Developing a Human Gut Microbiome Database: Alistipes, Bacteroides, Faecalibacterium, Parabacteroides, Ruminococcus and Their Association to Inflammatory Bowel Disease and Irritable Bowel Syndrome

Student: Erin Ledford Supervisor: Kate Cooper

University of Nebraska at Omaha

Author Note

Student Email: [eledford@unomaha.edu](mailto:eledford@unomaha.edu) Supervisor Email: [kmcooper@unomaha.edu](mailto:kmcooper@unomaha.edu)

Abstract

This project is to develop a database of diseases that are associated with a small selection of bacteria from the human gut microbiome and what the healthy and unhealthy profiles of these bacteria look like. A standardized database using microbiome metadata standards will be created with data on the relative abundance of bacteria in a healthy human compared to what the relative abundance looks like when humans suffer illnesses like irritable bowel syndrome (IBS), inflammatory bowel diseases (IBD): ulcerative colitis (UC), and Crohn’s disease (CD). A website will be created that is searchable and shows the data gathered into a format that is easy to understand. The code created from this development will be freely available for use on GitHub.

Keywords: IBS, CD, UC, IBD

Developing a Human Gut Microbiome Database: Alistipes, Bacteroides, Faecalibacterium, Parabacteroides, Ruminococcus and Their Association to Inflammatory Bowel Disease and Irritable Bowel Syndrome

# Introduction

The human microbiome consists of the microbes, mostly bacteria, that are in and on the human body. There are trillions of microbes on and in the human body [4, 5]. The human microbiome is diverse and unique to everyone. [5]

Most of the bacteria found in the human gut microbiome are from phyla *Firmicutes*, *Bacteroidetes*, *Actinobacteria*, *Proteobacteria*, and *Verrucomicrobia*. *Firmicutes* and *Bacteroidetes* represent most of the gut microbiota. [5] Choosing from the available data from the Human Microbiome Project, the genus’ *Bacteroides* (taxid816), *Parabacteroides* (taxid375288)*,* *Alistipes* (taxid239759), *Faecalibacterium* (taxid216851), and *Ruminococcus* (taxid1263) were chosen to look at for the development of this database. [6]

### Bacteria.

*Bacteroides* consists of greater than 20 species. [9] These species are anaerobic, bile-resistant, non-spore forming, gram-negative rods [9, 16]. Due to their physiology, *Bacteroides* species are antibiotic resistant as Bacteroides produce beta-lactamases, have ribosomal protection by tetQ class, and ribosomal modification. [16-17] *Bacteroides* *fragilis* is the species of *Bacteroides* that is seen most often in clinical studies due to its virulence. [9, 16-17] *Bacteroides* are also known for their ability to do carbohydrate fermentation within the gut that allows the host to utilize as an energy source. [9]

*Parabacteroides* are gram-negative, anaerobic, non-spore forming, non-motile and rod-shaped. [11-12] *Parabacteroides* species were once under the *Bacteroides* genus but were moved under the new genus due to the differences in phylogeny compared to other *Bacteroides* species. [11] There are 15 *Parabacteroides* species, but only 10 species have valid names. [12]

*Faecalibacterium* is a genus with only one species, *Faecalibacterium* *Prausnitzii*. This species is gram-positive, mesophilic, rod-shaped, and anaerobic in physiology. [10] *Faecalibacterium* *Prausnitzii* consumes acetate and produces butyrate and bioactive anti-inflammatory molecules in the gut. [10, 14]

*Alistipes* is a genus with 13 different species. Alistipes is anaerobic, gram-negative, rod-shaped, non-spore forming, and a relatively new genus as an offshoot of the Bacteroidetes phylum. Alistipes can hydrolyze tryptophan to indole. This genus is resistant to antimicrobials vanomycin, kanamycin, and colistin. [13]

*Rumincoccus* is a genus with 6 different species, as most of the *Ruminococcus* genus have been reclassified as *Blautia*. *Ruminococcus* is gram-positive and anaerobic. *Ruminococcus* degrade cellulose. [15, 30]

### Sequencing.

16S ribosomal RNA (rRNA) gene is used for sequencing the microbiome as it is in nearly all bacteria by using polymerase chain reaction (PCR). [5, 25, 26] Its function has not changed, which allows for understanding of bacterial evolution. This gene contains nine variable regions, V1-V3 and V3-V5 as used in the Human Microbiome Project. [2, 8, 26] 16S rRNA provides accurate genus identification but less accurate species identification. Other issues with 16S rRNA sequencing are its inability to tell how closely related by sequence similarity other species may be, its inability to tell between recently diverged species, and poor-quality data with greater than 2% random sequencing errors. [25]

Whole Genome Shotgun (WGS) sequencing uses random primers to sequence overlapping regions of a genome to provide analysis of the entire genome. WGS sequencing has better resolution of taxonomy at the genus and species level compared to 16S rRNA. WGS sequencing is less popular than 16S rRNA due to its cost and the time it takes to sort through the amount of data it produces. [27,28]

Relative abundance is the fraction of (observed taxon / all observed taxa), with this, relative abundance should come to a max of 1.0 or a percentage of 100%. [19] Differential abundance is the comparison of abundance between two different datasets (control vs. case). [19-20] The profiles of healthy individuals’ microbiomes compared to the profiles of patients with diseases may show a relationship between the microbiome and disease [4]. These profiles may be an indicator if someone is sick [4].

