SYNERGISTIC HEMOLYSIS IN "NONHEMOLYTIC" BACTEROIDES SPECIES

SYNERGISTIC HEMOLYSIS IN "NON-HEMOLYTIC" BACTEROIDES SPECIES

By Hiba Shareefdeen, BSc (Honours), MCMASTER UNIVERSITY 2020

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the
Requirements for the Degree Master of Science

Descriptive Note

McMaster University MASTER OF SCIENCE (2020) Hamilton, Ontario (Medical Sciences)
TITLE: Synergistic Hemolysis in “Nonhemolytic” Bacteroides Species

AUTHOR: Hiba Shareefdeen, BSc. (McMaster University)

SUPERVISOR: Dr. Alexander Hynes

NUMBER OF PAGES: xi, 81

Lay Abstract

Bacteroides is a genus of abundant bacteria in the gut that confer multiple health benefits to the host. They are currently classified as nonhemolytic; meaning they are unable to lyse red blood cells. Work with *Bacteroides* isolates led to the observation, by happenstance, of unexpected hemolytic activity in multiple species. After incubation on blood agar plates, we observed that some isolates were clearly hemolytic, but only in close proximity to certain, 'activating' isolates. When we probed a larger collection 94 of isolates, we identified one which was able to activate hemolysis in 30% of the collection! Furthermore, the asymmetrical patterns of hemolysis were unlike any bacterial synergy reported to date. We used both a comparative genomics and mutagenesis approach to better understand this unique phenotype. Unexpected phenotypes such as the one investigated here highlight the complexity of the bacterial interactions that occur in, on and around us.

Abstract

Bacteroides is a genus of anaerobic bacteria that are often found in high abundance in the human colon. They have a complex relationship with their human host, conferring health benefits as members of the gut microbiota but also acting as opportunistic pathogens. Though many of the virulence factors in *Bacteroides* have been characterized, it is currently classified as nonhemolytic. Work with *Bacteroides* isolates led to the observation, by happenstance, of unexpected hemolytic activity in multiple species. After incubation on blood agar plates, we observed that some isolates were clearly hemolytic, but only when plated in close proximity to certain, 'activating' isolates. We performed a systematic screen for hemolytic activity on a library of 94 *Bacteroides* isolates and identified one which was able to activate hemolysis in 30% of the collection.

Using timelapse photography, we show this zone of hemolysis begins between the two colonies and proceeds in a retrograde fashion towards the 'activated' colony. The asymmetrical patterns of hemolysis are unlike any bacterial synergy reported to date. To investigate the mechanism behind this pattern, we combined a comparative genomics and mutagenesis approach to narrow down the genetic basis of this phenotype. Here, we characterize a unique synergistic hemolysis phenotype in *Bacteroides*, a genus of bacteria currently classified as nonhemolytic. This novel phenotype may provide further insight into *Bacteroides* as an opportunistic pathogen and may have uncovered a new mechanism by which they interact.

Acknowledgements

First and foremost, I would like to acknowledge my supervisor Dr. Alexander Hynes for his support over the past two years. His passion and enthusiasm for science made my time in the lab an enriching experience. I am so grateful for the time and effort Dr. Hynes put into mentoring and encouraging me. Thank you for your patience and your constant guidance.

It was also a pleasure to work with the great team at the Hynes lab. All the members of the lab are such insightful and collaborative scientists. I would not have been able to complete this thesis without their support. Rachelle Di Tulio and Gayatri Nair in particular, provided me with help and support every day. I am so grateful to them. I would also like to thank Anisha Nandy for going through this exciting process with me. Thanks for always taking the time to listen.

My committee members Dr. Elena Verdu and Dr. Michael Surette were integral to the progression of this project. I appreciated their valuable input and their stimulating discussions. I am so thankful to have had the opportunity to learn from such knowledgeable experts.

To Dr. Surette, in particular, thank you for your mentorship over the past four years. To have had your input through both my undergraduate and Master's degree was an absolute privilege. Thank you for making the time to keep supporting me. My time in the Surette lab inspired me to pursue graduate school.

Finally, thank you to my wonderful parents and my sister for being a constant source of trust and encouragement. I am so lucky to have the most amazing family.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

BHI - Brain Heart Infusion

BFT – *Bacteroides fragilis* Toxin

CBA – Columbia Blood Agar

CBB – Columbia Blood Broth

CDS – Coding DNA Sequence

DNA – Deoxyribonucleic Acid

GIL – Gut Isolate Library

GWAS – Genome Wide Association Study

LB – Luria-Bertani

OD – Optical Density

OMV – Outer Membrane Vesicle

ORF – Open Reading Frame

PCR – Polymerase Chain Reaction

PFT – Pore-forming toxins

RNA – Ribonucleic Acid

rRNA – Ribosomal RNA

SCFA – Short Chain Fatty Acid

SNP – Single Nucleotide Polymorphism

SRP-PCR – Semi-random primer Polymerase Chain Reaction

TE – Transposable Element

TIS – Transposon Integration Site

WGS – Whole Genome Sequencing

Declaration of Academic Achievement

I, Hiba Shareefdeen, declare this thesis to be my own work. I am the sole author of this document. The research associated with this thesis was completed by myself, with help from Rachelle Di Tullio (research assistant, McMaster University, who assisted with bioinformatics).

Chapter 1: Introduction

1.1 THE GUT MICROBIOTA

The human gut microbiota acts as a complex ecosystem with several functions; it is composed of bacteria, yeasts, fungi and viruses, all of which contribute to its diversity (Lozupone et al. 2012). In recent years, there has been increased focus on the gut microbiota and its role in obesity, diabetes, and liver diseases as well as cancer, mental illness and neurodegenerative diseases (Singer-Englar et al. 2019; Cheung et al. 2019). Perturbations that result in shifts of microbial communities in the gut are connected with various disease states (Tamboli 2004). The identification of specific bacteria that are involved in these interactions with the host is of particular interest. Understanding these interactions, and the bacteria involved, provides an opportunity to modulate the microbiome. Many groups are attempting to manipulate the gut microbiota through probiotics, antibiotics, diet and Fecal Microbiota Transplants (FMT). The human gut microbiota is a potential source of novel therapeutics, and groups are characterizing the communities that exist in the gut environment (Uchiyama et al. 2019).

In terms of the bacterial density, the human colon has the largest population of bacteria in the body, and the majority of these organisms are anaerobes (Clemente et al. 2012). One group of these anaerobes that

populate the gut is the highly abundant genus *Bacteroides*, which are the focus of this project.

1.2 BACTEROIDES

Bacteroides is a genus of anaerobic bacteria that are often found in the human gut, and account for 30% of bacteria found in the colon (De Vos et al. 2009; Garrett and Onderdonk 2014). In fact, the order *Bacteroidales* is the most abundant Gram-negative bacteria in the colon with densities up to 8×10^{10} CFU per g of feces (Zitomersky et al. 2011). *Bacteroides* species are bile-resistant, non-spore-forming, Gram-negative rods that are passed from mother to child during vaginal birth and thus become part of the human microbiota in the earliest stages of life (Wexler 2007). Within the intestinal tract *Bacteroides* species can influence host function by promoting the development of the immune system and regulating Paneth cell protein production (Hooper et al. 2003). Paneth cells are found in the small intestinal epithelium and secrete antibacterial peptides that affect luminal microbial ecology. The principal molecules secreted by Paneth cells are alpha-defensins. Comparisons of germ-free and *B. thetaiotaomicron*-colonized transgenic mice lacking Paneth cells established that microbial regulation of angiogenesis is affected by *B. thetaiotaomicron* (Stappenbeck et al. 2002). The role of *Bacteroides* colonizing a mucosal surface could involve regulating underlying microvasculature in post-natal development by signaling through

epithelial cells (Stappenbeck et al. 2002). In general however, the small intestine has less bacterial diversity and lower bacterial density than the colon. Though *Bacteroides* can be found in the small intestine, they are more highly abundant in the colon (Kastl et al. 2020).

Bacteroides species are well-adapted to the large intestine, where they can utilize simple and complex sugars as well as polysaccharides for growth. In the colon, *Bacteroides* are known occupy the ecological niche of the crypts (Saffarian et al. 2019). The bacteria maintain a complex and generally beneficial relationship with the host in the gut, and their role as commensals has been extensively reviewed (Wexler and Goodman 2017).

1.2.1 Role of *Bacteroides* in Human Health

The role of *Bacteroides* in human health is still being investigated, though studies have shown that certain *Bacteroides* species confer benefits to the human host. They play a role in carbohydrate fermentation in the gut, are linked with lower rates of obesity and are involved in preventing the intestinal colonization of potential pathogens (Wells et al. 1988; Bessesen 2007; Wu et al. 2011). Specifically, *Bacteroides* are involved in colonization resistance, a process that prevents pathogen invasion in host environmental niches. Jacobson *et al.* found that production of propionate by *Bacteroides* mediates colonization resistance to *Salmonella* by disrupting intracellular pH homeostasis (Jacobson et al. 2018).

In general, *Bacteroides'* ability to produce short chain fatty acids (SCFAs) is important to their role as gut commensals. These SCFAs are integral to bacterial cross-feeding networks that maintain community structure and are also involved in microbiota-host interactions. For example, *Bacteroides thetaiotaomicron* starch utilization and interspecies cross-feeding promotes butyrate production by *Eubacterium ramulus* (Rodriguez-Castaño et al. 2019). Butyrate is used as an energy source by colonocytes and can improve epithelial barrier function (Donohoe et al. 2011; Rios-Covian et al. 2017).

In addition to their role in SCFA production, *Bacteroides* have been shown to impact atherosclerosis. Isolates of *Bacteroides vulgatus* and *Bacteroides dorei* reduce gut microbial lipopolysaccharide production and inhibit atherosclerosis. A 2018 study found that these species attenuated atherosclerotic lesion formation in atherosclerosis-prone mice, decreased gut microbial lipopolysaccharide production and suppressed proinflammatory immune responses (Yoshida et al. 2018).

