BIOINFORMATICS STUDY OF THE HUMAN GUT MICROBIOME AND RELATIONSHIP WITH DISEASE

Jayme McKinney
Natural Science Department
Middlesex College
Professor Phalguni Ghosh

Table of Contents

Title

Bioinformatics Study of the Human Gut Microbiome and Relationship with Disease

Author

Jayme McKinney

Abstract

The gut microbiome has many roles within the human body and that are known and speculated. One role is regulating homeostasis within the gut, and another is regulating the inflammatory response within the host human. The gut microbiome is the metabolic pathway for many different substrates and metabolites. Using data from gutMGene that focused on the gut microbiome-chemical interactions that occur in the human body a network was mapped out with Cytoscape to display the most prominent substrates, bacteria, and metabolites within the human gut. Then using research acquired from different articles and gutMDisorder I then accumulated data that showed the relationship between bacteria dysbiosis and disease, I found the diseases that had the most relationships with the gut microbiota from mapping the interactions in Cytoscape. This shows the bacteria that are affected by diseases and whether they have an increasing interaction or a decreasing interaction. This shows us that there is a relationship between the gut microbiome and disease and allows us to see the bacteria that increased and decreased and seeing its connection to chemicals I can see whether the gut microbiome affects inflammation within the gut.

Keywords

Gut microbiome, microbiota, metabolites, substrates, bacteria, dysbiosis, Cytoscape

Introduction

This paper focuses on and examines gut microbiota-chemical interactions as well as gut microbiota-disease-dysbiosis interactions between 19 different diseases. I used a bioinformatics software, Cytoscape, to learn about the effects of the gut microbiome when looking at the bacteria and chemicals that are metabolized and produced by specific gut microbiota as well as to see the relationship between microbiota and specific diseases.

I had several objectives for this research project. My main objective was to do a bioinformatics study with Cytoscape, to map the gut microbiota and their interactions with substrates, metabolites, and diseases. The 19 diseases I focused on were specifically from the following articles

- Kho, Zhi Y., and Sunil K. Lal. "The Human Gut Microbiome – a Potential Controller of Wellness and Disease." *Frontiers*, Frontiers, 1 Jan. 1AD, https://www.frontiersin.org/articles/10.3389/fmicb.2018.01835/full.
- Zhang, Yu-Jie et al. "Impacts of gut bacteria on human health and diseases." *International journal of molecular sciences* vol. 16,4 7493-519. 2 Apr. 2015, doi:10.3390/ijms16047493

I then identified the two most prominent substrates, metabolites from one network and the two diseases studied which had the most dysbiosis interactions with the gut microbiome from a second network. There were two different software that were used for my research; Cytoscape was used to map relationships between many different sources and targets. In order to gather and compile the date to map it in Cytoscape, Microsoft Excel was used. Many databases were used, including The Human Metabolome Database (HMDB), PubChem, Drug Bank Online, gutMGene, and gutMDisorder. HMDB, PubChem, and Drug Bank Online were used to gather information about the most common substrates and metabolites. gutMGene and gutMDisorder were used to gather information about the bacteria-chemical interactions and microbiota-disease interactions respectively.

Methodology

Bacteria-Chemical Network

First, I started by gathering data from gutMGene; where I downloaded all of the human bacteria-chemical interactions. Then I condensed the spreadsheet, removing spaces and gut bacteria that did not have an interaction with a metabolite or substrate as well as other information that was not key to the study. Once I had all the bacteria and chemicals that I wanted to study I gathered taxonomical information on each microbiota, specifically phylum and class. I also gathered the chemical classification for each chemical and the direct parent for each chemical from HMDB.

Gut Microbiota	Phylum	Bacteria Class	Chemical	Chemical Type	Chem Class	Direct Parent
Christensenella minuta YIT 12065T	firmicutes	clostridia	Acetate	metabolite	carboxylic acids and derivatives	acetates
Christensenella minuta YIT 12065T	firmicutes	clostridia	Acetate	metabolite	carboxylic acids and derivatives	acetates
Christensenella minuta YIT 12065T	firmicutes	clostridia	Acetate	metabolite	carboxylic acids and derivatives	acetates
Christensenella minuta YIT 12065T	firmicutes	clostridia	Acetate	metabolite	carboxylic acids and derivatives	acetates
Christensenella minuta YIT 12065T	firmicutes	clostridia	Acetate	metabolite	carboxylic acids and derivatives	acetates
Christensenella minuta YIT 12065T	firmicutes	clostridia	Acetate	metabolite	carboxylic acids and derivatives	acetates
Christensenella minuta YIT 12065T	firmicutes	clostridia	Butyrate	metabolite	fatty acyls	straight chain fatty acids
Christensenella minuta YIT 12065T	firmicutes	clostridia	Butyrate	metabolite	fatty acyls	straight chain fatty acids
Christensenella minuta YIT 12065T	firmicutes	clostridia	Butyrate	metabolite	fatty acyls	straight chain fatty acids
Christensenella minuta YIT 12065T	firmicutes	clostridia	Butyrate	metabolite	fatty acyls	straight chain fatty acids
Christensenella minuta YIT 12065T	firmicutes	clostridia	Butyrate	metabolite	fatty acyls	straight chain fatty acids
Christensenella minuta YIT 12065T	firmicutes	clostridia	Butyrate	metabolite	fatty acyls	straight chain fatty acids
Christensenella minuta YIT 12065T	firmicutes	clostridia	D-Glucose	substrate	organooxygen compound	hexoses

Once the spreadsheet was created, I was then able to load my data into Cytoscape and map the relationships between the bacteria and chemicals. For this bacteria-chemical network, the source node is bacteria with attributes being phylum and class. The target node is either a metabolite or substrate with the assigned attributes of chemical class and direct parent. The sizes of the note are varied based on the degree of interaction with other nodes.