### Diseases.

Gut dysbiosis is when the healthy microbial composition and function of the gut is disrupted. In illnesses like IBS and IBD, gut dysbiosis is when the beneficial bacteria in the gut is reduced. Irritable bowel syndrome is an intestinal disorder that causes bloating, abdominal pain, diarrhea, and constipation. It is believed that the gut microbiome may be one of the causes of irritable bowel syndrome, but the exact cause of IBS is not known. Ulcerative colitis is an inflammatory bowel disease that causes inflammation and ulcers in the inner lining of the colon. The causes of ulcerative colitis are not known but it is hypothesized that it may be caused by an immune response to a virus or bacteria that causes the immune system to attack the digestive tract. Crohn’s disease is a type of inflammatory bowel disease that causes inflammation in the digestive tract, most commonly in the small intestine. It is also hypothesized that this disease may be caused by an immune response to a virus or bacteria that causes the immune system to attack the digestive system. [14, 31]

IBS, UC, and Crohn’s disease usually have decreased or increased relative abundance of the microbiome genus. [18] Literature that has huge findings about the microbiome’s connection to disease may not be completely free of bias due to the newer nature of the research, and the statistical methods involved may cause false discovery rates. [20]

### Databases.

Databases like NCBI and PubMed provide a wealth of information in an easy to use and easy to understand way. [6] MicrobiomeDB helps users describe their data, where the sample was derived, how samples were treated during collection, and how they were processed. [7] The NIH HMP was a project that allowed researchers to have a wealth of resources to run their own research on the microbiome. This database allows the user to search for studies and their data. The Data Repository for Human Gut Microbiota (GMrepo) is a database of human gut metagenomes and aims to capture the full dynamics in the human gut, including abundances, prevalence, associations with disease, and co-occurrences. [22] Amadis is a human microbiome database that connects disease, molecular data, and the ecosystem. [23] mBodyMap is another microbiome database that associates health, disease, and the microbiome. This database shows a body map of where specific bacteria can be located.[24]

An easy-to-use tool to find which microbiome profile can be an indicator of what inflammatory bowel disease would not only be helpful to researchers, but to doctors and other medical fields when dealing with patients. This database is going to be more niche compared to the databases above, and instead of focusing primarily on disease, this database will instead focus primarily on the bacteria itself, and the disease as data. As such, all diseases listed above will be mentioned on each page, with their unhealthy data. Developing this database will allow for the ability to answer two specific aims: What does a healthy profile look like for the five chosen bacteria genera? Is IBS, UC and Crohn’s disease associated with the five chosen bacteria genera?

# Methods and Materials

Before creating a database, an entity-relationship (ER) diagram was built to show the flow of data within the database (Fig. 1). A software requirements specifications (SRS) document was created to help with workflow, design choice, and accessibility features for the database and website. It was decided that this database will fall under an MIT license and is provided “AS IS” without any warranty, with any type of use in mind. A license is provided within the source code on GitHub and as a text document.

Using MySQL and UNO’s Odin Linux server a database was created. The UNO Odin server limited to 1 GB of storage, 80 processes and 100 minutes per job. This database uses SQL version 15.1 for Debian Linux. There are no other software components or applications with which it must coexist. There was a limited ability to use tools with the UNO Odin server, so no such tools were used.

Using HTML, CSS, and PHP a website was created and connected to the database. The website contains a home page, frequently asked questions page, a side navigation to look at all bacteria, for seven pages in total. A search bar was considered at the beginning of the semester but was found to be unnecessary for such a small amount of data and bacteria studied. The website includes accessibility features like a bug report link, the website changes to fit the device it is used on, and all hyperlinks are provided in full text. All citations within the website are anchored to their reference further down the page. The website was designed to be used by naïve and sophisticated users. Naïve users would be those with no bioinformatics background and sophisticated users would be those with a bioinformatics background. The FAQ on the website includes information on the microbiome to help naïve users understand the analysis done within the website, most of this information comes from the introduction above. Each bacteria page describes the taxonomy of the bacteria from NCBI’s taxonomy finder, the healthy relative abundance data, the disease relative abundance data, and the analysis of said data.

Data for the database was obtained from the other microbiome databases described above and from NCBI Bioproject. Data was collected and statistical analysis was done to remove outliers and find relative abundances. Statistical analysis done includes mean, median, standard deviation, first quartile, third quartile, interquartile range, and percent difference using medians.