Bacteroides also may play a role in chronic gastrointestinal conditions such as Inflammatory Bowel Disease (IBD). One meta-analysis suggests that lower levels of Bacteroides are associated with IBD, especially in the active phase of the disease (Zhou and Zhi 2016). However, evidence for a pro-inflammatory role for *Bacteroides* also comes from animal models for IBD where *Bacteroides* spp. appear to play a role in the

induction of colitis (Onderdonk et al. 1983). *Bacteroides fragilis* enterotoxin induces human β-defensin-2 expression in intestinal epithelial cells (Yoon et al. 2010). Additionally, a high amount of adherent *Bacteroides fragilis* was found in the mucosal biofilm in patients with IBD (Swidsinski et al. 2005).

Bacteroides species are normally commensals in the gut environment yet they can also be opportunistic human pathogens (Wexler 2007). Given the opportunity to escape from the intestinal tract into the normally uncolonized peritoneal cavity, some *Bacteroides* species can cause life-threatening infection including bacteremia (Stone et al. 1975).

1.2.2 Virulence Factors in *Bacteroides*

Although the factors that determine virulence in *Bacteroides* are unclear, capsular polysaccharides, microbial adherence and production of proteases are considered important (Gibson et al. 1996). In general, virulence factors are components that facilitate a pathogen's ability to colonize the host. For example, some *Bacteroides* species are able to modulate their surface polysaccharides (Krinos et al. 2001). By changing its surface architecture, *Bacteroides* may avoid the host immune response (Krinos et al. 2001; TerAvest et al. 2014). Other potential effects of surface modification include the ability of *Bacteroides* to better colonize host tissues and form biofilms (Krinos et al. 2001). The expression of genes

involved in polysaccharide utilization and capsule biosynthesis is upregulated in biofilms (TerAvest et al. 2014).

Another virulence determinant is the presence of secretion systems. Some bacteria directly kill their competitors by releasing chemical compounds or by secreting effectors via specific secretion systems such as the type VI secretion system (T6SS) (Bingle et al. 2008). T6SSs are contact-dependent secretion machineries capable of directly injecting toxins into other bacteria as well as eukaryotic cells. *B. fragilis* uses its T6SS to antagonize numerous *Bacteroidales* species isolated from the human gut (Coyne and Comstock 2019). As *Bacteroidales* make up a large portion of colonic bacteria, it is suggested that T6SSs are distributed in about 25% of bacteria in the human colon (Coyne et al. 2016; Chatzidaki-Livanis et al. 2017). It is predicted that there are more than 10^9 T6SS-firing events per minute occurring per g of colonic contents, further supporting the importance of this system in shaping gut microbiota composition (Wexler et al., 2016).

Most of the work focusing on the virulence factors in *Bacteroides* looks at the species *Bacteroides fragilis*. *B. fragilis* infection is the leading cause of anaerobic bacteremia and sepsis (Wells et al. 1988; Cheng et al. 2019). Enterotoxigenic *B. fragilis* produces a secreted metalloprotease known as *B. fragilis* toxin (BFT), which contributes to anaerobic sepsis, colitis, and colonic malignancy in mouse models of disease. This zinc

metalloprotease may destroy the tight junctions in the intestinal epithelium by cleaving E-cadherin. The result is that this barrier leaks and results in diarrhea (Wick and Sears 2010) . These enterotoxigenic *B. fragilis* strains encode BFT on distinct bft loci, carried on the *B. fragilis* pathogenicity island (Franco 2004). A cysteine protease, fragipain (*Fpn*), directly activates BFT in the bacterial cell by removing a prodomain of the toxin. *Fpn* is a proenzyme and is autoactivated upon cleavage at an arginine residue in its activation loop (Wells et al. 1988; Herrou et al. 2016). BFT was shown not to be a freely secreted protease but is rather associated with outer membrane vesicles (OMVs) (Zakharzhevskaya et al. 2017). This study also indicated that only outer surface-exposed BFT causes epithelial cell disruption.

Bacteremia due to anaerobic microorganisms represents 2–3% of positive blood cultures performed in hospitalised patients (Robert et al. 2008). Enterotoxigenic strains that produce *B. fragilis* toxin contribute to colitis and intestinal malignancy and are also isolated in bloodstream infections. However, one virulence determinant that has not been well characterized in *Bacteroides* is hemolysis, the ability to lyse red blood cells. Although there have been reports of a few strains of *B. fragilis* being hemolytic, *Bacteroides* as a genus is classified as nonhemolytic (De Vos et al. 2009).

1.3 HEMOLYSIS

Hemolysis is the breakdown of red blood cells, also termed "erythrocytes". In the context of human body, hemolysis can be caused by pathogens, certain autoimmune disorders and some genetic disorders such as sickle-cell disease and thalassemia (Muncie and Campbell 2009). The release of hemolglobin into the blood plasma results in hemoglobinemia and increased risk of infection (Kalfa 2014).

Hemolysis is categorized into three types: alpha, beta and gamma. Hemolysis testing is performed by observing the effect of bacterial growth on an agar medium supplemented with blood. Alpha hemolysis, or partial hemolysis, is often caused by hydrogen peroxide production and results in a green hue around the colonies (Buxton 2005). Beta hemolysis, or complete hemolysis, is a complete lysis of red blood cells in the media around and under the colonies so the area appears lightened or transparent (Buxton 2005). The term gamma hemolysis is used when an organism is nonhemolytic and the agar is unchanged (Fischetti and Ryan 2015).

The production of hemolysin, a component which breaks down red blood cells, is a potential virulence factor produced by microorganisms which can be especially dangerous in the context of infection. In fact, hemolysins have been reported to be powerful virulence determinants in both Gram-positive and Gram-negative bacteria (Welch 1991). These are often cytotoxic proteins that target cell membranes, and can damage

membrane integrity. They are classified as enzymatics, pore-forming cytolysins, or surfactants (Robertson et al. 2006). Enzymatically active hemolysins such as α-toxin of *Clostridium perfringens*, are often lipases which can cleave lipids in plasma membranes of the host cells (Titball et al. 1999). Pore-forming toxins (PFTs) are the most common bacterial cytotoxic proteins and are important virulence factors in pathogens, including *Streptococcus pneumoniae*, *Staphylococcus aureus*, and *Mycobacterium tuberculosis* (Los et al. 2013). Pore-forming toxins work by disrupting cell membranes. The small-pore-forming repeat in (RTX) toxins, produced by Gram-negative bacteria, form a large group of PFTs. Some bacteria produce surfactants which cause hemolysis by osmotic lysis and solubilization (Manaargadoo-Catin et al. 2016). Hemolysins are a specific type of cytolysin. Cytolysins in general can lyse many different cell types. For example, cytolysins may offer an advantage to the bacteria by lysing incoming leukocytes in order to promote survival of the bacteria by weakening the immune system and gaining access to nutrients (Cavalieri and Snyder 1982; Sagar et al. 2013).

1.4 BACTERIAL INTERACTIONS

Bacteria can cooperate in many different ways, and the degree in which they do this can vary from only marginal support to mutual dependence. They can secrete metabolites that are used by a partner organism that lacks specific synthesis pathways and profits from this

support (Schink 2002). An example of this type of behaviour is bacterial cross-feeding in the infant gut. *Bifidobacterium bifidum* is able to degrade mucin and release monosaccharides which can be used by other gut microbes colonizing the infant gut (Bunesova et al. 2018). Even if a bacterium could generate the respective compound it needs on its own, it saves biosynthetic energy through synergistic interactions.

Some culturing methodologies may be missing these interactions if bacteria are isolated with media that selects for easy-to-cultivate organisms (Schink 2002). However, phenotypic culturing is now linked to large-scale whole-genome sequencing as well as phylogenetic analysis (Vartoukian et al. 2010). Combining culture-based methodologies with sequencing information has indicated that the majority of intestinal bacteria are culturable (Browne et al. 2016). We still may be overlooking other bacteria that are outcompeted in these conditions and that display novel interactions in the presence of other species. Certain species might depend largely on cooperation with partners, and this could account for why certain bacterial groups are difficult to cultivate. Members of the widespread photosynthetic marine bacterial genus *Prochlorococcus* have been cultured from the ocean many times, but only two variants had ever been grown in pure culture. Other isolates depend on heterotrophic bacteria for coculture, and it is difficult to grow these organisms from a single cell as a colony on a petri

dish (Morris et al. 2008). A number of efforts to culture "unculturable" bacteria has revealed multiple coculture-dependent isolates.

Bacteria in the environment do not function in isolation. Bacteria are capable of sustaining complex structures in polymicrobial environments and interact using a wide array of behaviours as a form of communication, cooperation or competition (Miller and Bassler 2001; Zheng et al. 2015). Bacterial interactions can be antagonistic, especially when bacteria compete for similar niches. They may secrete compounds such as bacteriocins to kill a competing population (Zheng et al. 2015). However in the majority of cases, it seems that most bacteria largely ignore one another and do not actively interact. By and large, we associate a bacterium's defining traits as a function of its chemical environment, only sometimes studying how it interacts with biological components such as the host in which it lives. By studying a bacterium in isolation, we fail to examine how the presence of other microorganisms in its environment might impact its behaviour.

1.4.1 Synergistic Hemolysis

Synergistic hemolysis, an interaction where a bacterium can potentiate the hemolytic activity of another, is most commonly characterized between *Staphylococcus* and *Streptococcus* species (Hebert and Hancock 1985). The hemolytic activity of staphylococcal beta-hemolysin on red blood cells is enhanced by an extracellular factor

produced by Group B *streptococci*, called the CAMP (Christie-Atkins-Munch-Peterson) factor (Munch-Petersen & Christie, 1945).

