencing data were trimmed using a quality control pipeline SHI7, OTUs were picked with BURST (Al-Ghalith and Knights, 2017) OTU picking algorithm.

Also, QIIME was used to measure alpha diversity with rarefied OTU tables. And the shotgun metagenomics sequences were aligned using the NCBI RefSeq database (Tatusova et al., 2013) and identified in the species level. Resources from Minnesota Supercomputing Institute (MSI) were used for further sequencing analysis (Vangay et al. 2018).

*Quality control of sequencing data*

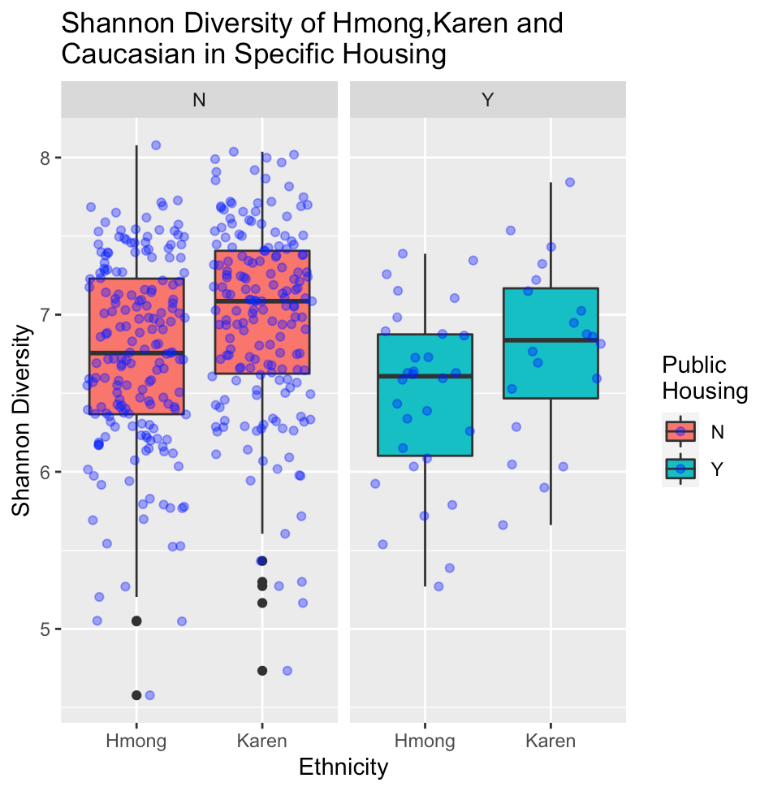
The raw DNA reads with .fastq format were converted to .qza done in the terminal using QIIME2. The DADA2 quality control software was then used to quality filter the sequences and a de novo OTU table was generated based on 97% similarity to other samples reads with no reference database. The OTU table and file with metadata information of Vangay dataset were then acquired from MSI. Based on our hypothesis, a subset of the mapping file was taken. Sample IDs which have non-certain value for the public housing entry were excluded from both OTU table and mapping file. Then certain sequencing depth was picking and the data on OTU table were rarefied to 10,000,000 reads using Qiime 2([github script](https://github.umn.edu/kram0247/Group7_3004_S21/blob/master/rarefaction%20code)). Then OTU and mapping files were used to calculate alpha (Shannon Index) and beta diversity metrics (Jaccard Index, and the Bray Curtis Dissimilarity measure). The rarefied OTU table was then converted to text file and loaded to R to generate various plots.

*Plotting and statistical analysis*

Based on our hypothesis, three plots related to alpha diversity, beta diversity and taxa summary were made using R. After loading to R, all the tables were edited and filtered using tidyverse package and plotted with ggplot2. Specifically, for the alpha diversity plot, the Shannon index was chosen and plotted related to different ethnicity groups. Individuals who lived in public housing and did not live in public housing were plotted separately. Tukey-Kramer test was used to find was there a significant difference for both groups due to their unequal sample number. Similarly, a PCoA plot was generated to cluster data into different group based on Bray Curtis dissimilarity measure. The overlap percentage for data of different groups was analyzed using Permutational multivariate analysis of variance (PERMANOVA). Also, a taxa summary plot was generated to visualize taxa composition found in samples using phylum taxonomy based on ethnicity groups and whether lived in public housing. Packages like reshape2 and plyr were used in this plotting process.