# Results

## Healthy Data

Using the Human Microbiome Project’s data, a collection of 544 samples collected from the colon, using 16s rRNA sequencing (V1-V3 and V3-V5) was analyzed using MicrobiomeDB to find healthy relative abundances. *Alistipes* had a first quartile of 0.02, a median of 0.05 and a third quartile of 0.08 for relative abundance. It was found that *Bacteroides* has a first quartile of 0.28, a median of 0.48, and a third quartile of 0.66 for relative abundance. *Faecalibacterium* had a first quartile of 0.01, a median of 0.03, and a third quartile of 0.06 for relative abundance. *Parabacteroides* had a first quartile of 0.01, a median of 0.03, and a third quartile of 0.06 for relative abundance. *Ruminococcus* had a first quartile of 0.00, a median of 0.01, and a third quartile of 0.03 for relative abundance. [2, 7-8]

Using the Human Microbiome Project’s data, a collection of 146 samples collected from stool using WSG sequencing and was analyzed using MicrobiomeDB to find healthy relative abundances. *Bacteroides* has a first quartile of 0.42, a median of 0.61, and a third quartile of 0.76 for relative abundance. *Alistipes* has a first quartile of 0.02, a median of 0.06, and a third quartile of 0.11 for relative abundance. *Parabacteroides* has a first quartile of 0.02, a median of 0.05, and a third quartile of 0.08 for relative abundance. *Faecalibacterium* has a first quartile of 0.0, a median of 0.02 and a third quartile of 0.04 for relative abundance. *Ruminococcus* has a first quartile of 0.00, a median of 0.01, and a third quartile of 0.03 for relative abundance. [2, 7-8]

Using Excel, interquartile ranges were found for all genera using the V1-V3 and V3-V5 dataset above. *Alistipes* had an IQR of 0.06. *Bacteroides* had an IQR of 0.38. *Faecalibacterium* had an IQR of 0.05. *Parabacteroides* had an IQR of 0.05. *Ruminococcus* had an IQR of .

Standard deviations were found using Excel as well. *Alistipes* had a standard deviation of 0.067779. *Bacteroides* had a standard deviation of 0.238521. *Faecalibacterium* had a standard deviation of 0.053988. *Parabacteroides* had a standard deviation of 0.042527. *Ruminococcus* had a standard deviation of 0.056998.

Means for healthy data were found using Excel for all genera using the V1-V3 and V3-V5 dataset above. Alistipes had a relative abundance mean of 0.060809. Bacteroides had a relative abundance mean of 0.474819. Faecalibacterium had a relative abundance mean of 0.046274. Parabacteroides had a relative abundance mean of 0.040188. Ruminococcus had a relative abundance mean of 0.026243.

## Disease Data

Using GMrepo relative abundances for IBS were found from 9 different datasets with 1,002 valid runs from stool samples (PRJEB11419, PRJNA268708, PRJNA46339, PRJNA302437, PRJNA392762, PRJNA386442, PRJDB5442, PRJNA373876, and PRJNA524547). *Alistipes* had a relative abundance mean of 0.0341229, median of 0.0211782, and a standard deviation of 0.0383598. *Bacteroides* had a relative abundance mean of 0.2230661, median of 0.1943600, and standard deviation of 0.1628405. *Faecalibacterium* had a relative abundance mean of 0.0576713, median of 0.0440823, and standard deviation of 0.0522330. *Parabacteroides* had a relative abundance mean of 0.0195690, median of 0.0118683, and standard deviation of 0.0250039. *Ruminococcus* had a relative abundance mean of 0.0183178, median of 0.0095594, and standard deviation of 0.0233402. [2, 8, 22, 34-43]

Using GMrepo relative abundances for IBD (Crohn’s Disease and Ulcerative Colitis) from 8 different datasets with 605 valid runs from stool samples using 16S rRNA sequencing (PRJEB11419, PRJNA385949, PRJNA427597, PRJNA296920, PRJNA368966, PRJNA380944, PRJEB7949, and PRJNA284397). *Alistipes* had a relative abundance mean of 0.0332610, median of 0.0162673, and standard deviation of 0.0496512. *Bacteroides* had a relative abundance mean of 0.2218626, median of 0.1953990, and standard deviation of 0.1651884. *Faecalibacterium* had a relative abundance mean of 0.0846279, median of 0.0780259, and standard deviation of 0.0586727. *Parabacteroides* had a relative abundance mean of 0.0324641, median of 0.0166488, and standard deviation of 0.0514218. *Ruminococcus* had a relative abundance mean of 0.0409284, median of 0.0260026, and standard deviation of 0.0498326. [2, 8, 22, 34, 38, 44-47]

Using mBodymap relative abundances for Crohn’s Disease were found from two different datasets with 396 valid runs from the rectum (PRJNA46879 and PRJNA46321). *Alistipes* has a relative abundance mean of 0.0417087, median of 0.0271654, and a standard deviation of 0.0491028. *Bacteroides* had a relative abundance mean of 0.0480192, median of 0.0117297, and a standard deviation of 0.1010771. *Faecalibacterium* had a relative abundance mean of 0.0247488, median of 0.0138938, and a standard deviation of 0.0340558. *Parabacteroides* had a relative abundance mean of 0.0264035, median of 0.0189056, and a standard deviation of 0.0272257. *Ruminococcus* had a relative abundance mean of 0.0110728, median of 0.0039666, and standard deviation of 0.0312523. [2, 8, 24, 32]

