5-2021

Identification and Characterization of Butyrate-Producing Species in the Human Gut Microbiome

Grace Maline
gmaline@unomaha.edu

Identification and Characterization of Butyrate-Producing Species in the Human Gut Microbiome

Grace Maline

***Abstract*—Inflammatory Bowel Diseases (IBD) including ulcerative colitis, Crohn's disease, and indeterminate colitis are increasingly common conditions that places a high physical and financial burden on individuals and global healthcare systems. Though treatments exist for these conditions, their unpredictable nature and causation make them difficult to manage consistently across the variety of IBD patients. Additionally, many of these treatments come with undesirable side effects or modes of delivery. Therefore, we consider the use of Short Chain Fatty Acids (SCFAs) such as butyrate, whose affects in the human gut include decreased inflammation and decreased risk of colorectal cancer. As butyrate is produced primarily by microbial species within the human gut microbiome, this project's aim was to identify butyrate producing bacteria via metagenomic analysis based on the presence of genes that are essential to pathways of butyrate production. In total, 73 different potentially butyrate-producing species were identified and characterized. Species identified include both gram-positive and gram-negative anaerobic bacteria belonging to the classes Clostridia, Fusobacteriia, Negativicutes, Bacilli, and Bacteroidia. Of these, Clostridia was the highest scoring class of bacteria.**

***Index Terms*—IBD - Inflammatory Bowel Disease, SCFA - Short Chain Fatty Acid**

I. Introduction

DISEASES involving inflammation of the digestive tract are a highly prevalent and debilitating reality for many individuals. Typically, these diseases are referred to as a collection of conditions known as Inflammatory Bowel Disease (IBD). Conditions that fall under the umbrella term of IBD include Crohn's disease, ulcerative colitis, and indeterminate colitis [1]. In 2017, around 6.8 million cases of IBD were reported worldwide, which is almost twice the number of cases reported just three decades prior [2]. As the prevalence of this disease increases, so does its physical and financial burden both on those who suffer from these conditions and to global healthcare systems. Therefore, it is important that special attention is paid to IBD to establish non-expensive and effective treatments as well as preventative measures.

A. Motivation

IBD conditions are characterized by chronic, persistent relapses of inflammation in the lower digestive tract [3]. They are commonly associated with abdominal pain, rectal bleeding, weight loss, and perianal disease. As a result, patients with IBD report a large amount of discomfort and a substantial reduction in quality of life [2]. Unfortunately, the causes of IBD are often unknown and are overall poorly understand. However, it is suspected that individuals contain genetic predispositions to immune triggers that result in the inflammatory response [3]. It is often speculated that these triggers may come in the form of environmental factors or as a result of the interaction of host and microbial cells in the gut microbiome [4].

Because of their high prevalence and mysterious origins, much work in recent years has been dedicated to studying these diseases, their causes, diagnosis, and treatments. It has been found that induction of inflammation in the gut is often related to two well-known pathways, namely the Nuclear Factor-kappa B (NF-kB) pathway and the Mitogen-activated protein kinase (MAPK) pathway [5]. Though targeting pathway intermediates such as TNF-alpha with anti-inflammatory drugs and monoclonal antibodies can reduce gut inflammation, drugs that are able to do so often come with significant challenges [6]. Because they are non-specific to cell type, they produce undesirable side effects. Additionally, administration of the treatments may be invasive or uncomfortable [6]. Therefore, alternative prevention and treatments as they relate to intervention in these pathways are actively sought and studied.

Potential alternatives include Short Chain Fatty Acids (SCFA), which have shown increasing popularity in recent years. SCFAs such as butyrate, propionate, and acetate can be shown to correlate with decreased gut inflammation and decreased risk of colorectal cancer [7]. This is because SCFAs are the preferred energy source for colonocytes [8]. Butyrate in particular is a widely studied SCFA and accounts for about 70 percent of the energy supply for colonocytes, and inadequate supplies of butyrate in the gut are closely tied to inflammation and increased risk for cancer [7]. Butyrate acts by inhibition of NF-kB related inflammatory pathways in epithelial cells as well as by decreasing release of inflammatory cytokines [8].

Because of these anti-inflammatory effects, butyrate, among other SCFAs can be utilized as dietary supplements for gut health. However, there may be limitations to the effectiveness and delivery of direct oral supplementation of butyrate. Orally administrated butyrate products are often difficult to palate due to intense odor and poor taste [9]. Especially for pediatric patients, this can make it difficult or impossible to facilitate regular daily doses [9]. Therefore, more work is needed to investigate the potential for alternative methods of butyrate delivery to cultivate its anti-inflammatory benefits.

Though less explored than direct oral supplementation and prebiotic supplementation for existing flora, the possibility for

a probiotic solution to this problem is very real. For this to be cultivated, deeper exploration of the individual species in the gut is needed both to identify the butyrate-producing bacteria in the gut and to characterize them by their potential to serve as probiotics [7]. Thus, this project will aim to make steps toward that possibility by identifying and characterizing microbial species in the human gut that could be butyrate-producers based on the presence of genes necessary for its production.