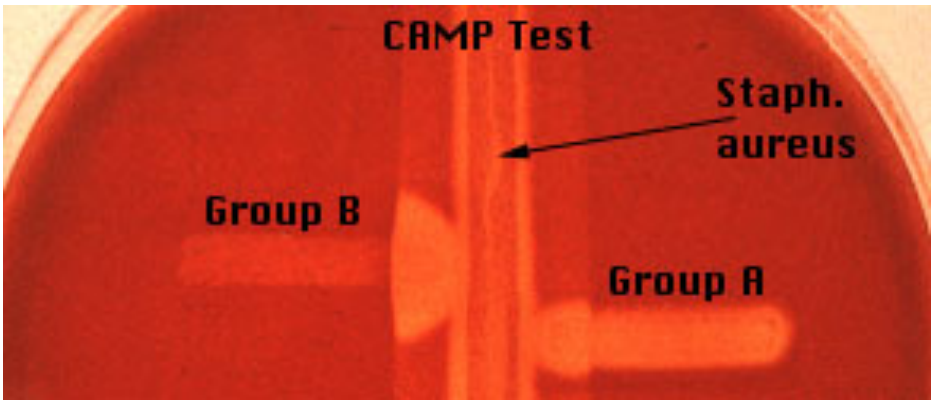

Figure 1. CAMP Test.

Hemolytic interactions between Group B streptococci species and *S. aureus* are well established. Group A and B *Streptococcus* are streaked perpendicularly to *S. aureus* (vertical) Wherever the two reactants overlap on agar plate with sheep's blood, a larger zone of beta-hemolysis in the shape of an arrowhead will be present.

Image used with permission, LSUMC/MIP/Harriet Thompson. (LSUMC 1995)

The CAMP test is used to identify Group B hemolytic streptococci based on their formation of this factor which enlarges the area of hemolysis formed by *Staphylococcus aureus* (Wilkinson 1977). The CAMP test detects the production of this diffusible, thermostable, extracellular protein produced by the Group B *Streptococcus*. The standard CAMP test results depend on the pattern of the two components during growth to form a

typical arrowhead-shaped clearing at the point where the two organisms are placed perpendicular to each other. In the reverse CAMP test, *Clostridium perfringens* is identified by streaking it perpendicularly to a Group B *Streptococcus* (Hansen and Elliott 1980). The positive reaction identifying *Clostridium perfringens* is marked by the "bow-tie" zone of enhanced hemolysis pointing towards the Group B *Streptococcus.* The established models of synergistic hemolysis have been well characterized between *Staphylococcus*, *Listeria*, and *Clostridium* species (Hansen and Elliott 1980; Holth-Haug and Olsvik 1981; McKellar 1994). To date, hemolysins have only been characterized in a few *B. fragilis* strains (Robertson et al. 2006) and synergistic hemolysis has not been reported in any *Bacteroides* species.

1.5 NOVEL SYNERGISTIC INTERACTION BETWEEN *BACTEROIDES* ISOLATES

Working with the *Bacteroides* isolates from the Surette Lab Gut Isolate Library collection led to the observation, by happenstance, of unexpected hemolytic activity in multiple species. After incubation on blood agar plates, which is often used for culturing *Bacteroides*, we observed that some isolates were clearly hemolytic, but only when plated in close proximity to certain other isolates.

1.6 HYPOTHESIS & AIMS

This thesis explores the hypothesis that **synergistic interactions in *Bacteroides* isolates from the human gut result in hemolytic phenotypes**. The work presented here investigates the discovery and characterization of a novel hemolytic phenotype discovered in multiple species of *Bacteroides*. As hemolysis in *Bacteroides* has not been well-characterized, I study both the prevalence of this phenotype and its unusual presentation. I attempt to identify key components driving this synergistic interaction. Through comparative genomics and mutagenesis approaches, I try to determine the genetic basis of this interaction.

Chapter 2: Materials and Methods

2.1 CULTURING BACTEROIDES

2.1.1 Culturing *Bacteroides* Isolates

The *Bacteroides* isolates used in this study were isolated from human stool from both healthy donors and patients with Irritable Bowel Syndrome. The isolates included 10 species and were obtained from the Surette Lab Gut Isolate Library (McMaster University) as stocks frozen in 10% skim milk (Table S1). Duplicates of frozen stocks were grown in BHI (Brain Heart Infusion) broth and were frozen with 15% glycerol and stored at -80°C. Work was done in an anaerobic chamber (Bactron IV, Sheldon Manufacturing, Oregon), using the manufacturer recommended gas mix: 5% CO2, 5% H2, 90% N2 (Praxair, Connecticut).

Growth on Solid Media

Bacteroides isolates were grown either on Columbia Blood Agar (CBA) plates (Becton Dickinson, Germany) supplemented with 5% defibrinated Sheeps' blood (Cedarlane, Canada) or BHI (Brain Heart Infusion) agar (Becton Dickinson, Germany) and incubated at 37°C in an anaerobic chamber.

Growth in Liquid

Bacteroides isolates were grown in test tubes with 10 ml of BHI supplemented with l-cysteine (0.5 g/L final concentration) and incubated at 37°C in an anaerobic chamber. For growth in 96-well plates, isolates

were inoculated into 150 µl of BHI supplemented with l-cysteine. After a sterile porous adhesive sheet was fixed to the top of the plate to prevent evaporation, the plate is incubated at 37°C in an anaerobic chamber.

2.1.2 DNA Extraction of Isolates

In preparation for whole genome sequencing, DNA was extracted from the isolates in Table S2. DNA was extracted from 1 ml of overnight culture using the Wizard Genomic DNA Purification Kit (Promega, Wisconsin).

2.2 HEMOLYSIS TESTING

To test for hemolytic activity I spotted 1 µl of overnight culture onto CBA supplemented with 5% Sheep's blood. After incubating for 36 hours at 37°C anaerobically I would observe the plates for zones of clearing which indicated red blood cell lysis.

2.2.1 Screening of Hemolysis Interactions in *Bacteroides* GIL Collection

To screen the *Bacteroides* isolates of the GIL collection for hemolysis, the isolates were arrayed in 96-well plates with each isolate next to "activating" isolate GIL820 (Figure S1). The arrayed plates were stamped onto CBA and incubated anaerobically for 36 hours at 37°C. After incubation, plates were imaged using an iPhone X (Apple, California).

2.3 CHARACTERIZATION OF HEMOLYSIS

2.3.1 Timelapse Imaging

To better understand the spatial and temporal variability behind this pattern of hemolysis, we created a custom camera set-up to create timelapse videos of hemolysis interactions. The camera, a Canon EOS Rebel T5 (Canon, Tokyo), was set to capture an image of the plate every 30 minutes using a Neewer LCD Timer Shutter Release Remote Control (Neewer, Shenzhen) intervalometer for 48 hours in a 37°C anaerobic environment. The plate was backlit with an LED lightpad.

2.3.2 Supernatant Experiments

Generating Culture Supernatants

B. vulgatus GIL820 and *B.ovatus* GIL819 were grown in both BHI broth, and Columbia Blood Broth (Becton Dickinson, Germany) supplemented with 5% Sheep blood. Cultures were grown anaerobically at 37°C overnight. 2 mL of culture was centrifuged at 8000 x g for 2 minutes. Supernatant was aspirated off and filter sterilized with 0.45 μm PES syringe filter. *B. vulgatus* GIL820 and *B. ovatus* GIL819 were also co-cultured in BHI and Columbia Blood broth to generate a co-culture supernatant.

Generating Concentrated Culture Supernatants

Concentrated culture supernatants were generated as above, but were concentrated six-fold using an Amicon Ultra 0.5ml 3K centrifugal filter (Millipore Sigma, Massachusetts) as per manufacturer's instructions with a 15 minute spin time at 14000 x g.

Determining the Effect of Supernatants on Hemolysis

To test if bacterial supernatants were sufficient to affect hemolysis, cultures were grown in different combinations (Table 1) and then filtered supernatants were prepared. Both Columbia Blood Broth (CBB) and BHI media types were used to generate supernatants. CBB was chosen to account for the differences between CBA and BHI. 2 µl of GIL819 culture was spotted next to 2 µl of each type supernatant of on a CBA plate. The same procedure was repeated for GIL820. The supernatants from the "GIL819 + GIL820 Mixed" consisted of GIL819 and GIL820 grown and filtered separately and then mixed afterwards in 1:1 ratio.

Table 1: Media Conditions used for Supernatant Experiments.

	Media			
	Filtered Only		**Filtered and Concentrated**	
GIL819	BHI	CBB + Blood	BHI	CBB + Blood
GIL820	BHI	CBB + Blood	BHI	CBB + Blood
GIL819 + GIL820 Co-cultured	BHI	CBB + Blood	BHI	CBB + Blood
GIL819 + GIL820 Mixed	BHI	CBB + Blood	BHI	CBB + Blood

2.3.3 Effect of pH Alteration on Hemolysis

Before autoclaving CBA media, the pH was adjusted with 5 M HCl and 5 M NaOH. I created CBA plates at 4 different pH points (pH 5.5, pH 6.5, pH 7.5, pH 8.5) to determine if altered pH was involved in activating

hemolysis in GIL819. 3 µl of GIL819 overnight culture were spotted on these plates.

2.4 BIOINFORMATICS

2.4.1 Bacterial Identification via Sanger Sequencing of 16S rRNA gene

Before proceeding with whole genome sequencing of isolates, species identity was confirmed by performing colony PCR of the 16S rRNA gene (Table S2) and performing BLASTn on the Sanger Sequencing results (McMaster Genome Sequencing Facility) using the NCBI nucleotide collection non-redundant (nr) database (Altschul et al. 1990; BLAST 2009).

2.4.2 Whole Genome Sequencing

A next-generation sequencing (NGS) library preparation was performed using the Nextera XT DNA Library Preparation kit (Illumina, San Diego) by the McMaster Genome Sequencing Facility. Sequencing was performed on a MiSeq platform with 2 x 250bp reads for all the samples in Table S2.