Using GMrepo relative abundances for Ulcerative Colitis were found from 18 different datasets with 1,840 valid runs from stool samples using 16S rRNA sequencing (PRJNA5063, PRJEB1220, PRJNA398089, PRJNA389280, PRJNA400072, PRJNA388210, PRJNA285502, PRJNA450340, PRJNA368966, PRJNA438164, PRJNA240346, PRJDB4871, PRJNA318788, PRJNA232056, PRJNA298762, PRJNA233411, PRJNA284397, and PRJNA324147). *Alistipes* has a relative abundance mean of 0.0273334, median of 0.0112793, and a standard deviation of 0.0442919. *Bacteroides* has a relative abundance mean of 0.2952188, median of 0.2491515, and standard deviation of 0.2471365. *Faecalibacterium* has a relative abundance mean of 0.0726973, median of 0.0427561, and standard deviation of 0.0907504. *Parabacteroides* has a relative abundance mean of 0.0284427, median of 0.0125685, and standard deviation of 0.0401865. *Ruminococcus* has a relative abundance mean of 0.0294842, median of 0.0080915, and a standard deviation of 0.0517572. [2, 6, 8, 36, 48-74]

Percent changes for all diseases were found using the healthy data medians and the disease medians as median is less likely to be affected by outliers in the data.

Table 1: Medians of Control (V1-V5) and Disease Relative Abundances

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Medians | Alistipes | Bacteroides | Faecalibacterium | Parabacteroides | Ruminococcus |
| Control | 0.05 | 0.48 | 0.03 | 0.03 | 0.01 |
| IBS | 0.0211782 | 0.1943600 | 0.0440823 | 0.0118683 | 0.0095594 |
| IBD | 0.0162673 | 0.1953990 | 0.0780259 | 0.0166488 | 0.0260026 |
| CD | 0.0271654 | 0.0117297 | 0.0138938 | 0.0189056 | 0.0039666 |
| UC | 0.0112793 | 0.2491515 | 0.042756 | 0.0125685 | 0.0080915 |

Table 2: Percent Change (Control vs Disease)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Percent Change (Control vs Disease) | Alistipes | Bacteroides | Faecalibacterium | Parabacteroides | Ruminococcus |
| IBS | -57.644 | -59.5083 | 46.941 | -60.439 | -4.406 |
| IBD | -67.4654 | -59.2919 | 160.0863 | -44.504 | 160.026 |
| CD | -45.6692 | -97.5563 | -53.6873 | -36.9813 | -60.334 |
| UC | -77.4414 | -48.0934375 | 42.52033 | -58.105 | -19.085 |

# Discussion

The findings of this study support that *Alistipes*, *Bacteroides*, *Faecalibacterium*, *Parabacteroides*, and *Ruminococcus* are associated with IBS, IBD, CD, and UC. As shown above, the relative abundances of these bacteria genus change when healthy is compared to disease (Table 1). *Alistipes* lost relative abundance when compared to the healthy relative abundance for all diseases. Bacteroides lost relative abundance when compared to the healthy relative abundance for all diseases. Faecalibacterium lost relative abundance in CD when compared to the healthy relative abundance but increased in IBS, IBD and UC. Parabacteroides lost relative abundance when compared to healthy relative abundances. Ruminococcus lost relative abundance when compared to healthy relative abundances for all diseases except IBD. (Table 2)

# Conclusion

It is important to understand how bacteria in the gut changes when exposed to disease. Alistipes, Bacteroides, Faecalibacterium, Parabacteroides, and Ruminococcus have been found within the intestines and do change in relative abundance in people with intestinal diseases. Lack of normal relative abundances may correlate to the symptoms experienced by those with intestinal diseases. This may mean that the gut microbiome may not function properly within the intestines.

Trying to create a database based solely on literature and curated data was not possible due to lack of responsibility for reproducibility of datasets. The accuracy of statistics done with un-curated data can be called into question. Regardless of these challenges, this project provides a proof-of-concept database and web interface demonstrating the potential for analysis of aggregation of the microbiome data in disease states should these challenges be addressed.