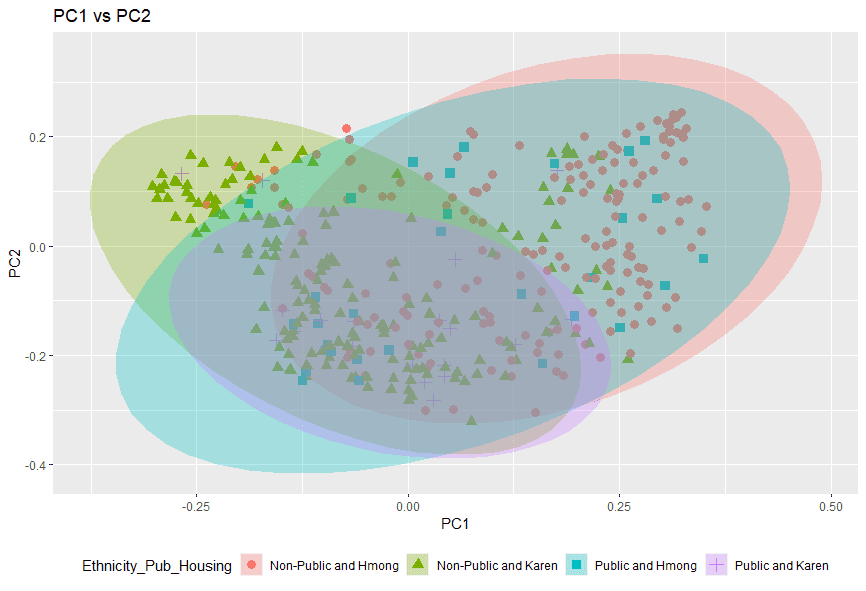
**Result**

**Alpha-diversity of both Hmong and Karen group show a significant difference based on public housing condition.** We generated box plot to show the distribution of OTU number for individuals in Hmong and Karen group regarding their public housing condition, and used Tukey-Kramer test to find whether there is a difference in the mean of OTU (Figure 1). The result shows that those in public housing have a significantly different microbiome diversity to those not in public housing, with a p-value samller than 0.05 and sample number of 408.



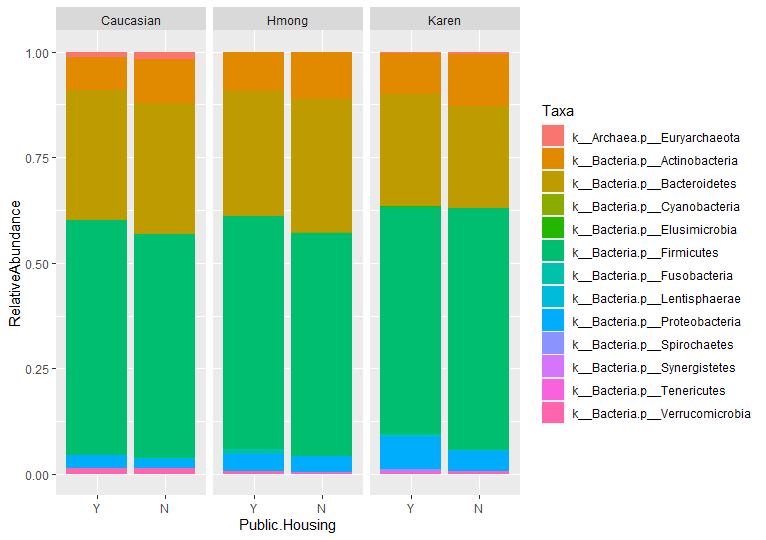
**Figure 1:** Box plot for Shannon index of Hmong and Karen in specific housing. Those in public housing did have a significantly different OTU number than those not in public housing based on Tukey-Kramer test (p-value = 0.00575, N = 408).

**Beta-diversity of Hmong and Karen group were close to each other regardless of public housing condition.** To test our hypothesis about difference of beta-diversity for individuals who live in public housing or not, a PCoA plot was generate in R based on the Bray-Curtis dissimilarity matrix provided in Vangay dataset (Figure 2). We tested for significance differences based on ethnicity groups and public housing condition using PERMANOVA analysis, and the result shows a large overlap between data of different group, which is supported by low value in the test. Also, ethnicity and public housing combined only account for about 6.86% of the variation (P<0.001). This may not support our hypothesis and indicate that microbiome composition of different groups is pretty similar to each other.



**Figure 2:** PCoA plot of beta diversity: shows significant overlap in microbiome composition between all ethnicities and housing groups studied. PERMANOVA analysis indicated that Ethnicity and Public Housing combined only account for about 6.86% of the variation (P<0.001).