B. Hypothesis

The objective of this project is to identify gut microbial species that have butyrate production capabilities. Their ability to produce butyrate will be based on the presence of pathway genes within their genomes. The expected results will be a list of human gut microbial species that possess such criteria.

In doing so, this project can provide a broad view for further exploration of species that may be probiotic candidates. Future in vitro and in vivo studies to further explore the resulting microbial species could be established based on the findings of this project. After identification, the resulting species will be further evaluated by examining classes of bacteria by phylogenetic analysis and ANOVA analysis.

It is currently speculated that butyrate-producing bacteria are typically anaerobic, gram-positive firmicutes – mostly belonging to clostridial clusters IV and XIVa [10]. It is hypothesized that the results of these analyses will also fall within these classes of bacteria.

C. Background

Progress has already been made to begin identifying butyrate-producing species. Early studies were limited to in vitro observations of butyrate production and acetate utilization by species collected from fecal samples [11]. Enzymatic assays and biochemical assays using primers and qPCR to target important late pathway enzymes were popular methods of studying butyrate production and identifying species [11], [12]. However, these studies are limited by resources and scope. These limitations can be supplemented by sequence analysis and in silico approaches as proposed for this project.

An example of a successful in silico approach couples qPCR enzyme analysis with metagenomic searches for the the butyryl-CoA:acetate CoA-transferase gene sequence in a handful of bacterial species suspected of butyrate production. Both methods arrived at primarily the same conclusions and same identified species, showing that both are adequate for prediction of butyrate producing species [12], [13]. Therefore, a widescale metagenomic analysis using pathway genes could provide valuable insight into more microbes with these capabilities.

To identify bacteria capable of producing butyrate, it is important to understand the pathways that bring about its production. Butyrate is typically produced in a pathway that branches from glucose fermentation during glycolysis [14]. Starting with acetyl-coA, a product of glucose fermentation, there are two ways to obtain butyrate. These two pathways shown in **Figure 1** stem from a single pathway diverging at its final steps. These diverging branches are known as the Butyrate Kinase pathway and the Butyryl-CoA:acetate CoA transferase pathway [15]. Of these, the butyryl-CoA:acetate CoA-transferase pathway is suspected to be more prevalent in microbial species. Therefore, the butyryl-CoA:acetate CoA-transferase gene, encoding one of the pathway's key enzymes, has been shown to be one of importance in identifying species that utilize this method of butyrate production [15]. Additionally, butyrate producing pathways with their essential enzymes can also be observed in their overall context in KEGG as a part of the Butanoate Metabolism pathway, (pathway id: ko00650).

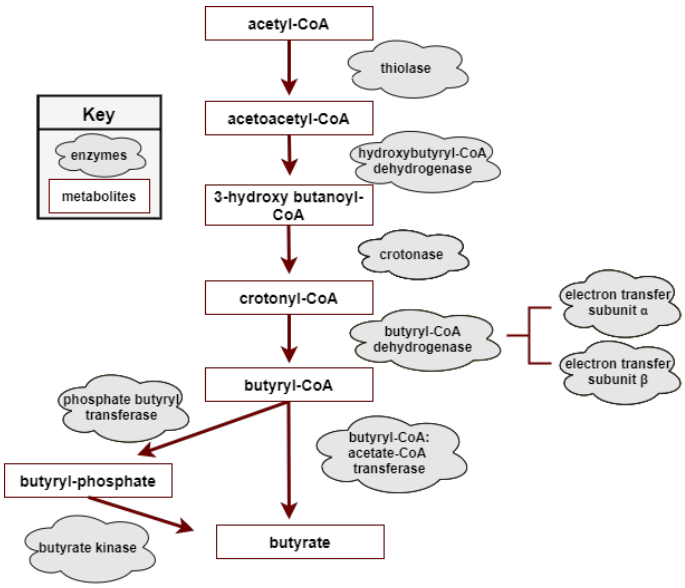


Fig. 1. **Butyrate Production Pathway Summary**. This figure illustrates the two common pathways of butyrate production in microbial gut species, starting with acetyl-CoA. Each gray bubble represents a pathway enzyme, and each white box represents a metabolite intermediate of the pathway. The pathway diverges in the final steps into the Butyryl-CoA:acetate CoA Transferase Pathway and the Butyrate Kinase Pathway.

II. Materials and Methods

The activities for this project can be broken down into multiple steps for data collection, data generation, and analysis. These activities and the flow of data associated with them are summarized in **Figure 2**. All code for these activities can be found in a public github repository (https://github.com/gmaline/SeniorProject.git).