2.4.3 Genome Assembly and Annotation

FastQC was used before and after read trimming to assess the quality of the Illumina paired end reads, [version: FastQC v0.11.8] (Andrews 2015). Trimmomatic was used to trim raw reads of low quality bases and remove adaptor sequences, [version: 0.39] (Bolger et al. 2014). Unicycler was used to assemble quality assessed reads into contigs [version: Unicycler v0.4.8](Wick et al. 2017). The assembled genomes of each

sequenced isolate was annotated using two different tools, RAST and Prokka (Aziz et al. 2008; Seemann 2014). RAST (Rapid Annotation using Subsystem Technology) identifies protein-encoding, rRNA and tRNA genes, assigns functions to the genes and predicts which subsystems are represented in the genome. Prokka which uses a suite of existing software tools to annotate sequences of bacterial genomes (Seemann 2014). Quast was used to assess the quality of the contigs assembled, and key quality metrics are highlighted in Table S2 (Gurevich et al. 2013). Six additional *B. vulgatus* genomes were obtained from an additional strain collection at the Surette Lab (GC Collection) and added to the analysis (Section 4.1.2).

2.4.4 Comparative Genomics

Roary was used to perform pangenome analysis with the default settings. This tool rapidly builds large-scale pan genomes and identifies the core and accessory genes (Page et al. 2015). The core genes as defined by the tool are genes that are in 99% of isolates. The pangenome is the entire gene set of all strains of a species. Scoary was used to perform assessment of trait-associated variants in pan-genome. This tool is used for studying the association between pan-genome gene presence or absence and observed phenotypes. Each candidate gene in the accessory genome is sequentially scored according to its apparent correlation to predefined traits (Brynildsrud et al. 2016) Based on the phenotype data of the sequenced isolates, we created a binary traits table, required by SCOARY to combine with the pangenome. Isolates were classified as “1” if they had been able

to activate any other isolate. Isolates that were only activatees were scored as “0”.

Table 2: Binary Traits Table of Sequenced Isolates

ISOLATE	ACTIVATOR PHENOTYPE
GIL726	1
GIL820	1
GIL849	1
GIL914	1
GC195	0
GC196	0
GC197	0
GIL819	0
GIL102	0
GIL97	1
GIL418	0
GIL504	0
GIL552	0
GIL629	0
GC268	0
GC242	0
GC397	0

2.5 MUTAGENESIS OF BACTEROIDES VULGATUS

In combination with whole genome sequencing, transposon mutagenesis is powerful technique that enables exhaustive identification of genes that contribute to a bacterial phenotype. It is a means of producing random mutations in bacterial genomes. This is usually accomplished

using a suicide vector including a transposon bearing a selectable marker, from this plasmid a transposon is extracted and inserted into the host chromosome (Hayes 2003).

2.5.1 Generating Mutant Library

Confirming Sensitivity

Before proceeding with the mutagenesis protocol, I confirmed antibiotic sensitivity by plating GIL819, GIL820 and *E.coli* S-17 λpir-pSAM-Bt on BHI + Erythromycin (10 μg/ml), BHI + Gentamycin (25 μg/ml) and BHI + Ampicillin (50 μg/ml) and checking for growth on the plates after incubation overnight at 37°C.

Generating Mutants

I performed transposon mutagenesis by adapting the *B. fragilis* mutagenesis protocol created by Veeranagouda et al (Veeranagouda et al. 2019). This protocol uses pSAMBt, a mariner transposon mutagenesis vector which has been used for the transposon mutagenesis of *B. thetaiotaomicron* (Goodman et al. 2009; Veeranagouda et al. 2012)

E. coli S-17 λpir-pSAM-Bt was inoculated from an overnight culture into 5 ml of LB broth. For *B. vulgatus*, 5 ml of culture was inoculated into 50 ml of sterilized BHI broth in three 250 ml Erlenmeyer flasks in the anaerobic chamber. Cultures were then incubated at 37°C for 4 hours. After incubation, 1 ml of *E. coli* S-17 λpir-pSAM-Bt was added to 10 ml of *B. vulgatus* culture in 15 ml falcon tubes. Culture broth was then centrifuged at 5000 x g for 5 minutes. Cells were pelleted and then resuspended in 100 μl of BHI broth. The cell suspension was spread on BHI plates and

incubated anaerobically at 37°C overnight. After incubation, 1 ml of sterile BHI + 10% glycerol broth was added to each culture plate and the cells were scraped off using a sterile spreader. This cell suspension contained the mutant library. The mutant library was aliquoted into sterile 2 ml tubes and stored at -80°C for long-term storage.

Creating Mutant Library

B. fragilis mutagenesis transposon vector pSAM-Bt is a mariner transposon vector (Figure S3). It replicates only in *E. coli*, and when mobilized into *Bacteroides*, acts as a suicide vector. Only the transposon within inverted repeats containing an erythromycin cassette is integrated at TA dinucleotide sites in the *Bacteroides* genome. Genes conferring ampicillin resistance serve as the selection marker for *E. coli*, whereas erythromycin genes serve as the selection marker for *Bacteroides* transposon mutants. *Bacteroides vulgatus* GIL820 is resistant to gentamycin but the *E. coli* is not. Accordingly, the mutant library suspension was plated on CBA containing erythromycin and gentamycin to select for transposon mutants which gain erythromycin resistance by transposon integration.

The mutant library was thawed on iced, and diluted in a 1:4 ratio with BHI + Erythromycin (10 μg/ml) + Gentamycin (25 μg/ml) broth. 100 μl of this mixture was plated on CBA + Erythromycin (10 μg/ml) + Gentamycin (25 μg/ml) using sterile glass beads and incubated anaerobically at 37°C for 48 hours.

After incubation, mutant colonies grew on the agar plates. 9600 individual colonies were picked into wells in a 96-well plate containing 150 µl of BHI + Erythromycin (10 µg/ml) + Gentamycin (25 µg/ml) broth. After growth for 48 hours anaerobically at 37°C, 125 µl of sterile 30% glycerol was added to the wells. The plate was then stored at -80°C.

2.5.2 Screening Mutant Library

To screen the mutant library for mutants of GIL820 where hemolysis was affected, I tested each mutant by spotting 1 µl of culture next to GIL819 on a CBA plate. Since the mutants were frozen in 96-well plates I stamped the isolates using a 96-pin replicator (Figure S2). Initial optimization screening revealed staggering the pin replicator to offset cultures of mutant isolates with GIL819 produced edge effects, perhaps due to the colonies being too close together (Figure S5). I also tested stamping the plates outside of the chamber and passing them into the anaerobic chamber within 15 minutes, and did not observe any effects on bacterial growth. Subsequent stamping experiments were performed outside of the chamber.

After optimization, I modified disposable polypropylene pin-replicators (V&P Scientific, California) to only access 48-wells at a time, from every other column. This resulted in pin replicators that would stamp either every odd column (Position A), or even column (Position B) from a 96-well plate. These replicators were used to stamp all the mutants from a plate using Position A next to GIL819 (in Position B) and vice versa.

Essentially, each plate of mutants was screened on two CBA plates in the two orientations as detailed below in Figure 2.

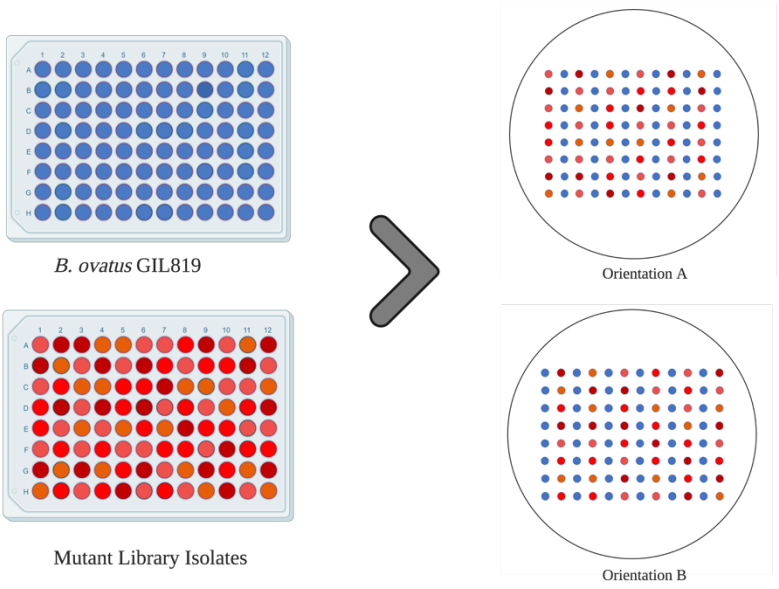

Figure 2. Schematic showing the positioning of the mutant library screening protocol.

After autoclaving the modified pin replicators, I stamped plates of mutants from the frozen plates onto CBA with GIL819 cultures grown overnight in broth in a 96-well plate. Plates were incubated for 48 hours at 37°C in the anaerobic chamber and then imaged.

2.5.3 Identification of Transposon Insertion Sites

After screening the isolates, I created lysates of the isolates of interest. A portion of the mutant colony was transferred into a tube containing 50 µl of water and put in thermocycler at 95°C for 10 min.

To perform the first step of the nested PCR to identify the transposon integration site, 23 µl of supernatant from the colony lysate was used as template. Primers SRP1, SAMSeq1 , and OneTaq Hot Start 2x master mix (NEB, Massachusetts) were added to perform first-round PCR (Table S4) as previously described (Veeranagouda et al. 2019). After the first cycle, I transferred 1 µl of the first PCR mix into a new tube with primers SRP2 and SAMSeq2 primer, and OneTaq Hot Start 2x master mix (NEB, Massachusetts) for the second-round of amplification (Table S5). Following the amplification, 5 µl of product was run on a 1% agarose gel and visualized with ethidium bromide.

In the first step of the nested PCR, SAMSeq1 primers bind to vector DNA, while the other one is a semirandom sequence (SRP1) which binds to random sequences on mutant DNA. SRP1 consists of an arm sequence, ten random bases (N) followed by five known bases (GATAT). The arm sequence serves as an anchor for the second PCR while ten random bases facilitate random binding to mutant genomic DNA. The steps in the thermocycler amplify the transposon integration site. In the second reaction SAMSeq2 binds downstream of SAMSeq1, and SPR-2 binds to the arm region of SPR-1. Thus, the second PCR preferentially enriches transposon integration site of mutant DNA (Table S4).