**Different groups shared similar composition of intestinal microbiome regradless of enthnity and public housing condition.** The taxa summary graph was generated to show the composition and significant difference of intestinal microbiome for each group based on phylum (Figure 3). The relative compositoin amouts of each bacteria were calculated based on OTU number. It can be found that different groups share similar microbiome composition, with three major phylum of microbe: *Firmicutes*, *Bacteroidetes* and *Actinobacteria*. While we could not find significant change of microbes composition based on public housing condition, which tended to show that people share similar composition regradless of public housing and may not support our hypothesis.



**Figure 3:** Taxa summary plot displays the bacterial taxa present in Caucasian, Hmong, and Karen participants based on whether they live in public housing. Different colors indicating different phylum of bacteria. All three groups had large amounts of Firmicutes and Bacteroides bacteria.

**Discussion**

In this project, our hypothesis is that individuals living in public housing have lower intestinal microbiome alpha and beta diversity than people live in other housing, regardless of their ethnicity. While our results show that the influence of public housing to human microbiome condition is weak, and no significant correlation between microbiome diversity and public housing could be found. From the alpha diversity figure, we found that Hmong and Karen participants in public housing did have a significantly different microbiome than those not in public housing. This result was supported by Tukey-Kramer test with a p-value equal to 0.00575 and sample size of 408. This is consistent with previous researches about living condition and human microbiome diversity. It is found that people lived in high latitude areas in India have a significant lower gut microbiome alpha diversity than people lived in low latitude (Das el al. 2018). Also, it is found that there is a substantial difference between the diversity and taxonomy in the gut microbiome of urban population with rural population ([J Soto-Girón](https://pubmed.ncbi.nlm.nih.gov/?term=Soto-Gir%C3%B3n+MJ&cauthor_id=33872206) et al. 2021). Though result reflects the potential influence of living environment to diversity of human microbiome, this significance may stem from the influence of ethnicity. Further research can be made to find out the mechanism for this difference.

From the PCoA plot, it could be found that there was a significant overlap in human intestinal microbiome composition for all the studied group with different ethnicities and public housing condition, which indicates similarity between public housing individuals is not substantially different with that of private housing individuals. This was supported by the PERMANOVA analysis using 1000 permutations with a p-value smaller than 0.001, indicating that the beta-diversity of individuals for group with different public housing condition did not have a significant difference which is contradict to our hypothesis. However, ethnicity could have more influence on the beta-diversity than housing condition based on the plot, which makes sense because same ethnicity group can share similar pattern of diet and habit, leading to unique microbiome pattern. Other studies also indicate that same ethnicity group identity tends to share similar microbiome diversity and composition. It is found that the Shannon alpha diversity and beta diversity have significant difference between different ethnicity groups, and the recurrent association between ethnicity and specific gut microbe taxa is also reported (Brooks et al. 2018). Further research can be made on testing how ethnicity affects beta-diversity for different groups using Vangay’s data.

For the taxa summary plot, it can be found that the composition of microbiome in each public housing group was similar with slight difference. It can also be found that all three ethnicity groups have large amount of *Bacteroides* and *Firmicutes*, which is consistent with the result in Wakita’s paper indicating that the microbiota of most healthy adults are dominated by these two bacterial phyla (Wakita et al. 2018). *Bacteroides* may have immunomodulatory effects and promote host health and *Firmicutes* is found more related with obesity and diabetes (Wexler & Goodman, 2017). This also shows that the effect of different ethnicity group to the human microbiome composition is not so significant in given dataset, thus both public housing condition and ethnicity cannot affect microbiome composition significantly.

The inconsistency of our expected results and actual results can be interpreted from many aspects. First, our study used dataset collected from Vangay’s paper, which was not targeted for public housing study. Non-public housing can include a variety of housing conditions which can exert different effects on the diversity and composition of microbiome. Also, the sample size for individuals lived in public housing is smaller than 100, while individuals lived in private housing have a sample size of more than 300 hundred. Such difference in sample size can potentially lead to larger error in statistics measurement. Besides, there was a mismatch of the relative OTU table and file with metadata information such that taxa cannot be directly mapped, which may make the taxa summary plot not accurate as we expected. We can conduct some further measures to make this project more complete. First, we can recollect data based on our research questions and try to have an equally distributed sample for individuals live in public housing or not. A complete OTU table with matched taxa information will be generated and analyzed. What’s more, the ratio of *Firmicutes* to *Bacteroidetes* can be calculated for public housing and private housing group, as this number can be interpreted as a marker for gut dysbiosis (Magne et al. 2020). Thus, the heath condition of individual can be estimated and related with living condition. It is also possible to use a machine learning model to estimate how each variable like public housing condition and ethnicity can affect microbiome diversity. In fact, there is study using support vector machines (SVMs) to predict host gut dysbiosis using metabolome data of human microbiome (Larsen & Dai, 2015), and it is possible to train our model with new calculated *Firmicutes* to *Bacteroidetes* ratios and recollected dataset to predict host living condition and gut health situation. Although a promising result cannot be acquired from this project, the topic related to housing condition and how it can affect human microbiome is still an intriguing topic to study due to its novelty and potential societal impact. The results of further study can be provided to government administration for public housing supervising and we can devote more effort on that in the future.