A. Data Sources

Sources of data for analysis will include whole genome sequences of microbial species of the human gut microbiome that have been made publicly available by the Human Microbiome Project (HMP) [16]. There are 457 microbial species from the gastrointestinal tract available at https://www.hmpdacc.org/hmp/HMRGD/. Much work has been put into sequencing and annotating genome data from HMP since its development [17]. Therefore, the quality of the reference genomes will be sufficient for this project.

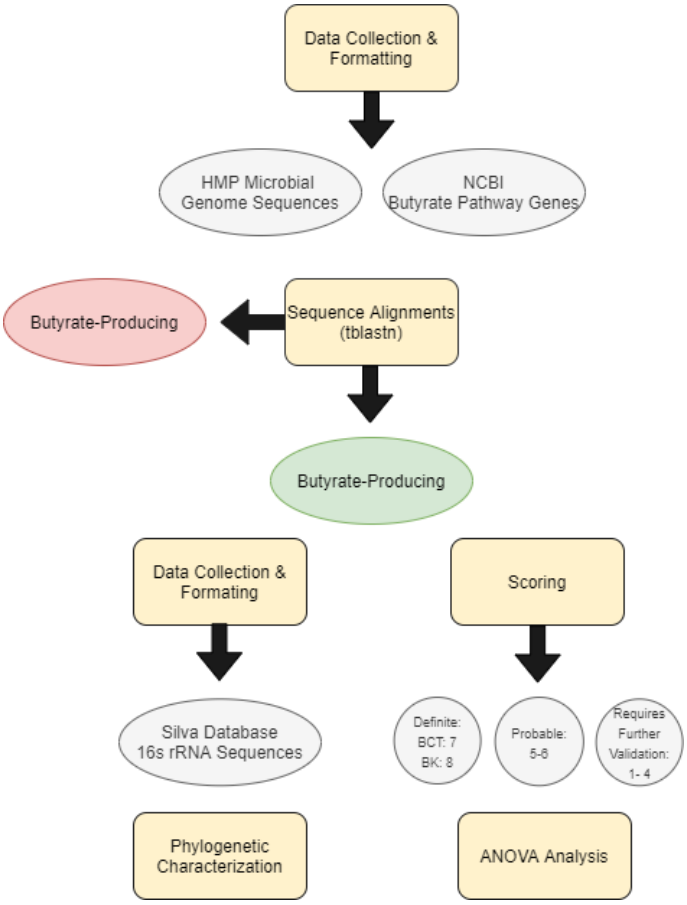


Fig. 2. **Flow Diagram for Project Execution.** This diagram illustrates step by step the process for completing this project as broken into 6 main tasks: 2 rounds of Data Collection, Sequence Alignments, Scoring, ANOVA Analysis and Phylogenetic Characterization.

In addition, sequences of the genes from the butyrate production pathways will be pulled and investigated against the genomes from HMP. The genes identified as important in the butyryl-CoA:acetate CoA-transferase pathway and the butyrate kinase pathway are summarized in **Table 1**.

Finally, for the phylogenetic analysis of the resulting species, 16s rRNA sequences will be obtained from Silva (https://www.arb-silva.de/). Silva is a high-quality ribosomal RNA database supported by de.NBI [18].

B. Collecting Microbial Genomes

From the HMP1 data portal, the CDS of 457 different microbial species of the gastrointestinal tract were batch downloaded in fasta format. To prep the sequences for further analysis, some pre-processing was performed. Using the SeqIO python library from Biopython, the fasta records in the larger CDS file were read in [19]. Regular expression pattern searching with the re python library allowed for the identification of species name from the header of the fasta record. Using the species name from the header, CDS sequences were separated into individual fasta files for each species. This allowed for the easy construction of blastdbs for each species for the subsequent alignment analyses.

Data Sources for Pathway Genes

Enzyme Name	Pathway(s)	Source
Thiolase	Both	NCBI: DQ987697.1
β-hydroxybutyryl-CoA dehydrogenase	Both	NCBI: DQ987697.1
Crotonase	Both	NCBI: DQ987697.1
butyryl-CoA dehydrogenase	Both	NCBI: DQ987697.1
electron transfer protein α	Both	NCBI: DQ987697.1
electron transfer protein β	Both	NCBI: DQ987697.1
Butyral-CoaA:acetate CoA transferase	Butyryl-CoA: acetate CoA Transferase Pathway	NCBI: DQ987697.1
Phosphate butyryltransferase	Butyrate Kinase Pathway	NCBI: WP_003420701.1
Butyrate Kinase	Butyrate Kinase Pathway	NCBI: WP_003722496.1

Table 1. Butyrate Production Pathway Enzymes. This table summarizes the genes involved, the pathways to which they belong, and the source from which their data will be pulled.