The PCR product was purified using Monarch PCR & DNA Cleanup Kit (NEB, Massachusetts) and sequenced using the SAMseq3 sequencing primer (Table S2). The mutated gene was identified by performing Blastn or Blastx analysis on the sequence against the non-redundant nucleotide collection as well as the sequenced genome of GIL820.

2.5.4 Analysis of Hits

The output from the Sanger sequencing of the hits was mapped to the annotated genome of GIL820 using Geneious [version 2019.2] by using the "Map to Reference" function with the Highest Sensitivity option selected (Biomatters Limited 2018). Sixteen out of nineteen mutants mapped back to GIL820. The remaining mutants were mapped back to a different *Bacteroides vulgatus* genome (*Bacteroides vulgatus* complete genome ATCC 8482). After genes were mapped to the reference, downstream analysis on the genes and operons of interest was performed using the following tools:

- BioCyc is a collection of 17043 Pathway/Genome Databases, with tools used to analyze them (Krummenacker et al. 2005; Karp et al. 2019). Protein sequences from mutants of interest were analyzed through the BioCyc online portal. BioCyc uses data curated from publications and computationally predicted metabolic pathways and operons. The output information includes data from other databases including essential genes, regulatory networks, protein features, and GO annotations.

- Phyre2 produces a set of potential 3D models of proteins based on alignment to known protein structures (Kelley et al. 2015). Translated proteins from mutants of interest with unknown function were uploaded onto the Phyre2 portal to look for structural similarity.

- SWISS-MODEL is a web-based integrated service dedicated to protein structure homology modelling (Schwede 2003; Biasini et al. 2014). Models are computed by the SWISS-MODEL server homology modelling pipeline which uses an in-house comparative modelling engine. Translated proteins from mutants of interest with unknown function were uploaded to the SWISS-MODEL pipeline using their online interface.

2.5.5 Mapping hits to Sequenced Isolates

I looked for hits in all sequenced isolates by building a custom BLAST database with the whole genome sequencing information we had in Table 2. Sequencing results of the TIS from the SRP-PCR were mapped to this database by using BLAST analysis to look for the gene of interest in the database of all the genomes.

Chapter 3: Identifying and Characterizing a Novel Hemolysis Phenotype

3.1 UNIQUE PATTERN OF HEMOLYSIS IN SUBSET OF ISOLATES

Unrelated work with the *Bacteroides* isolates from the Surette Lab Gut Isolate Library led to the observation, by happenstance, of unexpected hemolytic activity in multiple species. After incubation on blood agar plates, I observed that some isolates were clearly hemolytic, but only when plated in close proximity to certain, 'activating' isolates. The location and shape of the zones of hemolysis suggested an interaction far more complicated than simply induction of a new function in one isolate by the presence of the other. In particular, I discovered a synergistic hemolysis interaction between two *Bacteroides* isolates, numbered GIL819 and GIL820 (Figure 3A). After conducting Sanger sequencing of the full length 16s rRNA gene, we determined that the isolate GIL820, a *Bacteroides vulgatus* strain was able to 'activate' hemolysis in GIL819, *Bacteroides ovatus* (Figure 3B). Synergistic hemolysis has not previously been characterized in *Bacteroides*, prompting the focus of this project to be investigating this novel phenotype.

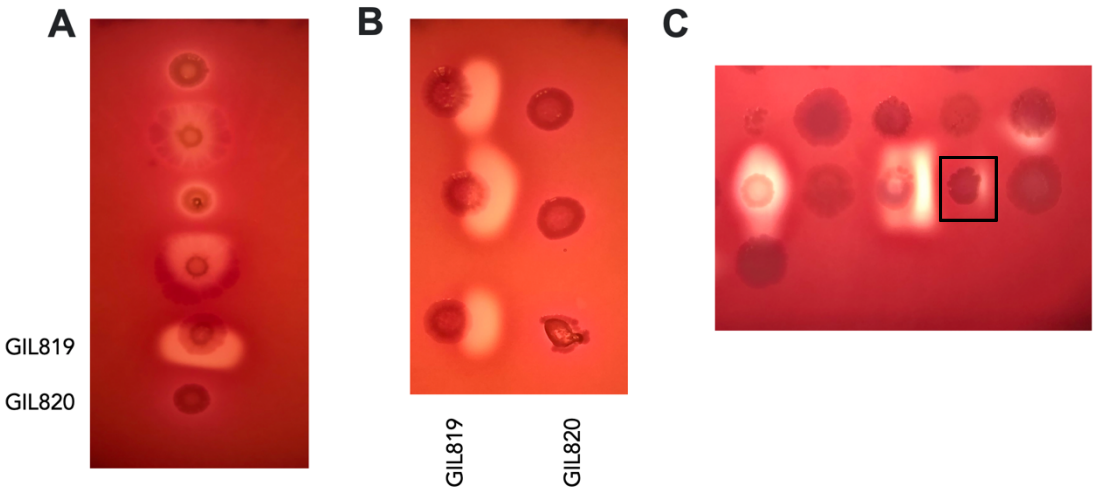

Figure 3. Diversity of hemolysis phenotypes in *Bacteroides* isolates.

(A) Six different *Bacteroides* isolates grown on blood agar plates show a diversity in hemolytic phenotypes. (B) Synergistic hemolysis between GIL819 (left) and GIL820 (right). (C) *Bacteroides* isolates display multiple patterns of hemolytic interactions. A black box highlights an isolate being both an activator and an activatee.

Figure 3B shows that this zone of clearing only occurs on one side of the activated colony and occurs in a semi-circular pattern. The directionality of this pattern is different from the interaction occurring in the CAMP test (Figure 1). These observations may indicate the presence of two substances, produced by two "nonhemolytic" colonies that act synergistically to produce hemolysis. The diffusion of substances in the agar may be involved in this process. Bacterial exotoxins, are an example of a diffusible toxin secreted by bacteria. Additionally, these observations are different from the induction of hemolysis in "normally nonhemolytic" *Staphylococci* on blood agar. In this case one group of *Staphylococci* are

only hemolytic in a secondary zone of hemolysis created by other *Staphylococcus* species (Smith et al. 1964)

Additionally, this phenotype was of particular interest due to the complexity of the interactions involving multiple isolates. In Figure 3C, there is evidence of an isolate activating hemolysis in a neighbouring colony, while, in turn, being activated by a third isolate. The directionality of the zones of hemolysis provide a clear indication of which isolate is in interaction with which.

3.2 CAMP TEST IN BACTEROIDES ISOLATES

To observe the differences between well-studied patterns of synergistic hemolysis and this novel phenotype, I streaked *B. vulgatus* GIL820 and *B. ovatus* GIL819 perpendicular to each other on a blood agar plate, in a manner similar to a CAMP test. The resulting zone of hemolysis was very different from the arrowhead pattern seen in the CAMP test. Instead, a semicircular zone of hemolysis is produced at the junction between GIL819 and GIL820.

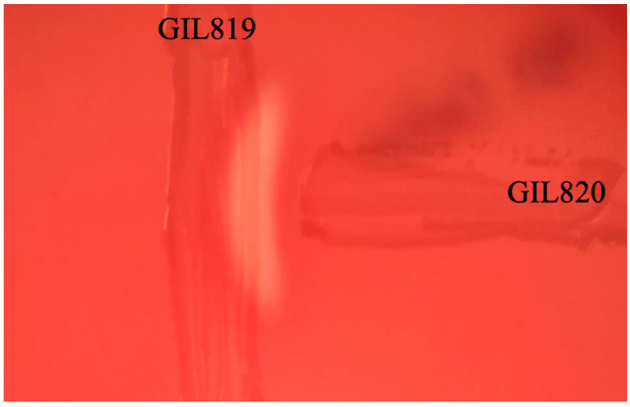

Figure 4. "CAMP" Test with *Bacteroides* GIL819 and GIL820.
B. vulgatus (horizontal) streaked perpendicular to *B. ovatus* (vertical) on CBA.

3.3 HEMOLYSIS WIDESPREAD IN BACTEROIDES ISOLATES

The incredible range of diversity in the hemolysis phenotypes prompted us to explore further. In particular, I was eager to characterize the activation of hemolysis in *Bacteroides* isolates by an "activator" strain. I encountered two main phenotypes, "activator" isolates that are able to turn on hemolysis in an "activatee". I focussed on determining if hemolysis was widespread in the collection of isolates. Specifically, I wanted to see how many isolates were able to become hemolytic when next to our activator GIL820. Our widespread screen showed that this activator isolate was able to turn on hemolysis in 30 out of 94 isolates it was spotted next to. The patterns of hemolysis show a great deal of diversity in the different shapes of the hemolytic zones. The peculiar patterns seemed dependent on which isolates were in close proximity to

one another. The shapes of the hemolytic zones were varied, including rectangular, semi-circular, and crescent clearings (Figure 5).

Through our screen, we found that there is some sort of hemolytic activity in six out of ten different species present in this collection. *Bacteroides xylanisolvens*, *Bacteroides vulgatus*, *Bacteroides ovatus* , *Bacteroides caccae*, *Bacteroides cellulosilyticus* and *Bacteroides thetaiotamicron* produced hemolytic phenotypes. The widespread nature of this phenotype suggests that there may be important conserved functions for the components of this interaction across the genus *Bacteroides*.