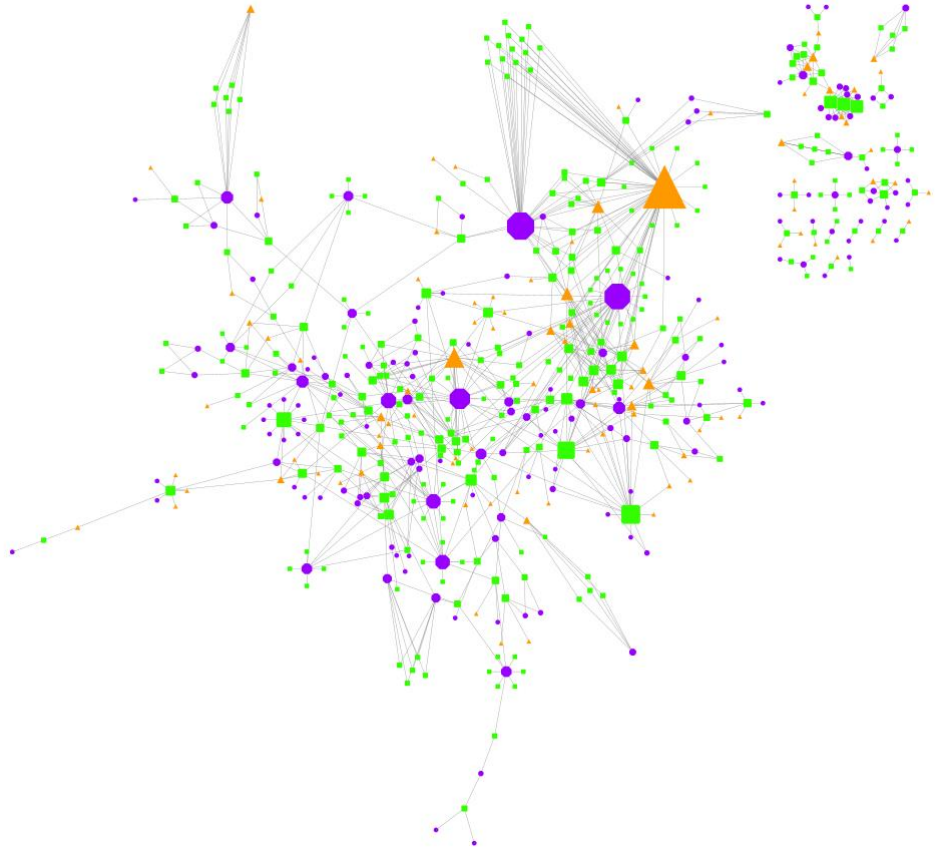


Figure 1 Full Bacteria-Chemical Network

Once the Cytoscape was created I then identified the two most prominent substrates and metabolites and did research on them to find out their roles in the human body.

Microbiota-Disease Network

For this second network, I collected microbiota dysbiosis-disease interactions for 19 different diseases from the two articles listed earlier and from gutMDisorder. After gathering the information between microbiota and their interactions with disease, I added the attributes phylum and class to each microbiota listed.

Specific Disease	Disease category	Microbe	Dysbiotic feature (increase or decrease)
atopic disease	immune system disease	clostridium difficile	increase
autism	neuropsychiatric	clostridium sp	increase
autism	neuropsychiatric	bacteroidetes	increase
autism	neuropsychiatric	lactobacillus desulfovibrio	increase
autism	neuropsychiatric	bifidobacteria	decrease
autism	neuropsychiatric	faecalibacterium	decrease
autism	neuropsychiatric	prevotella	decrease
autism	neuropsychiatric	coprococcus	decrease
autism	neuropsychiatric	haemophilus	decrease
autism	neuropsychiatric	prevoellaceae	decrease

After creating the full version of the table above, I created a Cytoscape network with the source node being specific disease and target node being microbiota with the interaction being dysbiotic feature – which can be increasing or decreasing. Nodes are different sizes to show the varying degrees of interaction between other nodes. The diseases are purple octagons while the squares are the microbiota which are color coded by phylum with the color key on the left.

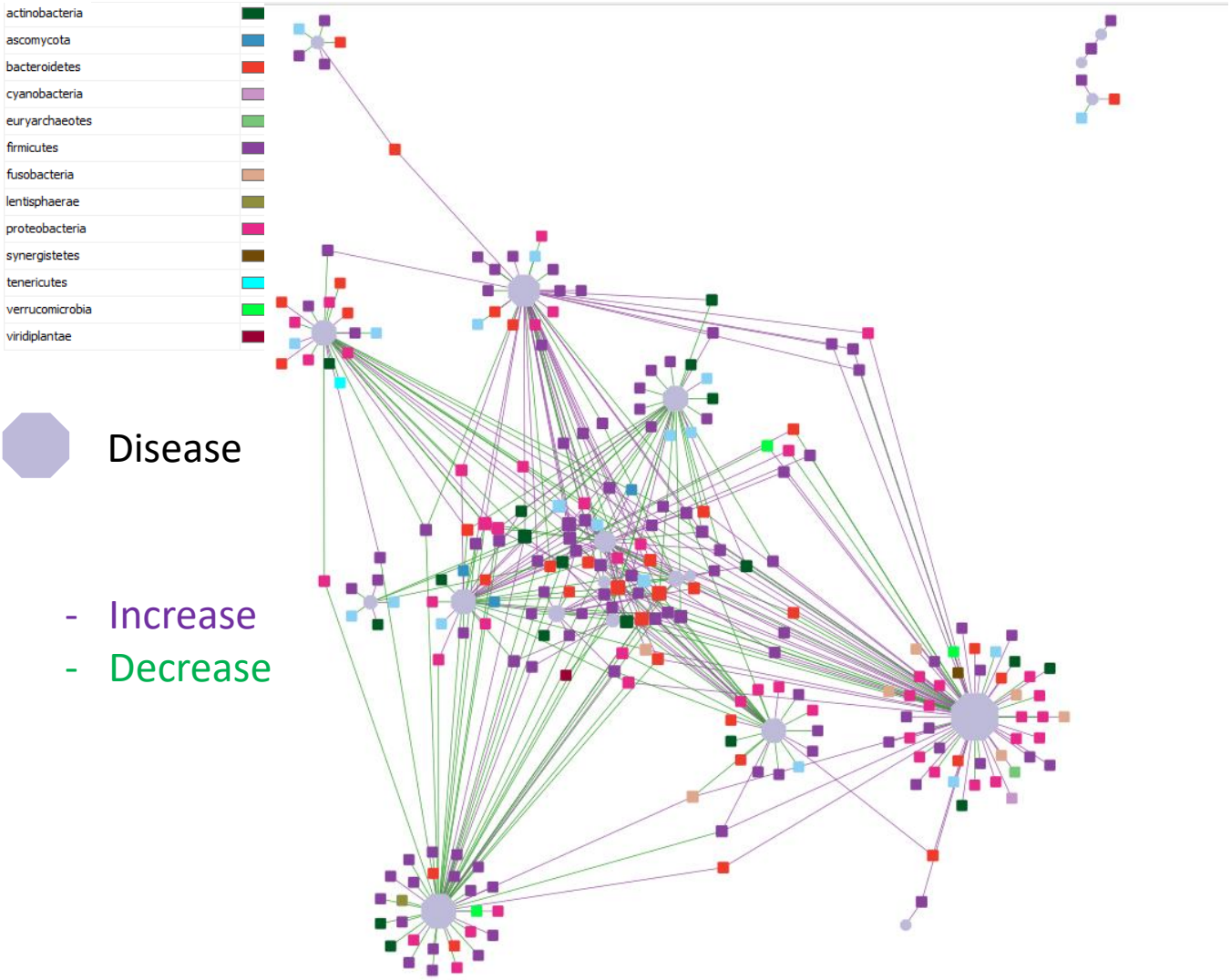


Figure 2 Microbiota-Disease Network

After the network was created, I analyzed the network by applying different formats to my network and seeing which bacteria and diseases had the highest degrees of interaction.

Discussion

Bacteria-Chemical Network

When analyzing the gut microbiota inside the human body catalogued on gutMGene with the program Cytoscape, the three most prevalent metabolites were butyrate, and equol. The three most common substrates were daidzein, D-glucose and dihydrodaidzein. Within this summary there will be use of Cytoscape, a networking program. The network data was taken from the human URL download on gutMGene for the bacteria, metabolites and substrates and The Human Metabolome Database for the chemical class and direct parent.