C. Collecting Pathway Enzyme Sequences

Additionally, the translated gene sequences of each enzyme were obtained from NCBI and processed such that each protein sequence was placed in its own fasta file with an identifiable header for subsequent analyses. Protein sequence fasta headers were as follows: thiolase, beta hydroxybutyryl-CoA dehydrogenase, crotonase, butyryl-CoA dehydrogenase, electron transfer flavoprotein alpha-subunit, electron transfer flavoprotein beta-subunit, butyryl-COA-transferase, phoasephate butyryltransferase, butyrate kinase.

D. Sequence Alignments

For each of the species CDS fasta files, a temporary nucleotide blastdb was built. Commands from the BLAST Command Line Application from NCBI were run on command line from within a python script using the os python library. For each species database built, a tblastn alignment of each protein sequence was performed using the python library NcbitblastnCommandline from Bio.Blast.Applications [19]. The 4,000+ alignment results files were stored in their own folder in an xml format.

Pseudocode for Alignments

```
1. FUNCTION findButyrateSpecies(Proteins, HMP):
2. butyrate_species ←empty list
3. FOR microbe in HMP:
4.         temporary_blastdb ← makeblastdb(microbe)
5.         FOR protein in Proteins:
6.                 alignment ← tblastn(query=protein, database=temporary_blastdb)
7.                 IF alignment.score >= 60%:
8.                         butyrate_species.add(microbe)
9. RETURN butyrate_species
```

Fig. 3. **Pseudocode for Alignments.** The pseudocode shown describes a high level overview of how the sequence alignments were performed for the purpose of this analysis.

E. Scoring

To parse the xml results, the NCBIXML python library was used from Bio.Blast [19]. Significant hits were stored as csv output based on high scoring pairs with an e-value less than .01. The results of parsing the xml files took note of species, protein, percent identity, percent query coverage, and the aligned query and subject sequences. After these results were obtained, the resulting csv files were easily parsed and scored such that only those species with one or more hits at or above a 60 percent identity level were kept.

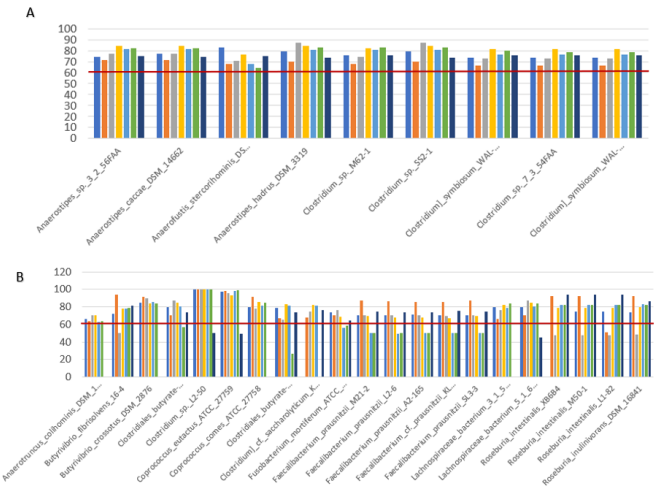
F. Phylogenetic and ANOVA Analyses

For comparison of the resulting potential butyrate-producing species, 16s ribosomal RNA sequences were obtained from Silva and compiled with identifiable headers [18]. A multiple sequence alignment and newick tree format were obtained from Clustal Omega [20]. The phylogenetic tree was visualized using Iroki, an automatic customization and visualization of phylogenetic trees tool [21].

In addition to creating the phylogenetic tree, the resulting scored species were labeled with taxon information. Breaking them down into classes of bacteria an ANOVA analysis of the scores was performed to look for differences in the strength of the representation of butyrate producing species by class of bacteria.

III. Results

At a 60 percent identity threshold, 73 species were found to have at least one of the pathway enzymes in their genomes. Though some species had partial hits for the butyrate kinase pathway, only the Butyryl-CoA:acetate CoA Transferase pathway was represented in the results. Only 9 species had all of the pathway enzymes present at the specified identity threshold. However, many of the species that did not make the cut do have some identity with the pathway enzymes below and/or around the threshold. This is best seen by considering **Figures 4A-D** in which the identities for each protein are plotted with the threshold shown as a red horizontal line at 60 percent identity.