# References

1. Durack, J., & Lynch, S. V. (2019, January 7). *The gut microbiome: Relationships with disease and opportunities for therapy*. The Journal of experimental medicine. Retrieved September 21, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6314516/
2. Human Microbiome Project Consortium. (2012, June 13). *Structure, function and diversity of the healthy human microbiome*. Nature. Retrieved September 21, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3564958/
3. Jia, W., Whitehead, R. N., Griffiths, L., Dawson, C., Waring, R. H., Ramsden, D. B., Hunter, J. O., & Cole, J. A. (2010, September 1). *Is the abundance of faecalibacterium prausnitzii relevant to crohn's disease?* FEMS microbiology letters. Retrieved September 21, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2962807/
4. Lloyd-Price, J., Mahurkar, A., Rahnavard, G., Crabtree, J., Orvis, J., Hall, A. B., Brady, A., Creasy, H. H., McCracken, C., Giglio, M. G., McDonald, D., Franzosa, E. A., Knight, R., White, O., & Huttenhower, C. (2017, October 5). *Strains, functions and dynamics in the expanded human microbiome project*. Nature. Retrieved September 21, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5831082/
5. Rinninella, E., Raoul, P., Cintoni, M., Franceschi, F., Miggiano, G. A. D., Gasbarrini, A., & Mele, M. C. (2019, January 10). *What is the healthy gut microbiota composition? A changing ecosystem across age, environment, diet, and diseases*. Microorganisms. Retrieved September 21, 2022, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6351938/#:~:text=The%20dominant%20gut%20microbial%20phyla,representing%2090%25%20of%20gut%20microbiota>
6. National Center for Biotechnology Information (NCBI)[Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; [1988] – [cited 2022 October 17]. Available from: <https://www.ncbi.nlm.nih.gov/>
7. Oliveira, F. S., Brestelli, J., Cade, S., Zheng, J., Iodice, J., Fischer, S., Aurrecoechea, C., Kissinger, J. C., Brunk, B. P., Stoeckert, C. J., Fernandes, G. R., Roos, D. S., & Beiting, D. P. (2017). MicrobiomeDB: A Systems Biology platform for integrating, mining and analyzing microbiome experiments. *Nucleic Acids Research*, *46*(D1). <https://doi.org/10.1093/nar/gkx1027>
8. Human Microbiome Project Consortium (2012). A framework for human microbiome research. *Nature*, *486*(7402), 215–221. <https://doi.org/10.1038/nature11209>
9. Wexler H. M. (2007). Bacteroides: the good, the bad, and the nitty-gritty. *Clinical microbiology reviews*, *20*(4), 593–621. <https://doi.org/10.1128/CMR.00008-07>
10. Parsaei, M., Sarafraz, N., Moaddab, S. Y., & Ebrahimzadeh Leylabadlo, H. (2021). The importance of *Faecalibacterium prausnitzii* in human health and diseases. *New microbes and new infections*, *43*, 100928. <https://doi.org/10.1016/j.nmni.2021.100928>
11. Sakamoto, M., & Benno, Y. (2006). Reclassification of Bacteroides distasonis, Bacteroides goldsteinii and Bacteroides merdae as Parabacteroides distasonis gen. nov., comb. nov., Parabacteroides goldsteinii comb. nov. and Parabacteroides merdae comb. nov. *International journal of systematic and evolutionary microbiology*, *56*(Pt 7), 1599–1605. <https://doi.org/10.1099/ijs.0.64192-0>
12. Ezeji, J. C., Sarikonda, D. K., Hopperton, A., Erkkila, H. L., Cohen, D. E., Martinez, S. P., Cominelli, F., Kuwahara, T., Dichosa, A., Good, C. E., Jacobs, M. R., Khoretonenko, M., Veloo, A., & Rodriguez-Palacios, A. (2021). *Parabacteroides distasonis*: intriguing aerotolerant gut anaerobe with emerging antimicrobial resistance and pathogenic and probiotic roles in human health. *Gut microbes*, *13*(1), 1922241. <https://doi.org/10.1080/19490976.2021.1922241>
13. Parker, B. J., Wearsch, P. A., Veloo, A., & Rodriguez-Palacios, A. (2020). The Genus *Alistipes*: Gut Bacteria With Emerging Implications to Inflammation, Cancer, and Mental Health. *Frontiers in immunology*, *11*, 906. <https://doi.org/10.3389/fimmu.2020.00906>
14. Nagao-Kitamoto, H., & Kamada, N. (2017). Host-microbial Cross-talk in Inflammatory Bowel Disease. *Immune network*, *17*(1), 1–12. <https://doi.org/10.4110/in.2017.17.1.1>
15. Henke, M. T., Kenny, D. J., Cassilly, C. D., Vlamakis, H., Xavier, R. J., & Clardy, J. (2019). *Ruminococcus gnavus*, a member of the human gut microbiome associated with Crohn's disease, produces an inflammatory polysaccharide. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(26), 12672–12677. <https://doi.org/10.1073/pnas.1904099116>