Butyrate is from the chemical class fatty acyls and its direct parent is straight chain fatty acids. Butyrate is a short-chain fatty acid anion with the molecular formula $C_4H_7O_2^-$. Butyrate displays anti-inflammatory and regenerative properties and provides relief to symptoms of colonic diseases when patients were given an oral supplement. (https://clinicaltrials.gov/ct2/show/NCT04879914). Butyrate is derived from the metabolism of indigestible carbohydrates by the gut microbiota. Butyrate contributes to homeostasis within the human gut and studies indicate that it might control inflammatory responses and host physiology in different types of tissues (https://pubmed.ncbi.nlm.nih.gov/31366236/). There are 29 different bacteria according to gutMGene that interact with and metabolize butyrate. Butyrate works by acting as an inhibitor for histone deacetylase (HDAC); it also works by signaling via various G protein-coupled receptors (GPCRs). There has been more and more evidence of the effect of butyrate on the gut-brain axis. (https://academic.oup.com/advances/article/9/1/21/4849000). GPCRs are the most diverse group of transmembrane proteins. Butyrate directly regulates GPR41-mediated sympathetic nervous system activity to regulate the body energy expenditure and preserve homeostasis of the human metabolism. GPR109A is activated by metabolite butyrate; the signaling of GPR109A activates the inflammasome pathway – inflammasomes are core signaling platforms that find pathogenic microorganisms and sterile stressors (stress induced sterile inflammation), they also activate the highly pro-inflammatory cytokines interleukin-1β and interleukin-18 (IL-18) (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3807999/) – in colonic macrophages and dendric cells, causing a difference between regulatory T cells and interleukin-10-producing T cells. The secretion of interleukin-18 is increased in intestinal epithelial cells via butyrate-stimulated signaling of GPR109A. Anti-inflammatory characteristics that butyrate is known for are accomplished by inhibiting the production of inflammatory-enzymes and cytokines (https://academic.oup.com/advances/article/9/1/21/4849000). Diseases and disorders associated with, as a marker or mechanism, butyrate include disease progression, Epstein-Barr Virus Infections, genomic instability, nasopharyngeal carcinoma, pain, pericardial effusion, and prenatal injuries (https://pubchem.ncbi.nlm.nih.gov/compound/104775#section=Associated-Disorders-and-

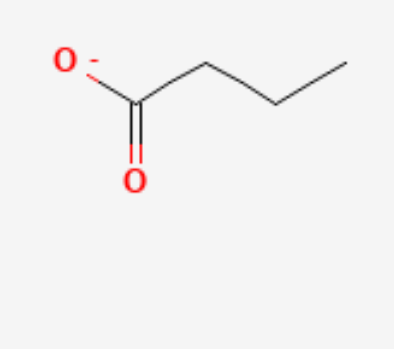


Figure 3Chemical Structure of Butyrate

Diseases).