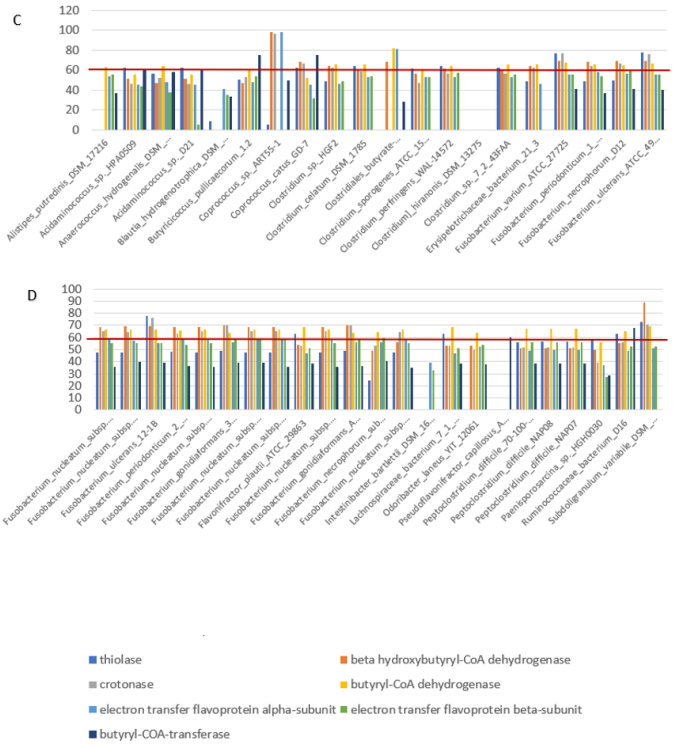


Fig. 4. **Alignment Percent Identities.** Results of Butyryl-CoA:acetate CoA Transferase Pathway Protein Sequence Alignments against the 457 HMP species. 73 total species were found to have at least 1 pathway protein present at a 60 percent identity threshold. A summary of those hits are shown in A-D. **A:** Species that contained all 7 pathway proteins at a 60 percent identity threshold. **B:** Species that had 5-6 proteins at a 60 percent identity threshold. **C-D:** Species that contained 1-4 proteins at a 60 percent identity threshold broken into two parts for readability.

The scored results were categorized into three distinct categories, Definite Capability, Probable Capability, and Requires Further Validation based on having scores of 7, 5-6, and 1-4 respectively. A breakdown of how many species fell into each category is shown in **Table 2**.

Breakdown of Scored Results

Category	Criteria	Count
Definite Capability	All	9
Probable Capability	5-6	21
Requires Further Validation	1-4	43

Table 2. Breakdown of Scored Results. Describes the distribution of categorical data for the scored results of protein pathway alignment for the 457 HMP species.

Though it was hypothesized that the results would be exclusively firmicutes, two other phyla of bacteria were also found among the results of these analyses. These results are

observed in **Figure 5A**. Firmicutes represented a majority of the results (74 percent). However, 23 percent of the results were Fusobacteria and 3 percent were Bacteroidota which could lend itself to further evaluation of these phyla and their metabolic activities. Similarly, it was suspected that the results would consist of bacteria of the clostridial class exclusively. The analysis results returned the representation of four additional classes of bacteria as summarized in **Figure 5B**. The classes found include Clostridia, Fusobacteriia, Negativicutes, Bacilli, and Bacteroidia.

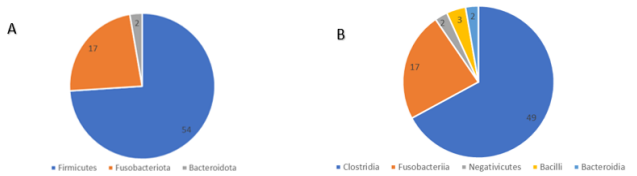


Fig. 5. **Breakdown of Phyla and Class from Potential Butyrate-Producing Bacterial Results** Three phyla and Five total classes were observed among the results of the scored alignment analyses.

Using the scored data, an ANOVA analysis was performed between the 5 classes of bacteria represented in **Figure 5B**. Its purpose was to identify if there was a significant difference in means among the scores for each class. The results of the ANOVA analysis were significant at a 95 percent confidence level with a p-vlaue of 0.0176. Thus, a Fischer's LSD was performed to identify the significant differences. As summarized in **Figure 6** and **Tables 3-4**, the significant differences indicate the Clostridia class is the highest scoring class among Bacteroidia, Bacilli, and Fusobacteriia. Though it does appear that this class is also higher scoring than Negativicutes, there were not enough results in that class to confer statistical significance.

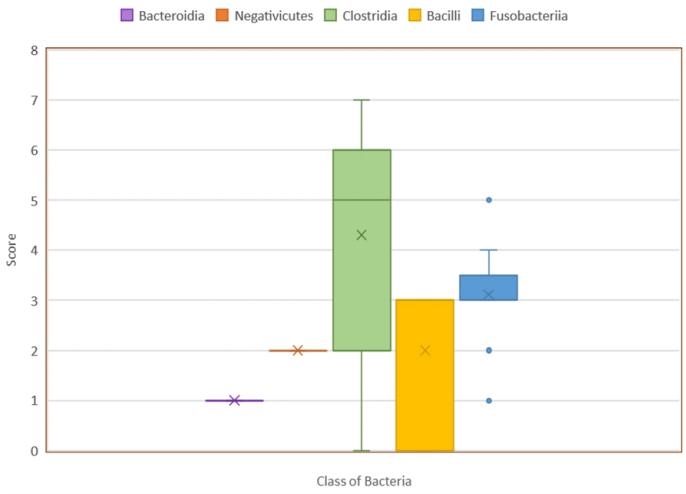


Fig. 6. **Statistical Distribution of Scores for 5 Classes of Potential Butyrate-Producing Bacteria.** Clostridia was among the highest scoring class of bacteria.