16. Rasmussen, B. A., Bush, K., & Tally, F. P. (1993). Antimicrobial resistance in Bacteroides. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, *16 Suppl 4*, S390–S400. <https://doi.org/10.1093/clinids/16.supplement_4.s390>
17. Jasemi, S., Emaneini, M., Ahmadinejad, Z., Fazeli, M. S., Sechi, L. A., Sadeghpour Heravi, F., & Feizabadi, M. M. (2021). Antibiotic resistance pattern of Bacteroides fragilis isolated from clinical and colorectal specimens. *Annals of clinical microbiology and antimicrobials*, *20*(1), 27. <https://doi.org/10.1186/s12941-021-00435-w>
18. Gevers, D., Kugathasan, S., Denson, L. A., Vázquez-Baeza, Y., Van Treuren, W., Ren, B., Schwager, E., Knights, D., Song, S. J., Yassour, M., Morgan, X. C., Kostic, A. D., Luo, C., González, A., McDonald, D., Haberman, Y., Walters, T., Baker, S., Rosh, J., Stephens, M., … Xavier, R. J. (2014). The treatment-naive microbiome in new-onset Crohn's disease. *Cell host & microbe*, *15*(3), 382–392. <https://doi.org/10.1016/j.chom.2014.02.005>
19. Lin, H., Peddada, S.D. Analysis of microbial compositions: a review of normalization and differential abundance analysis. *npj Biofilms Microbiomes* **6**, 60 (2020). <https://doi.org/10.1038/s41522-020-00160-w>
20. Stijn Hawinkel, Federico Mattiello, Luc Bijnens, Olivier Thas, A broken promise: microbiome differential abundance methods do not control the false discovery rate, *Briefings in Bioinformatics*, Volume 20, Issue 1, January 2019, Pages 210–221, <https://doi.org/10.1093/bib/bbx104>
21. Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., & Egozcue, J. J. (2017, November 15). *Microbiome datasets are compositional: And this is not optional*. Frontiers. Retrieved October 27, 2022, from https://www.frontiersin.org/articles/10.3389/fmicb.2017.02224/full
22. Die Dai, Jiaying Zhu, Chuqing Sun, Min Li, Jinxin Liu, Sicheng Wu, Kang Ning, Li-jie He, Xing-Ming Zhao, Wei-Hua Chen, GMrepo v2: a curated human gut microbiome database with special focus on disease markers and cross-dataset comparison, *Nucleic Acids Research*, Volume 50, Issue D1, 7 January 2022, Pages D777–D784, <https://doi.org/10.1093/nar/gkab1019>
23. Comprehensive Database for Association Between Microbiota and Disease. Frontiers in Physiology, 12, 697059. <https://www.frontiersin.org/articles/10.3389/fphys.2021.697059/full>
24. Jin, H., Hu, G., Sun, C., Duan, Y., Zhang, Z., Liu, Z., Zhao, X. M., & Chen, W. H. (2022). mBodyMap: a curated database for microbes across human body and their associations with health and diseases. *Nucleic acids research*, *50*(D1), D808–D816. <https://doi.org/10.1093/nar/gkab973>
25. Janda, J. M., & Abbott, S. L. (2007). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *Journal of clinical microbiology*, *45*(9), 2761–2764. <https://doi.org/10.1128/JCM.01228-07>
26. Johnson, J.S., Spakowicz, D.J., Hong, BY. *et al.* Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun* **10**, 5029 (2019). <https://doi.org/10.1038/s41467-019-13036-1>
27. Ranjan, R., Rani, A., Metwally, A., McGee, H. S., & Perkins, D. L. (2016). Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochemical and biophysical research communications*, *469*(4), 967–977. <https://doi.org/10.1016/j.bbrc.2015.12.083>
28. Li, Y., Jin, Y., Zhang, J. *et al.* Recovery of human gut microbiota genomes with third-generation sequencing. *Cell Death Dis* **12**, 569 (2021). <https://doi.org/10.1038/s41419-021-03829-y>
29. Cao, Y., Shen, J., & Ran, Z. H. (2014). Association between Faecalibacterium prausnitzii Reduction and Inflammatory Bowel Disease: A Meta-Analysis and Systematic Review of the Literature. *Gastroenterology research and practice*, *2014*, 872725. <https://doi.org/10.1155/2014/872725>
30. La Reau, A. J., Meier-Kolthoff, J. P., & Suen, G. (2016). Sequence-based analysis of the genus *Ruminococcus* resolves its phylogeny and reveals strong host association. *Microbial genomics*, *2*(12), e000099. <https://doi.org/10.1099/mgen.0.000099>
31. DeGruttola, A. K., Low, D., Mizoguchi, A., & Mizoguchi, E. (2016). Current Understanding of Dysbiosis in Disease in Human and Animal Models. *Inflammatory bowel diseases*, *22*(5), 1137–1150. <https://doi.org/10.1097/MIB.0000000000000750>