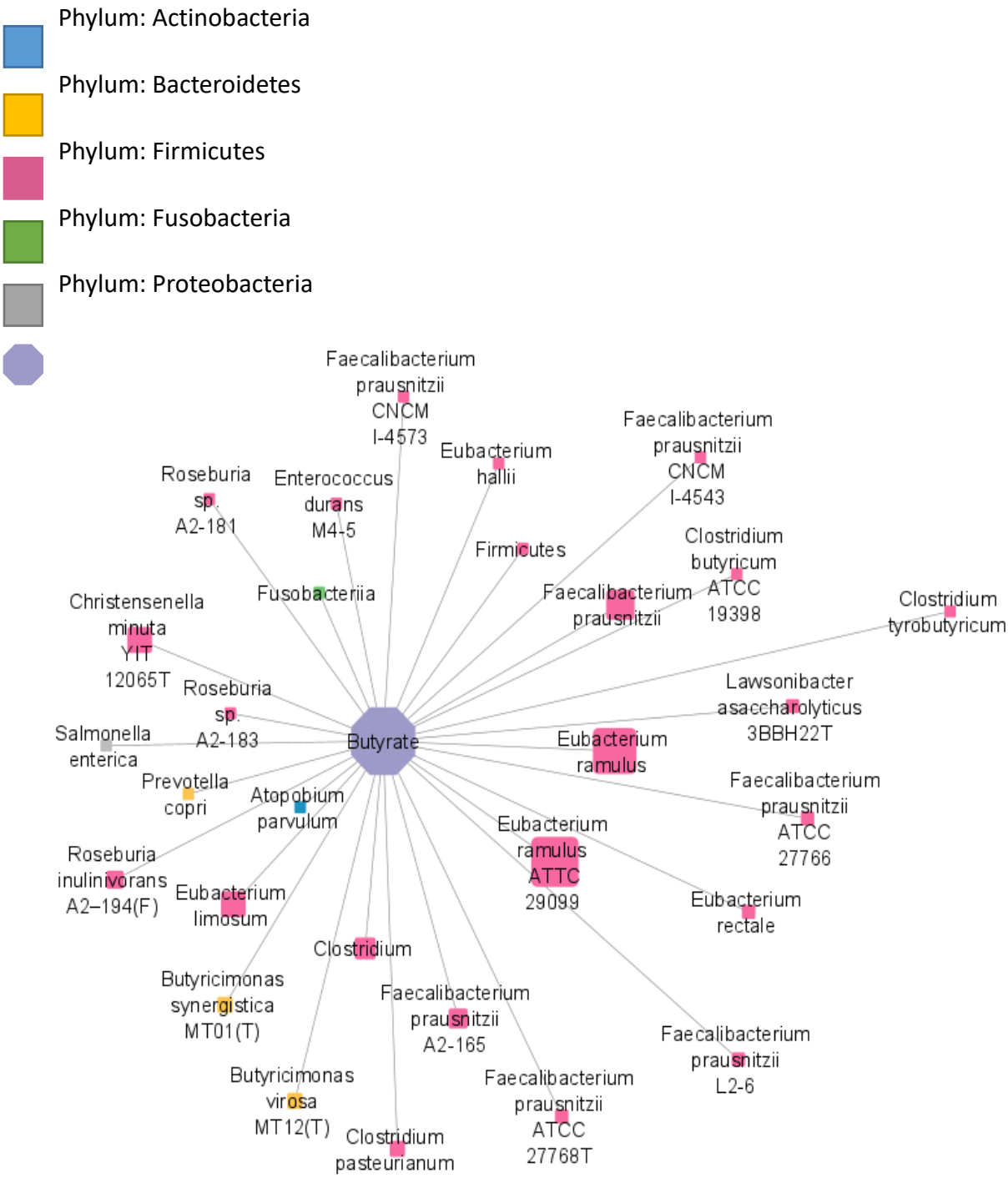


Figure 4 Cytoscape of Bacteria-Metabolite Interactions with Butyrate

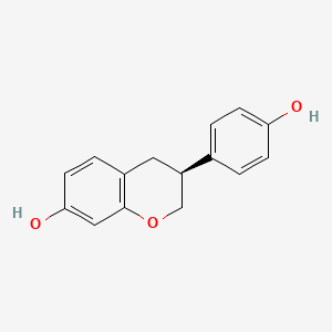


Figure 5 Chemical Structure of Equol

Equol belongs to the chemical class is isoflavonoids and its direct parent is isoflavanols (https://hmdb.ca/metabolites/HMDB0002209). Equol is a non-steroidal estrogen that comes from the metabolization of soybean products by the 31 bacteria (shown in the Cytoscape of equol) in the intestines. The functional uses of equol are as an antioxidant, hair conditioning and skin conditioning (https://pubchem.ncbi.nlm.nih.gov/compound/91469#section=Pathways). Equol is produced from daidzein and is the isoflavone-derived metabolite with the most estrogenic and antioxidant activity (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6770660/). Equol is produced through the formation of dihydrodaidzein via the gut microbiota. The mechanism – of which an unspecified gut bacterium is responsible for – the colonic biotransformation of daidzein to equol is unknown. One thing about metabolizing equol is that only ~30-40% of adults can do the transformation of daidzein to equol(https://www.cambridge.org/core/journals/proceedings-of-the-nutrition-society/article/is-equol-production-beneficial-to-health/105B5CE52EA2ECC354143A20CFA8F080). Equol might have greater biological activity when compared to equol's parent compound (daidzein). Equol's positive effects are debatable with limited studies showing improved symptomatic relief from hot-flashes, lower risk of breast and prostate cancer and improved bone health as well as equol having negative or no effect on the body. The diseases associated with equol therapy are breast and prostatic neoplasms (cancer) (https://pubchem.ncbi.nlm.nih.gov/compound/91469#section=Associated-Disorders-and-Diseases).

Bacteroidetes phylum

Actinobacteria phylum

Firmicutes phylum

Equol

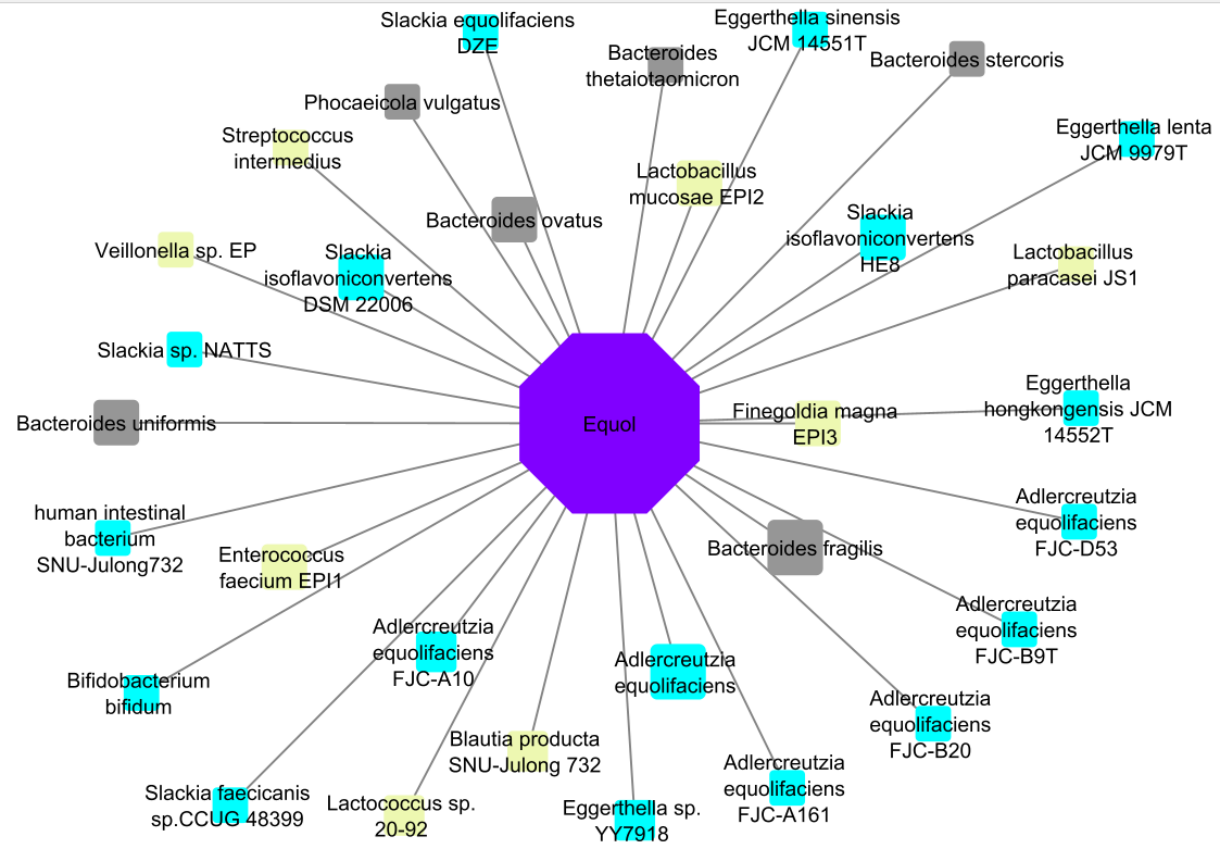
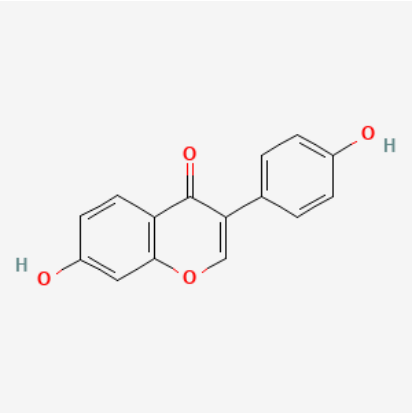


Figure 6 Cytoscape of Equol-Gut Bacteria Interactions

Daidzein belongs to the chemical class of isoflavonoids, and its direct parent is isoflavones. There are 54 different bacteria interactions with daidzein according to the Cytoscape. Daidzein's chemical formula is $C_{15}H_{10}O_4$. Daidzein in a human metabolite of formononetin. The main product – equol – of the daidzein metabolism and is produced by specific gut microbiota. Daidzein acts as a therapeutic treatment for a few diseases – adenocarcinoma, amnesia, cardiovascular diseases and type 2 diabetes mellitus, edema, endometrial neoplasms, hyperplasia, hypertension, leukemic infiltration, memory disorders and ventricular fibrillation.



Daidzein is also a marker for gynecomastia, erectile dysfunction and congenital abnormalities (pubchem daidzein). Daidzein can be found in soy products.