Mean Scores

Class	value	std	Min	Max	Q25	Q50	Q75
Bacilli	2	1.7320508	0	3	1.5	3	3
Bacteroidia	1	0	1	1	1	1	1
Clostridia	4.306122	2.3202378	0	7	2	5	6
Fusobacteriia	3.117647	0.8574929	1	5	3	3	3
Negativicutes	2	0	2	2	2	2	2

Table 3. Distibution of Score Means and Descriptive Statistics for 5 Classes of Potential Butyrate Producing Bacteria.

Fischer's LSD

Class		difference	pvalue
Bacilli	Clostridia	-2.306122	0.0585
Bacteroidia	Clostridia	-3.306122	0.0261
Clostridia	Fusobacteriia	1.188475	0.0399

Table 4. Significant Results of Fischer's LSD to Identify Differences in Scores of 5 Classes of Potential Butyrate Producing Bacteria.

Finally, a phylogenetic analysis of the resulting species revealed the following phylogenetic tree (**Figure 7**). The tree ultimately divides into two major groups: Group I and Group II. Group I contains the Clostridia, Bacilli, Negativicutes, and Bacteroidia, suggesting that these groups share more similarities with the class of bacteria that were proposed in this project's original hypothesis. On the other hand, Group II is composed entirely of the Fusobacteria class.

IV. Discussion

The results of this project found phyla and classes of potential butyrate producing bacteria outside of those expected from the literature. It was hypothesized that potential butyrate producers are within the realm of the firmicutes phyla and are majorly found in clostridial class clusters. However, the results of these analyses indicated that Fusobacteria and Bacterioidota phyla as well as the classes Bacilli, Negativicutes, Fusobacteriia, and Bacteroidia were also present.

Interestingly, though it was suspected that butyrate producing species include solely gram-positive anaerobes, the classes identified in this analysis also indicate the help of anaerobic gram-negative bacteria. Like Clostridium, the Bacillus class is an anerobic gram-positive bacteria [22]. However, Negativicutes, Bacteroides, and Fusobaceteria are all negative straining anaerobic bacteria [23], [24], [25]. Therefore, more exploration is needed to better understand the type of bacteria with these capabilities.

This project was limited due to the scope and time constraints of completion this Spring. Ways that this project could have been improved with an extended timeline include the inclusion of additional microbial genome datasets, such as the integrated Human Microbiome Project (iHMP) or Culturalable Reference Genome (CGR) for a wider breadth of species to analyze [26], [27]. Additionally, the pathway proteins used for the alignments could have provided a better standard for comparison if a consensus sequence of multiple difference versions of the pathway enzyme genes could have been used

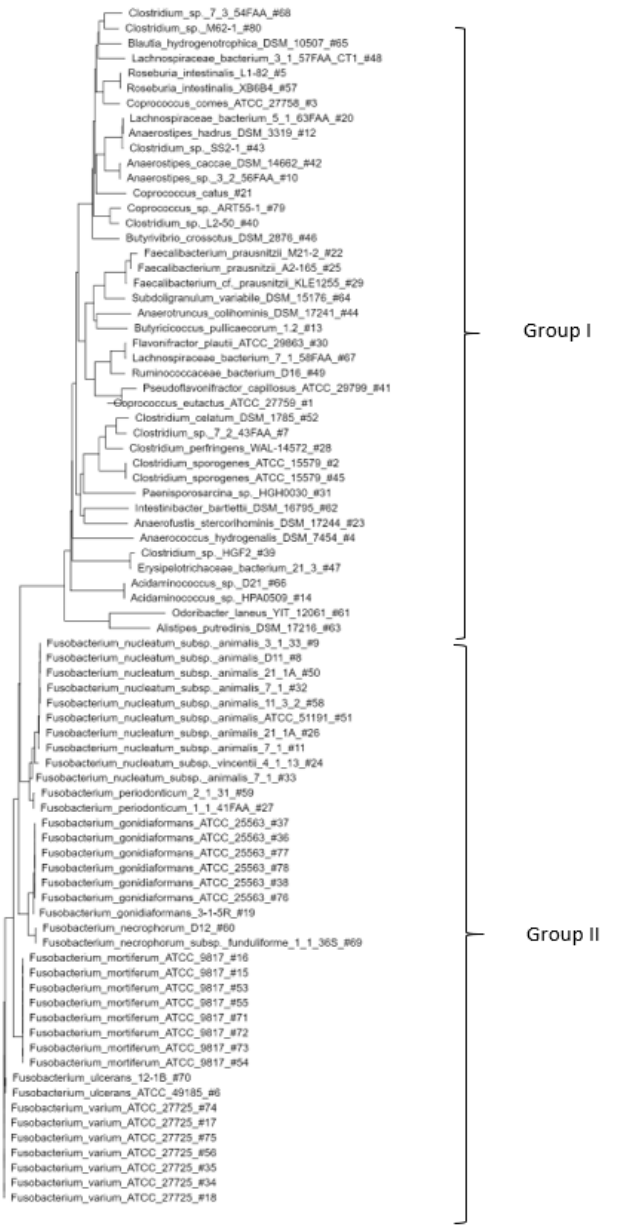


Fig. 7. **Phylogenetic Tree of Potential Butyrate-Producing Bacteria.** The identified species can be broken down into two high level groups.Group I contains the Clostridia, Bacilli, Negativicutes, and Bacteroidia classes whereas Group II contains the Fusobacteria class.