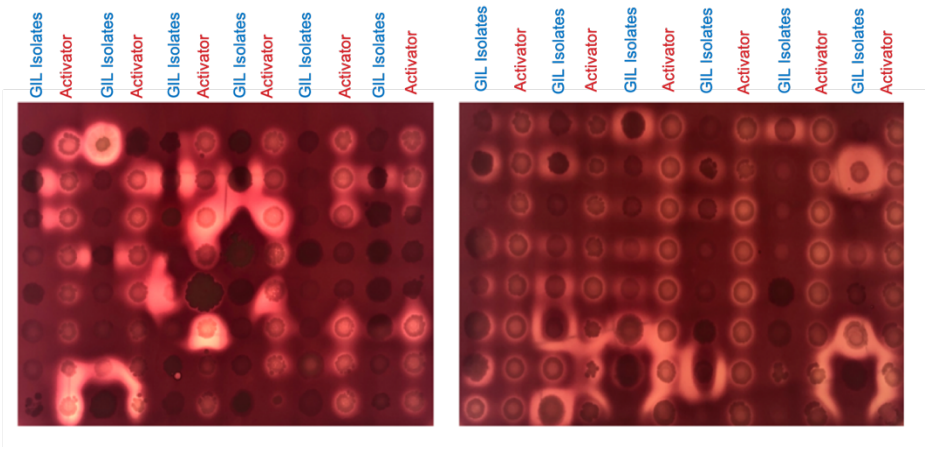

Figure 5. Synergistic hemolysis interactions observed in 94 *Bacteroides* isolates with *Bacteroides vulgatus* GIL820**.**

Bacteroides isolates were grown in BHI broth for 48 hours On one set of plates, every other column was our activator strain, while on another, every other column was our activatee strain. Isolates were stamped onto Columbia Blood Agar supplemented with 5% Sheep Blood and then incubated in an anaerobic chamber for 36 hours.

To investigate how many other strains had the ability to activate hemolysis, I used *B. ovatus* GIL819 (activatee) which became hemolytic when plated next to *B. vulgatus* GIL820 (activator). We screened all the *Bacteroides* isolates next to *B. ovatus* GIL819 on blood agar plates and looked for induction of hemolysis in activatee *B. ovatus* GIL819. 17 strains of *Bacteroides* were able to activate this synergistic hemolysis phenotype. These interactions are summarized in a schematic (Figure 6).

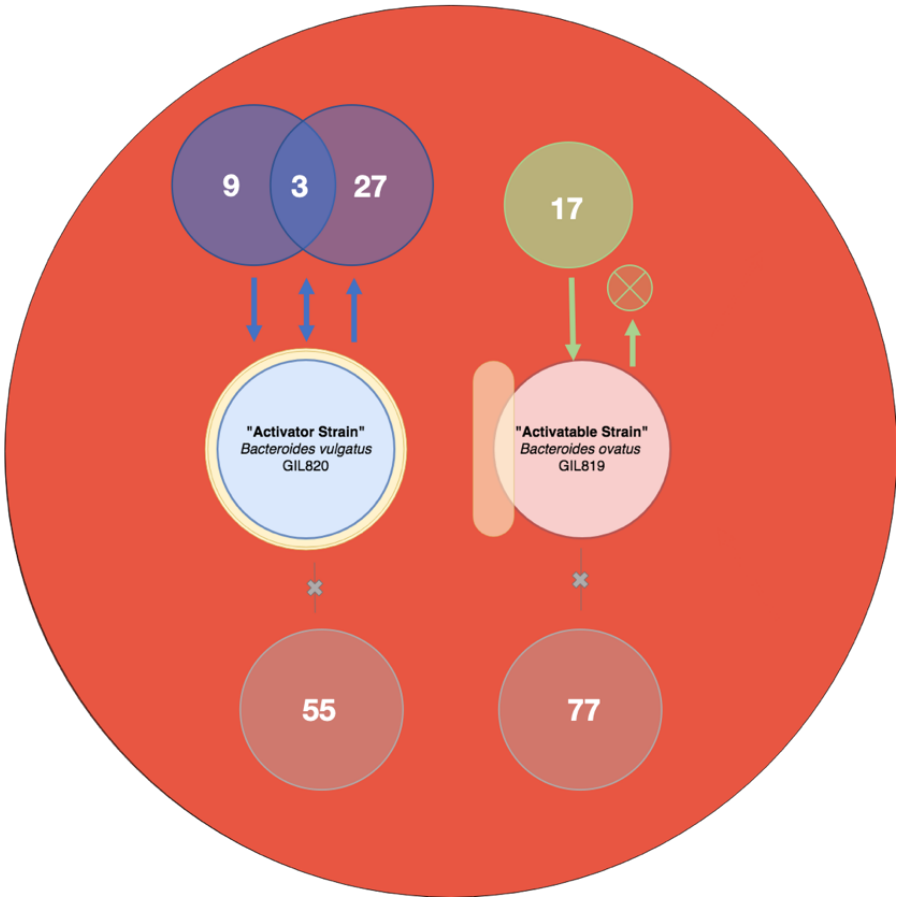

Figure 6. A schematic of the synergistic hemolysis interactions observed in 94 *Bacteroides* isolates.

Circles represent the number of *Bacteroides* isolates in each category. Arrows point to isolates where hemolysis being induced. Nine isolates induced hemolysis in *B. vulgatus* 820. *B. vulgatus* 820 induced hemolysis in 27 isolates, three isolates produced a bidirectional hemolytic phenotype. Fifty-five isolates did not interact with *B. ovatus* GIL820. Seventeen isolates activated hemolysis in *B. ovatus* 819. *B. ovatus* 819 did not induce hemolysis in any isolates. Seventy-seven isolates did not interact with *B. ovatus* 819.

3.3.1 Timelapse Imaging Shows Spatial and Temporal Dynamics of Interaction

To better understand the spatial and temporal variability behind this pattern of hemolysis, we created a custom camera set-up to create timelapse videos of hemolysis interactions. We were interested in characterizing the diffusion pattern and the size of the zone of hemolysis over time between *B. ovatus* GIL819 and *B. vulgatus* GIL820. Timelapse images show the zone of hemolysis beginning between the two colonies (Figure 7) and proceeding in a retrograde fashion towards the 'activated' colony. Our timelapse analysis shows that the proximity between the colonies and the incubation time of the agar plate are factors that influence the appearance of the zone of hemolysis.

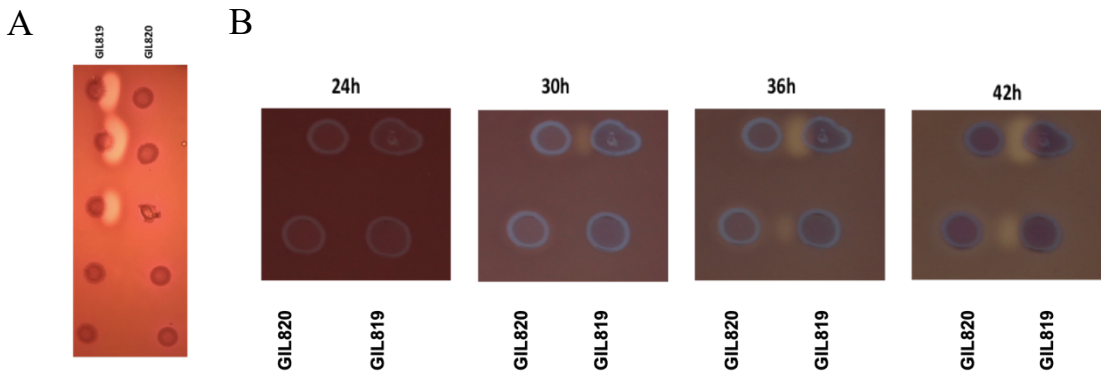

Figure 7. Distance and incubation time affect the synergistic hemolysis phenotype between *B. vulgatus* GIL819 and *B. ovatus* GIL820.

(A)*B. ovatus* GIL819 (left) and *B. vulgatus* GIL820 (right) plated at increasing distances from one another. (B) *B. ovatus* GIL819 (right) and *B.vulgatus* GIL820 (left) shown in duplicate in a series of timelapse images at 24 hours, 30 hours, 36 hours and 42 hours

3.3.2 Effect of Changing Media Conditions on Hemolysis

I wanted to understand the effects of pH, blood source and source of the "activating" component in this interaction. To investigate the effect of pH on this interaction, I grew *Bacteroides* GIL819 (activatee) on CBA plates made at different pH points including one that that might be more physiologically relevant. The pH of the gut varies from pH 5.5 to pH 7 (Fallingborg 1999). CBA plates at pH 5.5, 6.5, 7.5 were not capable of activating hemolysis in GIL819 and did not enhance the interaction between GIL819 and GIL820. Colonies grew poorly on CBA at pH 5.5 and pH 6.5.

I wanted to determine if the hemolysis was specific to Sheep's blood. I repeated this screen on CBA plates supplemented with human blood sourced from a single donor instead of Sheep's blood and observed a similar pattern of phenotypes, although the zones of hemolysis appeared to be smaller on the human blood plates. *Bacteroides* isolates are able to lyse both human and sheep blood (Figure 8).

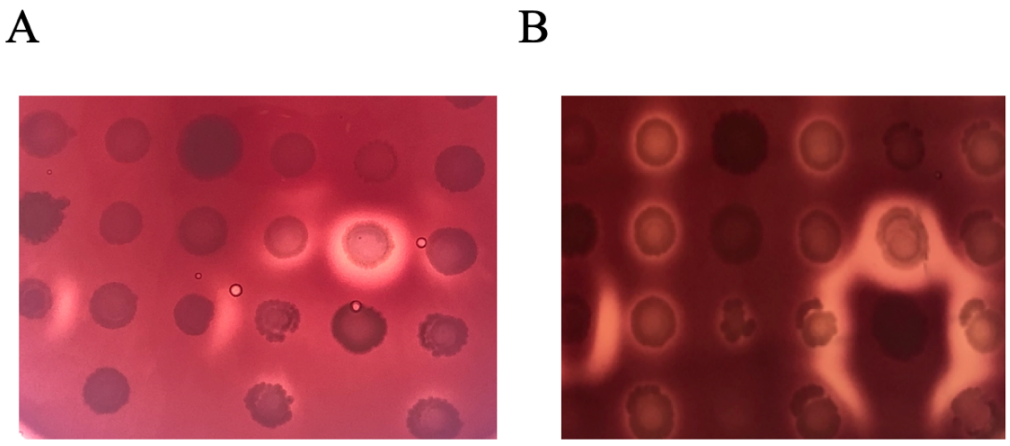

Figure 8. Blood source affects synergistic hemolysis phenotypes.
The same *Bacteroides* isolates grown on human blood (A) and Sheep's blood (B).