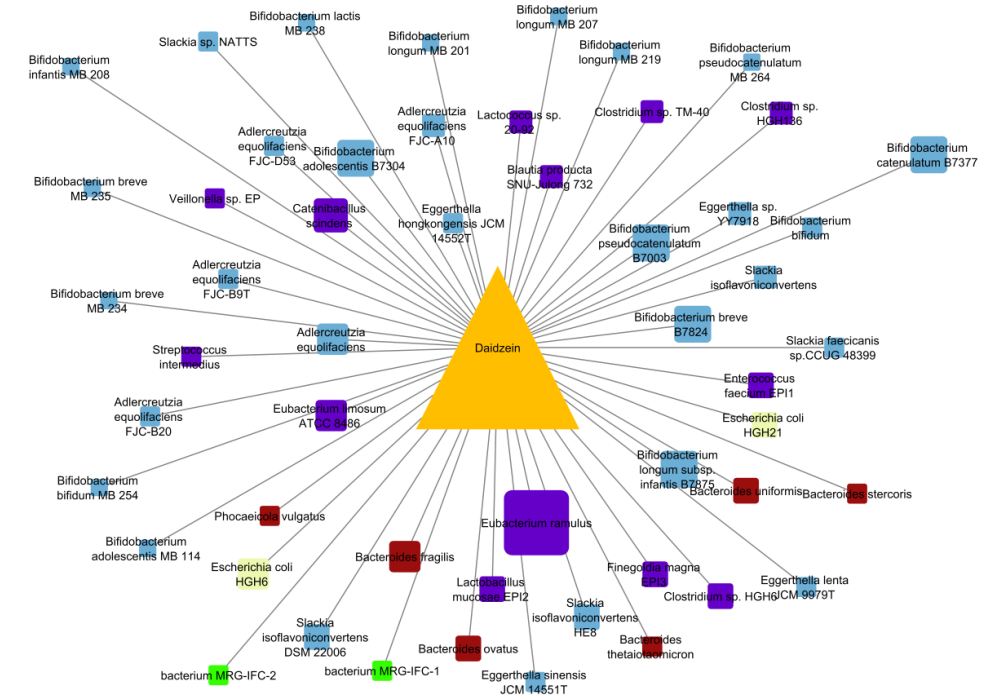


Figure 8 Cytoscape of Daidzein-Bacteria Interactions

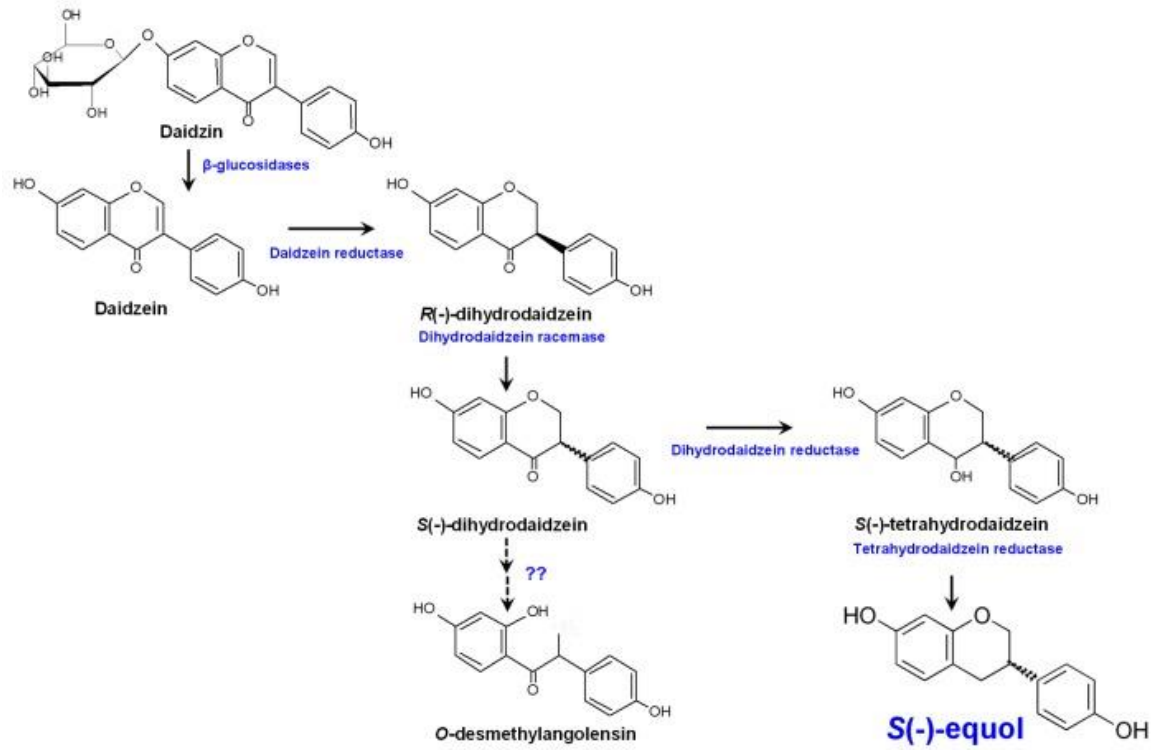
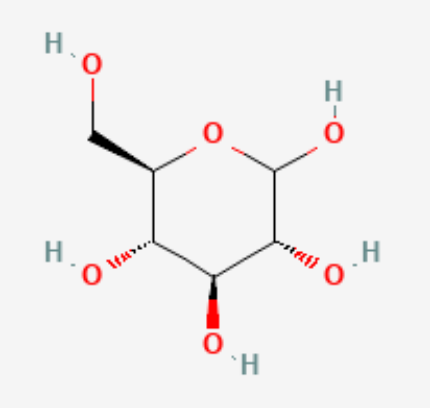


Figure 9 Pathway from Daidzein to Equol



D-Glucose is an organooxygen compound with hexoses as its direct parent. There are 22 different gut-bacteria interaction with D-glucose according to the Cytoscape. The chemical formula for D-glucose is $C_6H_{12}O_6$. D-glucose is the most regularly occurring isomer of glucose in nature. D-glucose is often referred to as Dextrose and is uses as a medicine for treatment of metabolic disorders and nutrient deprivation. D-glucose's role in the human body is to act as a signaling molecule to control glucose and energy homeostasis. Also, D-Glucose can regulate gene transcription, enzyme activity, hormone secretion and the activity of glucoregulatory neurons (https://go.drugbank.com/drugs/DB01914). D-glucose is associated with the following disorders and disease by being a marker/mechanism: chemical and drug induced liver injury, fatty liver, hepatomegaly, insulin resistance and liver cirrhosis (https://pubchem.ncbi.nlm.nih.gov/compound/5793#section=Associated-Disorders-and-Diseases).