for each analysis. It is possible that the low results for the Butyrate Kinase pathway could have been a result of pulling the core pathway enzyme sequences from a more distant species that utilizes the Butyryl-CoA:acetate CoA Transferase Pathway.

To build on this project, there are multiple future steps that could add to the work done here. First, the butyrate-producing species could be confirmed in vivo. This could help to validate or exclude some of the lower scoring species from the analyses. Additionally, the resulting species could be further characterized by analyzing them for probiotic characteristics. This could provide guided directions for further analysis by targeting only those species that have the properties to fare well as a probiotic supplement. For instance, the class of Bacteroidia would likely not be worth pursuing as a probiotic as these types of bacteria are typically opportunistic pathogens [24]. Their potential anti-inflammatory benefits would be overshadowed by the risk of infection that they may pose if they become unbalanced in the microbiome or if they invade undesired tissues.

Some additional characteristics of a potential probiotic could include ensuring the bacteria is capable of adhesion and colonization in the colonic environment, having tolerance to acid and bile salts, and maintaining stability and vitality in the digestive tract [28]. If there are identifiable genes associated with these characteristics, preliminary analyses could be performed in silico before testing in in vivo models.

V. Conclusion

In conclusion, butyrate production via microbial species in the human gut is a natural alternative to treating and preventing inflammatory bowel disease, avoiding the various undesirable side effects of current treatments. This project identified 73 potential butyrate producing species. Though the majority of the high scoring results fell within the scope of the hypothesized phyla and class of bacteria (firmicutes; clostridia), a handful of others were identified that are phylogenetically related to clostridia and might be worth further exploration. With further analysis and validation, these bacteria could be candidates for probiotic supplements for the treatment and prevention of IBD.

References

[1] T. Sairenji, K. L. Collins, and D. V. Evans, "An update on inflammatory bowel disease." *Primary care*, vol. 44, no. 4, pp. 673–692, 2017.

[2] S. Alatab, S. G. Sepanlou, K. Ikuta, H. Vahedi, C. Bisignano, S. Safiri, A. Sadeghi, M. R. Nixon, A. Abdoli, H. Abolhassani *et al.*, "The global, regional, and national burden of inflammatory bowel disease in 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017," *The Lancet gastroenterology & hepatology*, vol. 5, no. 1, pp. 17–30, 2020.

[3] B. ZHANG, L.-c. LV, and L.-j. ZHUANG, "Litchi's regional production layout and development trend in china," *Guangdong Agricultural Sciences*, vol. 2011, p. 23, 2011.

[4] A. Geremia, P. Biancheri, P. Allan, G. R. Corazza, and A. Di Sabatino, "Innate and adaptive immunity in inflammatory bowel disease," *Autoimmunity reviews*, vol. 13, no. 1, pp. 3–10, 2014.

[5] S. Tedelind, F. Westberg, M. Kjerrulf, and A. Vidal, "Anti-inflammatory properties of the short-chain fatty acids acetate and propionate: a study with relevance to inflammatory bowel disease," *World journal of gastroenterology: WJG*, vol. 13, no. 20, p. 2826, 2007.

[6] M. Fakhoury, R. Negrulj, A. Mooranian, and H. Al-Salami, "Inflammatory bowel disease: clinical aspects and treatments," *Journal of inflammation research*, vol. 7, p. 113, 2014.

[7] S. E. Pryde, S. H. Duncan, G. L. Hold, C. S. Stewart, and H. J. Flint, "The microbiology of butyrate formation in the human colon," *FEMS microbiology letters*, vol. 217, no. 2, pp. 133–139, 2002.

[8] M. Vinolo, "a r, rodrigues hg, nachbar rt, curi r," *Regulation of inflammation by short chain fatty acids. Nutrients*, vol. 3, no. 10, pp. 858–76, 2011.

[9] R. B. Canani, M. Di Costanzo, L. Leone, M. Pedata, R. Meli, and A. Calignano, "Potential beneficial effects of butyrate in intestinal and extraintestinal diseases," *World journal of gastroenterology: WJG*, vol. 17, no. 12, p. 1519, 2011.

[10] P. Louis and H. J. Flint, "Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine," *FEMS microbiology letters*, vol. 294, no. 1, pp. 1–8, 2009.