**Reference**

1. Al-Ghalith, Gabriel & Knights, Dan. (2017). BURST enables mathematically optimal short-read alignment for big data. 10.1101/2020.09.08.287128.
2. Brooks, A. W., Priya, S., Blekhman, R., & Bordenstein, S. R. (2018). Gut microbiota diversity across ethnicities in the United States. *PLoS biology*, 16(12), e2006842. <https://doi.org/10.1371/journal.pbio.2006842>
3. Das, B., Ghosh, T. S., Kedia, S., Rampal, R., Saxena, S., Bag, S., Mitra, R., Dayal, M., Mehta, O., Surendranath, A., Travis, S., Tripathi, P., Nair, G. B., & Ahuja, V. (2018). Analysis of the Gut Microbiome of Rural and Urban Healthy Indians Living in Sea Level and High Altitude Areas. *Scientific reports*, 8(1), 10104. <https://doi.org/10.1038/s41598-018-28550-3>.
4. Krieger, J., & Higgins, D. L. (2002). Housing and health: time again for public health action. *American journal of public health*, 92(5), 758–768. <https://doi.org/10.2105/ajph.92.5.758>
5. Larsen, P. E., & Dai, Y. (2015). Metabolome of human gut microbiome is predictive of host dysbiosis. *GigaScience*, 4, 42. https://doi.org/10.1186/s13742-015-0084-3
6. Magne F, Gotteland M, Gauthier L, et al. The Firmicutes/Bacteroidetes Ratio: A Relevant Marker of Gut Dysbiosis in Obese Patients?. *Nutrients*. 2020;12(5):1474. Published 2020 May 19. doi:10.3390/nu12051474
7. Pearson, A. L., Pechal, J., Lin, Z., Benbow, M. E., Schmidt, C., & Mavoa, S. (2020). Associations detected between measures of neighborhood environmental conditions and human microbiome diversity. *The Science of the total environment*, 745, 141029. <https://doi.org/10.1016/j.scitotenv.2020.141029>
8. Ruel, E., Oakley, D., Wilson, G. E., & Maddox, R. (2010). Is public housing the cause of poor health or a safety net for the unhealthy poor?. *Journal of urban health : bulletin of the New York Academy of Medicine*, 87(5), 827–838. <https://doi.org/10.1007/s11524-010-9484-y>
9. Soto-Girón, M. J., Peña-Gonzalez, A., Hatt, J. K., Montero, L., Páez, M., Ortega, E., Smith, S., Cevallos, W., Trueba, G., Konstantinidis, K. T., & Levy, K. (2021). Gut Microbiome Changes with Acute Diarrheal Disease in Urban Versus Rural Settings in Northern Ecuador. *The American journal of tropical medicine and hygiene*, tpmd200831. Advance online publication. <https://doi.org/10.4269/ajtmh.20-0831>.
10. Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Ciufo, S., and Li, W. (2013). *The NCBI handbook* (National Center for Biotechnology Information).
11. Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R., & Gordon, J. I. (2007). The human microbiome project. *Nature*, 449(7164), 804–810.
12. Vangay P, Johnson AJ, Ward TL, Al-Ghalith GA, Shields-Cutler RR, Hillmann BM, Lucas SK, Beura LK, Thompson EA, Till LM, Batres R, Paw B, Pergament SL, Saenyakul P, Xiong M, Kim AD, Kim G, Masopust D, Martens EC, Angkurawaranon C, McGready R, Kashyap PC, Culhane-Pera KA, Knights D. US Immigration Westernizes the Human Gut Microbiome. *Cell*. 2018 Nov 1;175(4):962-972.e10. doi: 10.1016/j.cell.2018.10.029. PMID: 30388453; PMCID: PMC6498444.
13. Wakita, Y., Shimomura, Y., Kitada, Y. et al. Taxonomic classification for microbiome analysis, which correlates well with the metabolite milieu of the gut. *BMC Microbiol*18, 188 (2018). <https://doi.org/10.1186/s12866-018-1311-8>
14. Wexler, A. G., & Goodman, A. L. (2017). An insider's perspective: Bacteroides as a window into the microbiome. *Nature microbiology*, 2, 17026. <https://doi.org/10.1038/nmicrobiol.2017.26>.