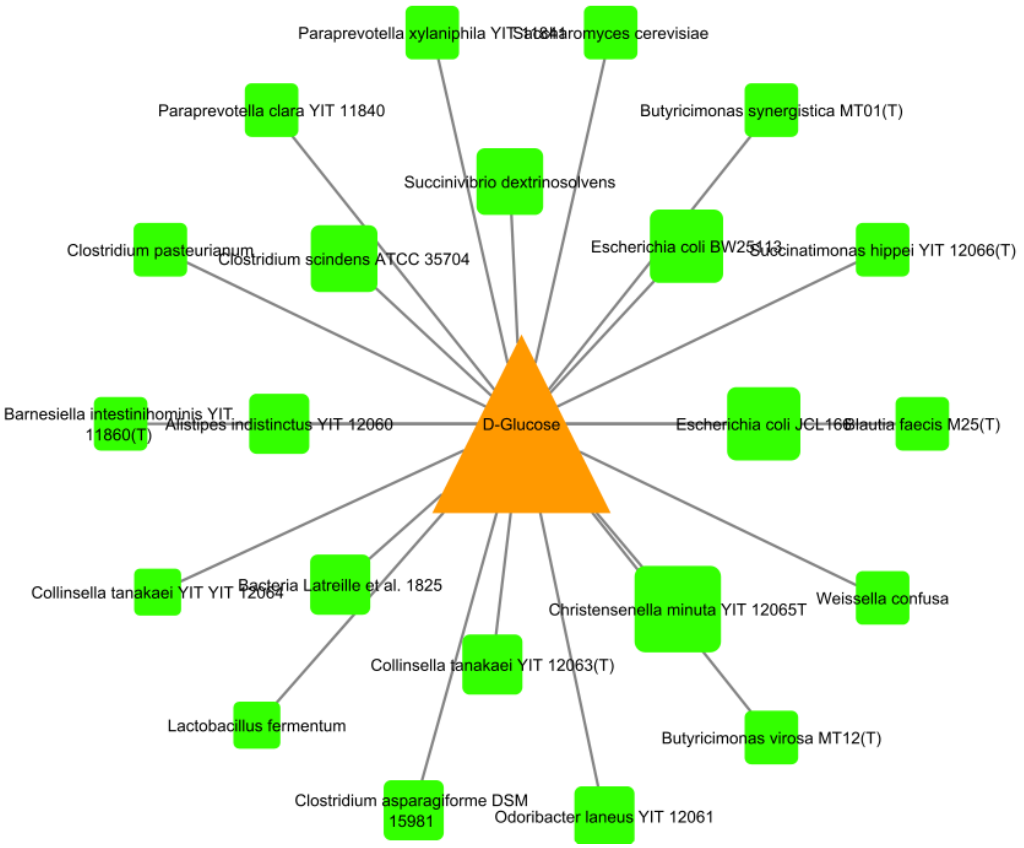
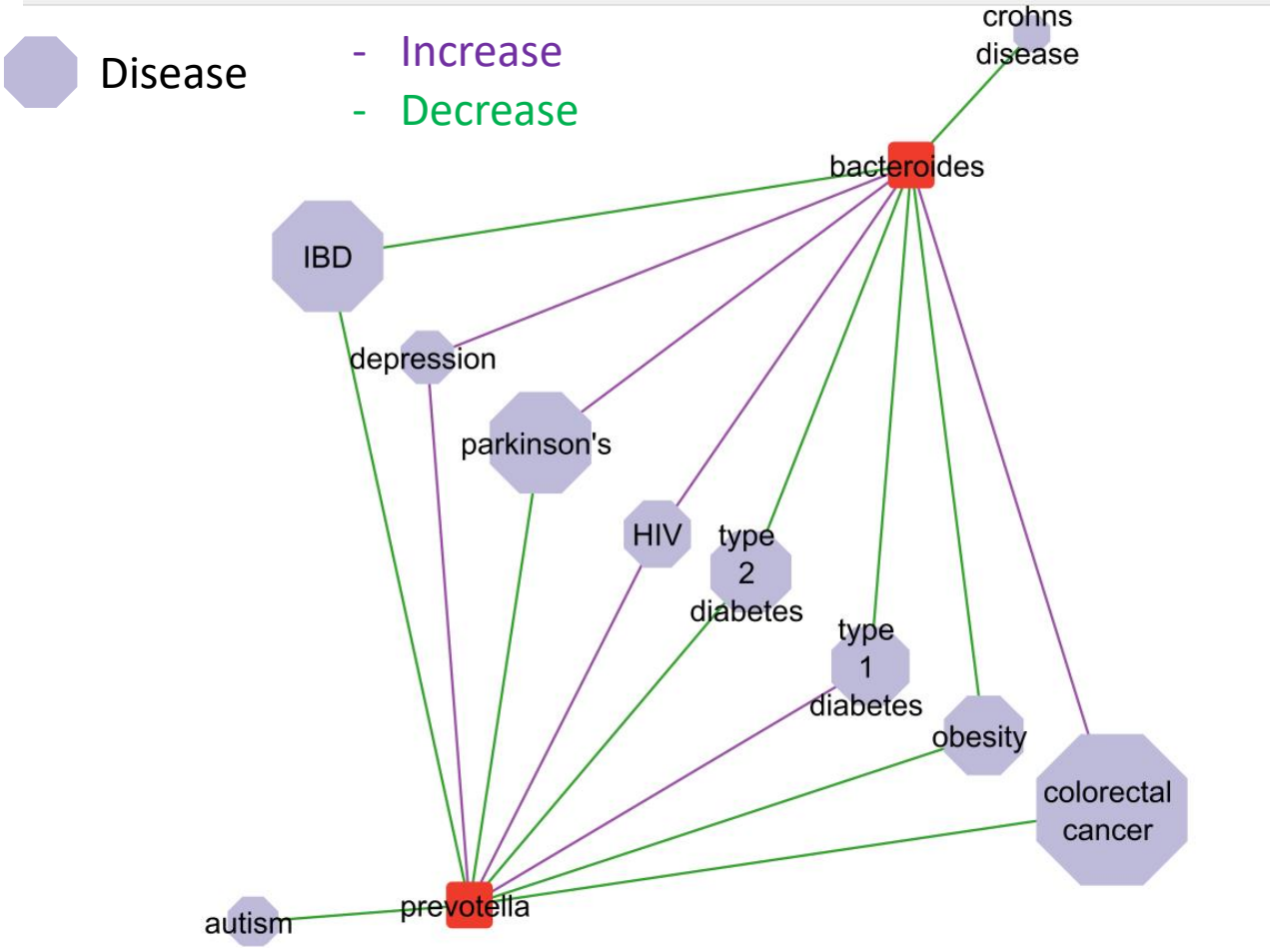


Figure 10 Cytoscape of D-Glucose-Bacteria Interactions

Microbiota-Disease Network

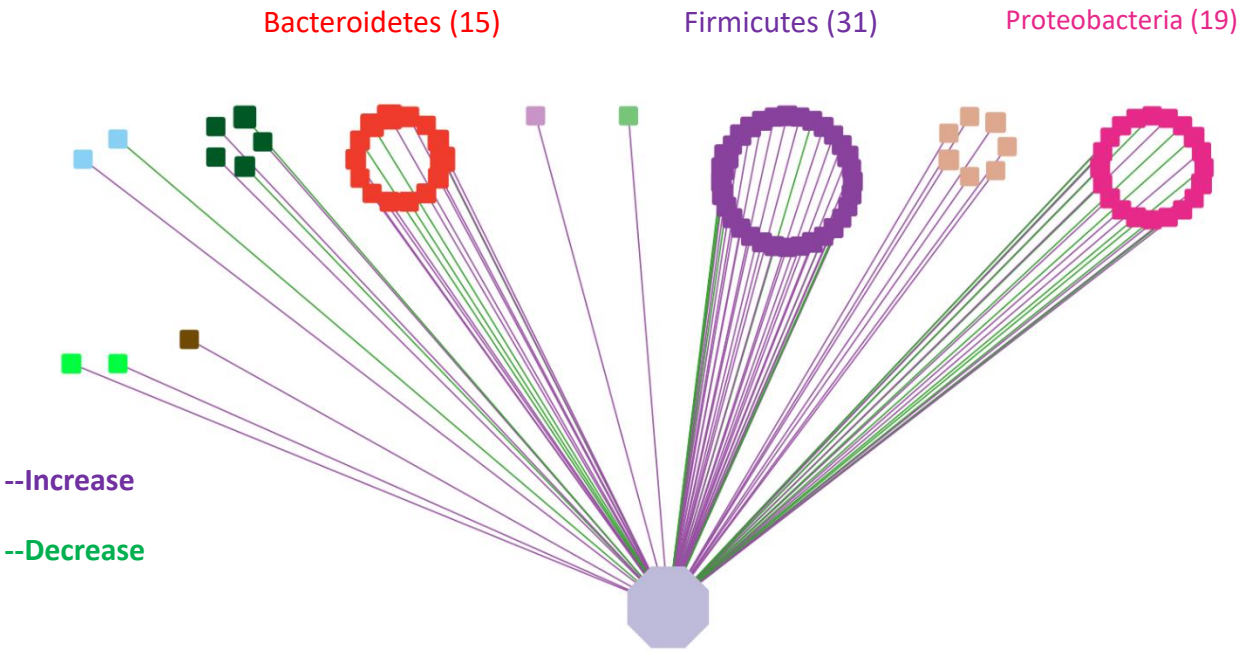
Of the 19 different diseases studied where the gut microbiome was affected, two stood out as having the most dysbiosis interactions with the most amount of microbiota. These diseases were Inflammatory Bowel Disease (IBD) and colorectal cancer. There were also two bacteria with the