3.3.3 Bacterial Supernatants Are Not Sufficient To Cause Hemolysis

To understand if the activating component from GIL820 is present in the BHI broth when grown in suspension, I generated supernatants from activator GIL820. Supernatants from GIL820 were added to the media before growing GIL819. The supernatant from the activator species was not sufficient to induce hemolysis in the activatee GIL819. When bacterial supernatants of activator GIL820 were added to cultures of GIL819 we did not see activation of the hemolysis phenotype. Additional combinations of activatee GIL819 and activator GIL820 (Table 1) were also insufficient to see the phenotype. I hypothesized that there was an interaction between components produced by in both supernatants that were interacting to produce hemolysis. However, supernatants of co-cultured bacteria as well as mixed supernatants of GIL819 and GIL820

did not produce hemolysis. Concentrating the supernatants six-fold using a centrifugal filter before adding them to the growing culture also was not sufficient to activate hemolysis in GIL819. It is important to note that the BHI broth *Bacteroides* is cultured in is a "rich" media and contains glucose while CBA does not. A broth media similar to CBA, CBB was also used to account for some of the differences in media composition. However, CBB also contains glucose. The presence of glucose in the liquid media could be affecting bacterial gene expression and the synergistic hemolysis phenotype in the supernatant experiments.

Chapter 4: Exploring The Genetic Basis Of Synergistic Hemolysis In *Bacteroides vulgatus*

The novel interaction characterized in the previous section revealed a phenotype unlike any other synergistic interaction. The phenotype was complex and involved interactions between multiple isolates of *Bacteroides,* spanning six species. It was difficult to pin down the physical components behind this interaction. The supernatant experiments (Section 3.3.3) in particular suggested that we should take a different approach to characterizing the mechanism. I chose to focus on the specific genes that were responsible in the hopes that they would provide new insight into the dynamics of this interaction. To accomplish this, I took both a comparative genomics and mutagenesis approach.

4.1 COMPARATIVE GENOMICS OF BACTEROIDES ISOLATES

Comparative genomics is a field of biological research which compares the genomic features of different organisms (de Crécy-Lagard and Hanson 2013). The genomic features may include the DNA sequence, genes, gene order and regulatory sequences. Common features of two organisms will often be encoded within the DNA that is conserved between the species and may provide insight into the evolution of the organism. DNA sequences encoding the proteins and RNAs responsible for conserved functions from the last common

ancestor should be preserved in genome sequences (Hardison 2003). ROARY was used to look at the pangenome, to identify genes shared by isolates of interest.

4.1.1 Assembly and Annotation

Quality metrics of the assemblies obtained by QUAST are summarized in Table S2. Annotation is an important part of genomics for downstream analysis as the biological interpretation of results depends on the annotation of samples. Annotation predicts the presence of genes and classifies them. Software tools including RAST (Aziz et al.2008) and Prokka (Seemann 2014) are used to predict open reading frames and annotate genes. The output from Prokka predicted 4296 CDS regions were predicted in the genome of GIL820. In GIL819, 5159 CDS regions are present. The output files from the annotation steps were used for subsequent analysis.

4.1.2 Pangenome Matrix reveals a large accessory genome

To better understand the genes that could be contributing to the differences in these phenotypes we looked the pangenome of the *Bacteroides vulgatus* isolates. I focused on the species *Bacteroides vulgatus* because a large proportion of isolates were either activators or activatees. By focusing on a single species we could potentially look across smaller phylogenetic space for differences, even SNPs, that differed between the isolates. For pangenome

analysis, we included sequencing information from nine B. *vulgatus* isolates we sequenced ourselves as well as six assembled isolates from the GC Collection (Surette Lab). We confirmed the phenotypes of these isolates prior to including them in the pangenome analysis (Table 3)

Table 3. Summary of Isolates Used in Comparative Genomics

Isolate	**Phenotype**
B. vulgatus GIL97	Activator/ Activatee
B. vulgatus GIL102	Activatee
B. vulgatus GIl504	Activatee
B. vulgatus GIL552	Activatee
B. vulgatus GIL629	Activatee
B. vulgatus GIL726	Activator
B. vulgatus GIL849	Activator/Activatee
B. vulgatus GIL914	Activator
***B. vulgatus* GIL820**	**Activator**
B. vulgatus GC186	Activatee
B. vulgatus GC195_B11	Activatee
B. vulgatus GC196	Activatee
B. vulgatus GC197	Activatee
B. vulgatus GC397	Activatee

The pangenome is the entire gene set of all strains of a species. It includes genes present in all strains, or the core genome, and genes present only in some

strains of a species, known as the accessory genome. *Bacteroides vulgatus* has a large accessory genome suggesting that this species has a lot inter-strain variability (Figure 9). Combined with the large genome size of this species, narrowing down key targets responsible for the synergistic hemolysis phenotype would be complex. The clustering of the isolates in the tree suggests that isolates with the activator phenotype may be more phylogenetically related.

There are a large number of genes that are variable even within the same species (Figure 9). There are two main categories generated by ROARY. One category represents the genes of the core genome where a gene is categorized and shared by 99% or more of the isolates, which in the case of this limited set of genomes will account for all the genes shared by the fifteen isolates. There are 2490 genes that are shared in all 15 isolates (Figure S4). The remaining genes are classified as the accessory genes, and differ between isolates. There are 6507 accessory genes. To further refine these categories, more *Bacteroides vulgatus* genomes would have to be included in the pangenome analysis. SNP analysis on such a large and variable genome would have yielded very noisy data and was therefore not pursued. SNP analysis is usually performed on much larger datasets when sequencing data is available for many strains.

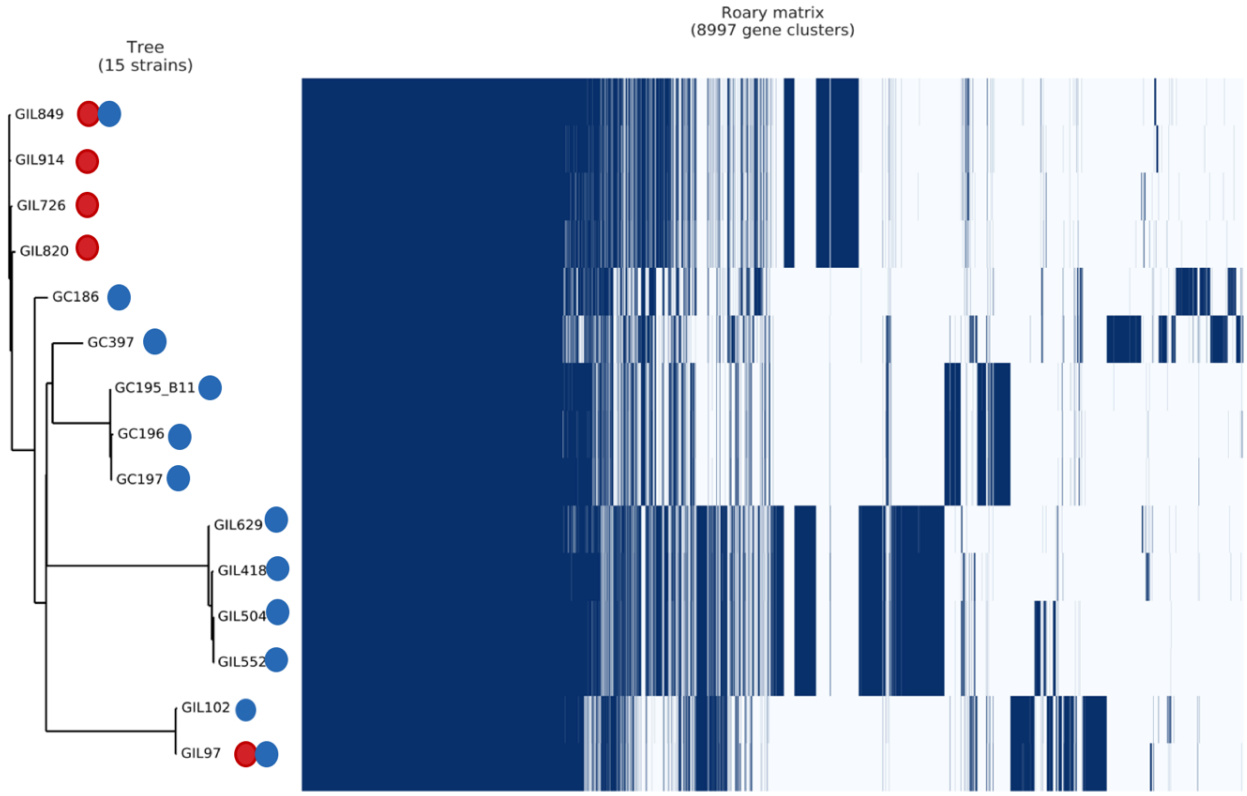

Figure 9. A pangenome matrix generated by ROARY of 15 *Bacteroides vulgatus* isolates.
The pangenome matrix show the presence (blue) and absence (white) of genes in the core and accessory genome. Red circles identify isolates that display an activator phenotype while blue circles identify isolates displaying an activatee phenotype. The tree is maximum-likelihood phylogeny representing the isolates are clustered based on an alignment of the core genes.

4.1.3 Pan-GWAS of *Bacteroides vulgatus*

With the pangenome information generated by ROARY, I next wanted to see if there were any genes that were only present in our "activator" isolates. The genome wide association by SCOARY revealed only one gene that was present in all the activators and absent in the

activatees, N-acetylmannosaminyltransferase. However, there are 17 genes that are present in all of the activators but only one of the activatees. Finally, there are two genes that are present in only four of the activators.