32. Noecker C, Eng A, Srinivasan S, Theriot CM, Young VB, Jansson JK, Fredricks DN, Borenstein E. Metabolic Model-Based Integration of Microbiome Taxonomic and Metabolomic Profiles Elucidates Mechanistic Links between Ecological and Metabolic Variation. mSystems. 2016 Jan-Feb;1(1):e00013-15. doi: 10.1128/mSystems.00013-15. Epub 2016 Jan 19. PMID: 27239563; PMCID: PMC4883586.
33. PRJNA46879 [human metagenome (ID 46879) - BioProject - NCBI (nih.gov)](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA46879) (CD)
34. McDonald, D., Hyde, E., Debelius, J. W., Morton, J. T., Gonzalez, A., Ackermann, G., Aksenov, A. A., Behsaz, B., Brennan, C., Chen, Y., DeRight Goldasich, L., Dorrestein, P. C., Dunn, R. R., Fahimipour, A. K., Gaffney, J., Gilbert, J. A., Gogul, G., Green, J. L., Hugenholtz, P., Humphrey, G., … Knight, R. (2018). American Gut: an Open Platform for Citizen Science Microbiome Research. *mSystems*, *3*(3), e00031-18. <https://doi.org/10.1128/mSystems.00031-18>
35. Pozuelo, M., Panda, S., Santiago, A., Mendez, S., Accarino, A., Santos, J., Guarner, F., Azpiroz, F., & Manichanh, C. (2015). Reduction of butyrate- and methane-producing microorganisms in patients with Irritable Bowel Syndrome. *Scientific reports*, *5*, 12693. <https://doi.org/10.1038/srep12693>
36. NIH HMP Working Group, Peterson, J., Garges, S., Giovanni, M., McInnes, P., Wang, L., Schloss, J. A., Bonazzi, V., McEwen, J. E., Wetterstrand, K. A., Deal, C., Baker, C. C., Di Francesco, V., Howcroft, T. K., Karp, R. W., Lunsford, R. D., Wellington, C. R., Belachew, T., Wright, M., Giblin, C., … Guyer, M. (2009). The NIH Human Microbiome Project. *Genome research*, *19*(12), 2317–2323. <https://doi.org/10.1101/gr.096651.109>
37. Mizuno, S., Masaoka, T., Naganuma, M., Kishimoto, T., Kitazawa, M., Kurokawa, S., Nakashima, M., Takeshita, K., Suda, W., Mimura, M., Hattori, M., & Kanai, T. (2017). Bifidobacterium-Rich Fecal Donor May Be a Positive Predictor for Successful Fecal Microbiota Transplantation in Patients with Irritable Bowel Syndrome. *Digestion*, *96*(1), 29–38. <https://doi.org/10.1159/000471919>
38. Sicheng Wu, Chuqing Sun, Yanze Li, Teng Wang, Longhao Jia, Senying Lai, Yaling Yang, Pengyu Luo, Die Dai, Yong-Qing Yang, Qibin Luo, Na L Gao, Kang Ning, Li-jie He, Xing-Ming Zhao, Wei-Hua Chen, GMrepo: a database of curated and consistently annotated human gut metagenomes, *Nucleic Acids Research*, Volume 48, Issue D1, 08 January 2020, Pages D545–D553, <https://doi.org/10.1093/nar/gkz764>
39. PRJNA302437 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA302437> (IBD)
40. PRJNA392762 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA392762> (IBD)
41. PRJNA386442 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA386442> (IBD)
42. PRJNA373876 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA373876> (IBD)
43. PRJNA524547 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA524547> (IBD)
44. Hall, A. B., Yassour, M., Sauk, J., Garner, A., Jiang, X., Arthur, T., Lagoudas, G. K., Vatanen, T., Fornelos, N., Wilson, R., Bertha, M., Cohen, M., Garber, J., Khalili, H., Gevers, D., Ananthakrishnan, A. N., Kugathasan, S., Lander, E. S., Blainey, P., Vlamakis, H., … Huttenhower, C. (2017). A novel Ruminococcus gnavus clade enriched in inflammatory bowel disease patients. *Genome medicine*, *9*(1), 103. <https://doi.org/10.1186/s13073-017-0490-5>
45. Staley, C., Hamilton, M. J., Vaughn, B. P., Graiziger, C. T., Newman, K. M., Kabage, A. J., Sadowsky, M. J., & Khoruts, A. (2017). Successful Resolution of Recurrent Clostridium difficile Infection using Freeze-Dried, Encapsulated Fecal Microbiota; Pragmatic Cohort Study. *The American journal of gastroenterology*, *112*(6), 940–947. <https://doi.org/10.1038/ajg.2017.6>
46. Goyal, A., Yeh, A., Bush, B. R., Firek, B. A., Siebold, L. M., Rogers, M. B., Kufen, A. D., & Morowitz, M. J. (2018). Safety, Clinical Response, and Microbiome Findings Following Fecal Microbiota Transplant in Children With Inflammatory Bowel Disease. *Inflammatory bowel diseases*, *24*(2), 410–421. <https://doi.org/10.1093/ibd/izx035>
47. Harris, R. A., Shah, R., Hollister, E. B., Tronstad, R. R., Hovdenak, N., Szigeti, R., Versalovic, J., & Kellermayer, R. (2016). Colonic Mucosal Epigenome and Microbiome Development in Children and Adolescents. *Journal of immunology research*, *2016*, 9170162. <https://doi.org/10.1155/2016/9170162>