highest degree of interaction with different diseases, Prevotella and Bacteroides. Prevotella and Bacteroides are both from the phylum Bacteroidetes, which is why in the network they are both red. Increase in dysbiotic feature of the bacteria is shown with purple lines while decreasing dysbiotic features are shown with green lines. Of the 9 diseases that each bacteria have an interaction with, 8 of them are the same with Crohn's Disease and autism being individual to Bacteroides and Prevotella respectively.

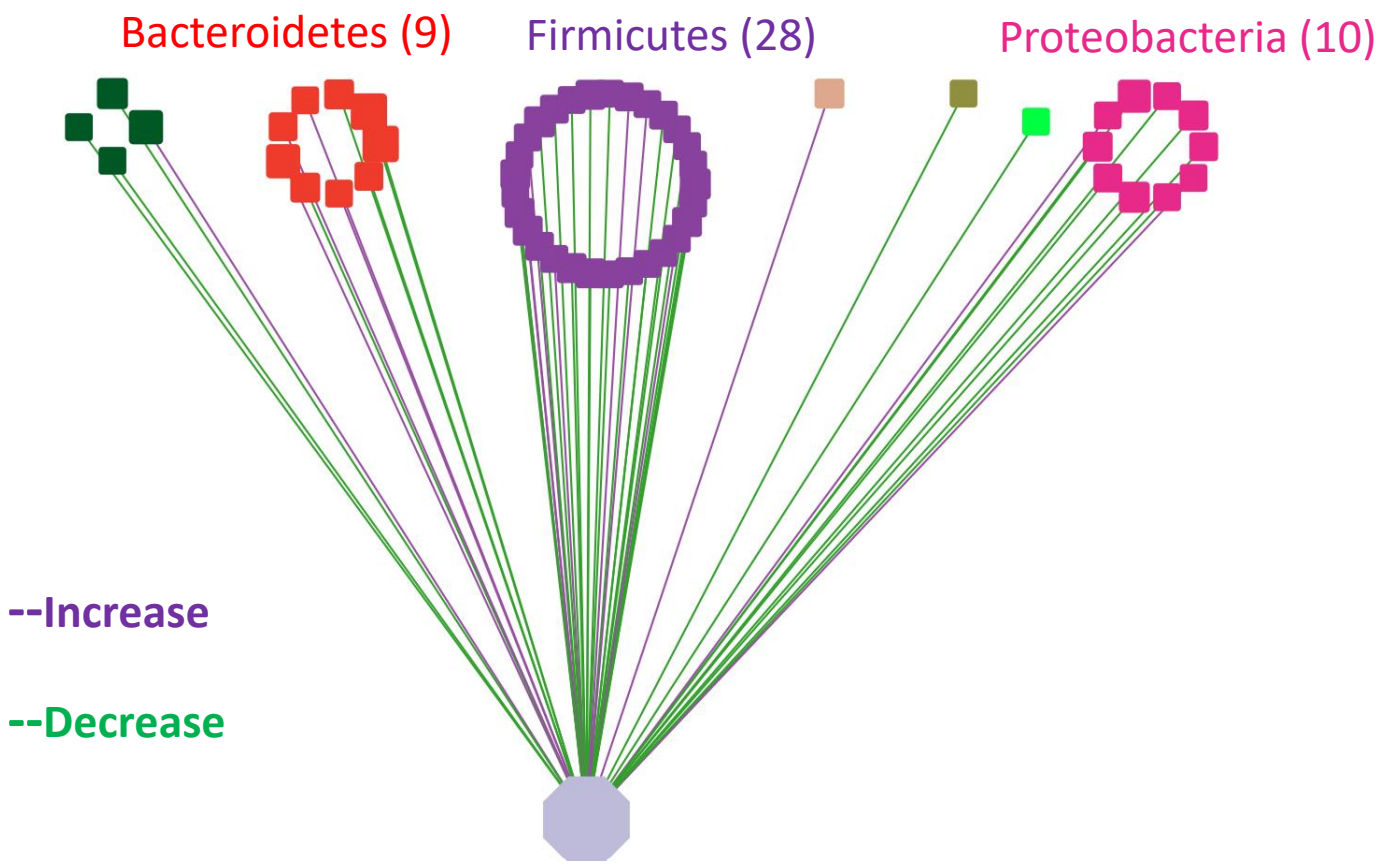


Colorectal cancer is a one of the leading causes of cancer deaths in the world and is a disease associated with genetic and environmental factors. Studies show that bacterial blooms can contribute to colorectal cancer. Examples of enrichment contributing to colorectal cancer is when enterotoxigenic strains of Bacteroides fragilis contributes to colorectal cancer through boosted production of *B. fragilis* enterotoxin. Non-colitogenic Fusobacterium nucleatum is also highly present in colorectal cancer patients, promoting inflammation and proliferation of the cancer, and suggests that overrepresentation of the gut microbiota might be a driver for the disease. This suggestion corresponds to the Cytoscape network where there is an increase in many of the microbiota within the gut. Microbial metabolites such as butyrate have been reported to inhibit the proliferation and survival of tumor cells. I found 85 different interactions between gut microbiota and colorectal cancer, and here they are displayed in Cytoscape sorted by phylum. The three most populous

bacteria phylum are labeled. There are 10 phyla known to be affected by colorectal cancer. There is a decrease in the butyrate producing bacteria Faecalibacterium, Fusobacteriia, and butyrate producing Firmicutes. The decrease in butyrate producing bacteria may be a sign that there is less butyrate being produced and therefore there is more inflammation within the gut.