[11] P. Louis, P. Young, G. Holtrop, and H. J. Flint, "Diversity of human colonic butyrate-producing bacteria revealed by analysis of the butyryl-coa: acetate coa-transferase gene," *Environmental microbiology*, vol. 12, no. 2, pp. 304–314, 2010.
[12] C. J. Walsh, C. M. Guinane, C. Hill, R. P. Ross, P. W. O'Toole, and P. D. Cotter, "In silico identification of bacteriocin gene clusters in the gastrointestinal tract, based on the human microbiome project's reference genome database," *BMC microbiology*, vol. 15, no. 1, pp. 1–11, 2015.
[13] S. H. Duncan, A. Barcenilla, C. S. Stewart, S. E. Pryde, and H. J. Flint, "Acetate utilization and butyryl coenzyme a (coa): acetate-coa transferase in butyrate-producing bacteria from the human large intestine," *Applied and environmental microbiology*, vol. 68, no. 10, pp. 5186–5190, 2002.
[14] T. L. Miller and M. J. Wolin, "Pathways of acetate, propionate, and butyrate formation by the human fecal microbial flora." *Applied and environmental microbiology*, vol. 62, no. 5, pp. 1589–1592, 1996.
[15] H. J. Flint, S. H. Duncan, K. P. Scott, and P. Louis, "Links between diet, gut microbiota composition and gut metabolism," *Proceedings of the Nutrition Society*, vol. 74, no. 1, pp. 13–22, 2015.
[16] J. Lloyd-Price, A. Mahurkar, G. Rahnavard, J. Crabtree, J. Orvis, A. B. Hall, A. Brady, H. H. Creasy, C. McCracken, M. G. Giglio *et al.*, "Strains, functions and dynamics in the expanded human microbiome project," *Nature*, vol. 550, no. 7674, pp. 61–66, 2017.
[17] K. Nelson, G. Weinstock, S. Highlander, K. Worley, H. Creasy, J. Wortman, D. Rusch, M. Mitreva, E. Sodergren, A. Chinwalla *et al.*, "lenvercoe e, earl am, farmer cn, liolios k, surette mg, xu q, pohl c, wilczekboney k, zhu d," *A catalog of reference genomes from the human microbiome. Science*, vol. 328, pp. 994–999, 2010.
[18] C. Quast, E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, and F. O. Glöckner, "The silva ribosomal rna gene database project: improved data processing and web-based tools," *Nucleic acids research*, vol. 41, no. D1, pp. D590–D596, 2012.
[19] P. J. Cock, T. Antao, J. T. Chang, B. A. Chapman, C. J. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, B. Wilczynski *et al.*, "Biopython: freely available python tools for computational molecular biology and bioinformatics," *Bioinformatics*, vol. 25, no. 11, pp. 1422–1423, 2009.
[20] F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding *et al.*, "Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega," *Molecular systems biology*, vol. 7, no. 1, p. 539, 2011.
[21] R. M. Moore, A. O. Harrison, S. M. McAllister, S. W. Polson, and K. E. Wommack, "Iroki: automatic customization and visualization of phylogenetic trees," *PeerJ*, vol. 8, p. e8584, 2020.
[22] S. Baron, *Leptospira–Medical Microbiology*. University of Texas Medical Branch at Galveston, 1996.
[23] T. Vesth, A. Ozen, S. C. Andersen, R. S. Kaas, O. Lukjancenko, J. Bohlin, I. Nookaew, T. M. Wassenaar, and D. W. Ussery, "Veillonella, firmicutes: microbes disguised as gram negatives," *Standards in genomic sciences*, vol. 9, no. 2, pp. 431–448, 2013.
[24] H. M. Wexler, "Bacteroides: the good, the bad, and the nitty-gritty," *Clinical microbiology reviews*, vol. 20, no. 4, pp. 593–621, 2007.
[25] R. Garcia-Carretero, M. Lopez-Lomba, B. Carrasco-Fernandez, and M. T. Duran-Valle, "Clinical features and outcomes of fusobacterium species infections in a ten-year follow-up," *The Journal of Critical Care Medicine*, vol. 3, no. 4, p. 141, 2017.
[26] H. Integrative, "Research network consortium the integrative human microbiome project: dynamic analysis of microbiome-host omics profiles during periods of human health and disease," *Cell Host Microbe*, vol. 16, no. 3, pp. 276–289, 2014.
[27] Y. Zou, W. Xue, G. Luo, Z. Deng, P. Qin, R. Guo, H. Sun, Y. Xia, S. Liang, Y. Dai *et al.*, "1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses," *Nature biotechnology*, vol. 37, no. 2, pp. 179–185, 2019.
[28] M. De Vrese, Schrezenmeir, and J, "Probiotics, prebiotics, and synbiotics," *Food biotechnology*, pp. 1–66, 2008.