48. Nielsen, H. B., Almeida, M., Juncker, A. S., Rasmussen, S., Li, J., Sunagawa, S., Plichta, D. R., Gautier, L., Pedersen, A. G., Le Chatelier, E., Pelletier, E., Bonde, I., Nielsen, T., Manichanh, C., Arumugam, M., Batto, J. M., Quintanilha Dos Santos, M. B., Blom, N., Borruel, N., Burgdorf, K. S., … MetaHIT Consortium (2014). Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nature biotechnology*, *32*(8), 822–828. <https://doi.org/10.1038/nbt.2939>
49. Lloyd-Price, J., Arze, C., Ananthakrishnan, A. N., Schirmer, M., Avila-Pacheco, J., Poon, T. W., Andrews, E., Ajami, N. J., Bonham, K. S., Brislawn, C. J., Casero, D., Courtney, H., Gonzalez, A., Graeber, T. G., Hall, A. B., Lake, K., Landers, C. J., Mallick, H., Plichta, D. R., Prasad, M., … Huttenhower, C. (2019). Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature*, *569*(7758), 655–662. <https://doi.org/10.1038/s41586-019-1237-9>
50. Franzosa, E. A., Sirota-Madi, A., Avila-Pacheco, J., Fornelos, N., Haiser, H. J., Reinker, S., Vatanen, T., Hall, A. B., Mallick, H., McIver, L. J., Sauk, J. S., Wilson, R. G., Stevens, B. W., Scott, J. M., Pierce, K., Deik, A. A., Bullock, K., Imhann, F., Porter, J. A., Zhernakova, A., … Xavier, R. J. (2019). Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nature microbiology*, *4*(2), 293–305. <https://doi.org/10.1038/s41564-018-0306-4>
51. Damman, C. J., Brittnacher, M. J., Westerhoff, M., Hayden, H. S., Radey, M., Hager, K. R., Marquis, S. R., Miller, S. I., & Zisman, T. L. (2015). Low Level Engraftment and Improvement following a Single Colonoscopic Administration of Fecal Microbiota to Patients with Ulcerative Colitis. *PloS one*, *10*(8), e0133925. <https://doi.org/10.1371/journal.pone.0133925>
52. Forbes, J. D., Chen, C. Y., Knox, N. C., Marrie, R. A., El-Gabalawy, H., de Kievit, T., Alfa, M., Bernstein, C. N., & Van Domselaar, G. (2018). A comparative study of the gut microbiota in immune-mediated inflammatory diseases-does a common dysbiosis exist?. *Microbiome*, *6*(1), 221. <https://doi.org/10.1186/s40168-018-0603-4>
53. Del Chierico, F., Gnani, D., Vernocchi, P., Petrucca, A., Alisi, A., Dallapiccola, B., Nobili, V., & Lorenza, P. (2014). Meta-omic platforms to assist in the understanding of NAFLD gut microbiota alterations: tools and applications. *International journal of molecular sciences*, *15*(1), 684–711. <https://doi.org/10.3390/ijms15010684>
54. Chehoud, C., Dryga, A., Hwang, Y., Nagy-Szakal, D., Hollister, E. B., Luna, R. A
55. ., Versalovic, J., Kellermayer, R., & Bushman, F. D. (2016). Transfer of Viral Communities between Human Individuals during Fecal Microbiota Transplantation. *mBio*, *7*(2), e00322. <https://doi.org/10.1128/mBio.00322-16>
56. Harris, R. A., Shah, R., Hollister, E. B., Tronstad, R. R., Hovdenak, N., Szigeti, R., Versalovic, J., & Kellermayer, R. (2016). Colonic Mucosal Epigenome and Microbiome Development in Children and Adolescents. *Journal of immunology research*, *2016*, 9170162. <https://doi.org/10.1155/2016/9170162>
57. PRJNA50637 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA50637> (UC)
58. PRJEB1220 <https://www.ncbi.nlm.nih.gov/bioproject/PRJEB1220> (UC)
59. PRJNA398089 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA398089> (UC)
60. PRJNA389280 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA389280> (UC)
61. PRJNA400072 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA400072> (UC)
62. PRJNA388210 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA388210> (UC)
63. PRJNA285502 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA285502> (UC)
64. PRJNA450340 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA450340> (UC)
65. PRJNA368966 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA368966> (UC)
66. PRJNA438164 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA438164> (UC)
67. PRJNA240346 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA240346> (UC)
68. PRJDB4871 <https://www.ncbi.nlm.nih.gov/bioproject/PRJDB4871> (UC)
69. PRJNA318788 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA318788> (UC)
70. PRJNA232056 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA232056> (UC)
71. PRJNA298762 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA298762> (UC)
72. PRJNA233411 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA233411> (UC)
73. PRJNA284397 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA284397> (UC)
74. PRJNA324147 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA324147> (UC)

# Links

**Link to website** (requires UNO VPN to access off campus): odin.unomaha.edu/~eledford/course\_project\_4560.php

**Link to GitHub:**