IBD is a very common disease only getting more common within the human population. IBD is a gut-microbiome associated disease. It is said that the unnecessary immune response that happens against microbiota in people who are genetically predisposed to IBD is the cause of severe inflammation. Gut microbiota dysbiosis is suggested to contribute to IBD pathogenesis. Now, while most therapies deal with the inflammatory response, these treatments are not too safe. Meanwhile, probiotics have shown to help significantly with the symptoms of IBD, but, when taken off the probiotics, the disease can come back with worse symptoms and a person might require long-term treatment. IBD is linked to 55 different microbiota-dysbiosis interactions. IBD is represented by a purple octagon with the microbiota being squares that are color coded by their phylum. Green lines mean a decrease in microbiota while purple lines mean increase in microbiota. There are 7 phylum of gut microbiota affected by IBD, with the most prominent being Bacteroidetes, Firmicutes and Proteobacteria. Most of the microbiota within the gut are decreasing suggesting that the immune system is inhibiting the survival of the microbiome when a person has IBD. There is a decrease in the butyrate producing firmicutes, Faecalibaterium prausnitzii and Roseburia sp., all of which have interactions with butyrate and can be used to infer that the decrease in butyrate producing microbiota leads to a pro-inflammatory response that further inhibits the survival of bacteria within the gastrointestinal tract.



Conclusion

In colorectal cancer and in IBD cases, there is evidence that butyrate producing bacteria decreases within the gut. Butyrate was found to have potential and known anti-inflammatory and anti-cancer properties. In colorectal cancer there are more microbiota-dysbiosis interactions where the microbiota is increasing, supporting that bacterial blooms are a cause for colorectal cancer and a cause for the symptoms of colorectal cancer. IBD has many decreasing microbiota-dysbiosis interactions, and the data shows that the increase in inflammation and altered immune responses lowers the number of bacteria in the gut. After studying the most common microbiota effected by disease, I hypothesize that probiotics composed of butyrate-producing bacteria can be used to offset the negative dysbiosis occurring in cases of colorectal cancer and IBD and to aid in the treatment of symptoms of the diseases.

References

- https://www.gutmicrobiotaforhealth.com/about-gut-microbiota-info/#:~:text=It's%20home%20to%20more%20than,is%20unique%20to%20each%20individual.
- Rowland, Ian, et al. “Gut Microbiota Functions: Metabolism of Nutrients and Other Food Components.” *European Journal of Nutrition*, vol. 57, no. 1, 2017, pp. 1–24., https://doi.org/10.1007/s00394-017-1445-8.
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4290017/
- https://clinicaltrials.gov/ct2/show/NCT04879914
- McNabney, Sean & Henagan, Tara. (2017). Short Chain Fatty Acids in the Colon and Peripheral Tissues: A Focus on Butyrate, Colon Cancer, Obesity and Insulin Resistance. Nutrients. 9. 1348. 10.3390/nu9121348.
- Soy, isoflavones, and prostate cancer - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Proposed-metabolic-pathway-leading-to-formation-of-equol-from-daidzein-in-a-urine-sample_fig2_23452889 [accessed 21 Nov, 2021]
- National Center for Biotechnology Information. "PubChem Compound Summary for CID 91469, Equol" *PubChem*, https://pubchem.ncbi.nlm.nih.gov/compound/Equol. Accessed 21 November, 2021.
- National Center for Biotechnology Information. "PubChem Compound Summary for CID 104775, Butyrate" *PubChem*, https://pubchem.ncbi.nlm.nih.gov/compound/Butyrate. Accessed 21 November, 2021.
- National Center for Biotechnology Information. "PubChem Compound Summary for CID 5281708, Daidzein" *PubChem*, https://pubchem.ncbi.nlm.nih.gov/compound/Daidzein. Accessed 21 November, 2021.
- National Center for Biotechnology Information. "PubChem Compound Summary for CID 5793, D-Glucose" *PubChem*, https://pubchem.ncbi.nlm.nih.gov/compound/D-Glucose. Accessed 21 November, 2021.
- Lee, Ho-Jae & Cha, Ji-Young. (2018). Recent insights into the role of ChREBP in intestinal fructose absorption and metabolism. BMB reports. 51. 10.5483/BMBRep.2018.51.9.197.
- https://go.drugbank.com/drugs/DB01914#
- Durack J, Lynch SV. The gut microbiome: Relationships with disease and opportunities for therapy. J Exp Med. 2019 Jan 7;216(1):20-40. doi: 10.1084/jem.20180448. Epub 2018 Oct 15. PMID: 30322864; PMCID: PMC6314516.

- Kho, Zhi Y, and Sunil K Lal. “The Human Gut Microbiome – A Potential Controller of Wellness and Disease.” *Frontiers in Microbiology*, vol. 9, no. 1664-302X, 2018, p. 1835., doi:10.3389/fmicb.2014.00075.
- https://www.optibacprobiotics.com/learning-lab/in-depth/gut-health/microbiome-all-you-need-to-know
- https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.biologydiscussion.com%2Fbacteria%2Fbacterial-taxonomy%2Fbacterial-taxonomy-meaning-importance-and-levels%2F54679&psig=AOvVaw3M2qPXFmp7uz9GEd94nVrD&ust=1637607935223000&source=images&cd=vfe&ved=0CAsQjRxqFwoTCMC88ouTqvQCFQAAAAAdAAAAABAD
- https://www.mayo.edu/research/labs/gut-microbiome/projects/gut-microbiota-host-physiology
- https://www.mdpi.com/2072-6643/11/8/1712/htm
- Zhang, Yu-Jie et al. “Impacts of gut bacteria on human health and diseases.” *International journal of molecular sciences* vol. 16,4 7493-519. 2 Apr. 2015, doi:10.3390/ijms16047493
- https://clinicaltrials.gov/ct2/show/NCT04879914
- https://pubmed.ncbi.nlm.nih.gov/31366236/
- https://academic.oup.com/advances/article/9/1/21/4849000
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3807999/
- https://academic.oup.com/advances/article/9/1/21/4849000
- https://pubchem.ncbi.nlm.nih.gov/compound/104775#section=Associated-Disorders-and-Diseases
- https://hmdb.ca/metabolites/HMDB0002209
- https://pubchem.ncbi.nlm.nih.gov/compound/91469#section=Pathways
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6770660/
- https://www.cambridge.org/core/journals/proceedings-of-the-nutrition-society/article/is-equol-production-beneficial-to-health/105B5CE52EA2ECC354143A20CFA8F080
- https://pubchem.ncbi.nlm.nih.gov/compound/91469#section=Associated-Disorders-and-Diseases
- pubchem daidzein
- https://go.drugbank.com/drugs/DB01914
- https://pubchem.ncbi.nlm.nih.gov/compound/5793#section=Associated-Disorders-and-Diseases
- pubchem dihydrodaidzein
- https://hmdb.ca/metabolites/HMDB0005760

- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6274073/
- https://ars.els-cdn.com/content/image/1-s2.0-S2090123219301833-ga1.jpg
- https://www.cell.com/trends/endocrinology-metabolism/fulltext/S1043-2760(20)30252-6
- https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.biologydiscussion.com%2Fbacteria%2Fbacterial-taxonomy%2Fbacterial-taxonomy-meaning-importance-and-levels%2F54679&psig=AOvVaw3M2qPXFmp7uz9GEd94nVrD&ust=1637607935223000&source=images&cd=vfe&ved=0CAsQjRxqFwoTCMC88ouTqvQCFQAAAAAdAAAAABAD
- Sun, Meng-Yao & Ye, Ying & Xiao, Ling & Rahman, Khalid & Xia, Wei & Zhang, Hong. (2016). Daidzein: A review of pharmacological effects. African Journal of Traditional, Complementary and Alternative Medicines. 13. 117. 10.4314/ajtcam.v13i3.15.