Table 4. Summary of Top 20 SCOARY Results

Gene	Annotation	Number_pos_present_in	Number_neg_present_in
group_1444	N-acetylmannosaminyltransferase	5	0
group_1294	hypothetical protein	5	1
group_6690	Glycosyl transferase, group 1	5	1
group_505	hypothetical protein	5	1
group_1274	Rhamnogalacturonides degradation protein RhiN	5	1
group_1338	plasmid recombination enzyme	5	1
group_1334	putative protein involved in transposition	5	1
group_1280	FIG00897206: hypothetical protein	5	1
group_6689	putative glycosyltransferase protein	5	1
group_1281	hypothetical protein	5	1
group_510	Class A beta-lactamase => CfxA family, broad-spectrum	5	1
group_6543	hypothetical protein	5	1
group_6545	excisionase	5	1
group_6544	hypothetical protein	5	1
group_6547	Transposase	5	1
group_6546	Integrase	5	1
group_6686	Membrane protein involved in the export of O-antigen, teichoic acid lipoteichoic acids	5	1
group_6542	Mobilization protein	5	1
group_8389	beta-galactosidase	4	0
group_8388	Outer membrane TonB-dependent transporter, utilization system for glycans and polysaccharides(PUL), SusC family	4	0

4.2 SCREENING FOR INTERACTIONS BETWEEN GIL820 MUTANT LIBRARY AND WILDTYPE GIL819

The complexity of the *Bacteroides vulgatus* accessory genome, combined with its large genome size and inter-strain variability led us to the mutagenesis approach for understanding the genetic basis for this phenotype.

Transposon mutagenesis was chosen to produce randomized gene mutations in the *Bacteroides* isolates of interest. Tranposable elements (TEs), or transposons are mobile genetic elements that can relocate from one genomic location to another (Hayes 2003). Transposon mutagenesis was chosen because this method produces a library of mutant bacteria which could be screened for alterations in the hemolysis phenotype. The *Himar1* mariner transposon, which was used for this approach, yields random distribution of insertions in the genome with a requirement for a TA dinucleotide target (Munoz-Lopez and Garcia-Perez 2010).

Although mutagenesis of GIL819 was attempted, the conjugation efficiency was very low with only about 15 mutants present in the undiluted mutant library sample. Therefore, the mutant library screen was only performed with GIL820.

The comparative genomics approach provided us with a lot of potential targets but no clues to a mechanism behind this phenotype. Through the mutagenesis protocol described in section 2.5., I created a mutant library of 9600 *B. vulgatus* GIL820 isolates which I then screened for an altered synergistic hemolysis phenotype. In particular, I was looking for mutants that were constitutively hemolytic on their own, or were unable to activate hemolysis in *B. ovatus* GIL819 (Figure 10). Most of the mutants did not display an altered hemolysis phenotype, and were consistently able to activate GIL819. Some

mutants did display altered growth phenotypes including slower growth, and differences in morphology from the wildtype.

Screening 2880 mutants, I was able to find 19 that displayed either constitutive hemolysis or loss of activation. Fifteen were loss of activation mutants and four were constitutive hemolysis mutants. These mutants were selected for further investigation to determine their transposon integration site. A nested PCR using semirandom primers was used to amplify the region with the mutation. After identifying the insertion site, I used Blastx to determine which gene the insertion was in. I was able to map 16/19 isolates back to GIL820 (Table 3). One of these hits, 19B8E, was not able to be identified by either Blastx or mapping back to the GIL820 genome. Hit 1B4E and 10B2G were also unable to be mapped back to the GIL820 genome but were identified as clostripain after using Blastx. Some of these mutants had insertions in the same area of the *Bacteroides vulgatus* genome. Next, I needed to consider which genes may be affected by the mutation, by looking at the operon as a whole.

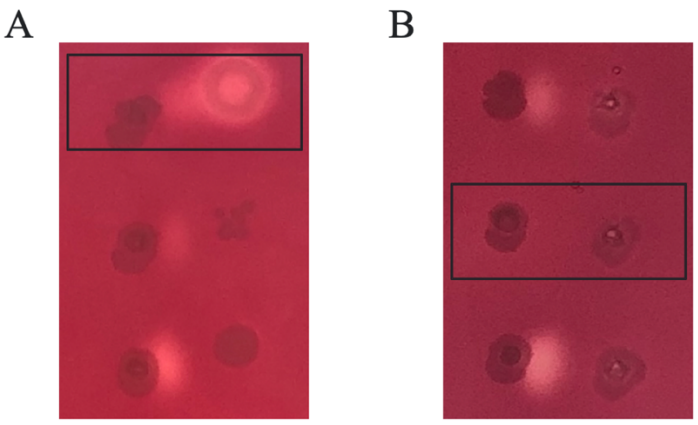

Figure 10. Hemolysis screen of mutants revealed two phenotypes of interest. (A) Constitutive Hemolysis Phenotype (B) Loss of Activation Phenotype

Table 5: Summary of Hits from Mutant Library Screen

Mutant	Phenotype	Blastx (nr) Results
21A11C	Constitutive Hemolysis	DNA/RNA non-specific endonuclease
9B12C	Constitutive Hemolysis	DUF4837 family protein
13B2D	Constitutive Hemolysis	peptide chain release factor 3*
8B10A	Constitutive Hemolysis	peptide chain release factor 3*
13B8G	Loss of Activation	[FeFe] hydrogenase H-cluster maturation GTPase HydF
16A10B	Loss of Activation	2,3-bisphosphoglycerate-independent phosphoglycerate mutase
9B2E	Loss of Activation	2,3-bisphosphoglycerate-independent phosphoglycerate mutase
6A3C	Loss of Activation	acyl-[acyl-carrier-protein]-UDP-N-acetylglucosamine O-acyltransferase
6A11F	Loss of Activation	beta-ketoacyl-ACP synthase III
9B2C	Loss of Activation	clostripain
1B4E	Loss of Activation	clostripain
13A9F	Loss of Activation	clostripain
10B2G	Loss of Activation	clostripain
21B10B	Loss of Activation	IS982 family transposase
22A5C	Loss of Activation	lipoprotein
19B8E	Loss of Activation	No Match
21B8G	Loss of Activation	putative transmembrane protein
12B4D	Loss of Activation	thioredoxin-disulfide reductase
21A9A	Loss of Activation	type II 3-dehydroquinate dehydratase

*Mutants at identical sites.

4.2.1 Analysis of Hits

Area with multiple hits close together in the genome were selected for further analysis and characterization. The goal of this approach was to

find clusters of genes that may be involved in this phenotype. Interestingly, there are no insertion hits in the genes that were hits from the SCOARY pan-GWAS analysis. The discrepancy between these two methods may be a function of the complexity of these genomes. However, the pangenome matrix suggests that four of the activators are quite phylogenetically close. Additionally, whether a strain is an activator or activatee may depend on other differences such as SNPs. The activator isolates may not all function in the exact same way and it is possible that different genes with similar functions may be involved. However, the output from SCOARY includes hypothetical proteins that may need to be further characterized. Continued screening of the mutant library could reveal additional genes of interest.

4.2.2 Analysis of Hit 6A3C

Hit 6A3C resulted in a loss of activation and was in *lpxa*, an acyl-[acyl-carrier-protein]-UDP-N-acetylglucosamine O-acyltransferase. When mapped to *Bacteroides vulgatus* complete genome ATCC 8482 for more completeness, the mutation appears to be in a cluster of genes involved in transport. There are downstream hypothetical proteins that could be affected by the mutation here. Phyre2 and SWISS-MODEL analysis reveal that the one of the downstream proteins is a putative lipase with one low coverage match (Coverage 0.36) to a hemolysin (Table S6). Additionally, the fact that this hit in an ACP-acyltransferase is interesting in the context of hemolysins.

Hemolysins secreted by pathogenic *E. coli* bind to cell membranes and lyse cells by pore-formation. Hemolysins in *E.coli* are synthesized as nontoxic prohemolysin, which is activated intracellularly by a mechanism dependent on HlyC (Issartel et al. 1991). HlyC is an internal protein acyltransferase that activates the hemolysin (Trent et al. 1999). HlyC alone cannot activate this prohemolysin but requires a cytosolic activating factor which is identical to the acyl carrier protein. Issartel *et al.* showed that activation of the toxin is achieved by the transfer of a fatty acyl group from acyl carrier protein to the prohemolysin and that only acyl carrier protein, not acyl-CoA, can perform this acylation (Issartel et al. 1991). Although the process of hemolysis in *E. coli* is most likely very different from hemolysis in *Bacteroides*, the activation of the hemolysin by an acyltransferase may be relevant.

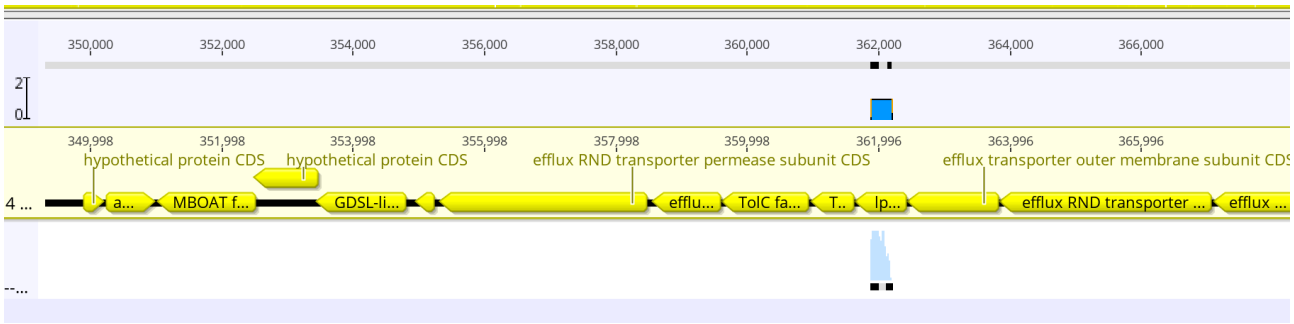

Figure 11. Genomic context for Hit 6A3C
Hit 6A3C mapped to reference genome *Bacteroides vulgatus* ATCC 8482.The insertion is located in the *lpxA* gene coding for acyl-[acyl-carrier-protein]-UDP-N-acetylglucosamine O-acyltransferase. The area in blue represents the coverage of the hit mapped back to the reference. *Geneious version 2019.2 created by Biomatters. Available from https://www.geneious.com*

4.2.3 Analysis of Hit 8B10A and 13B2D

Mutants 8B10A and 13B2D were both hits at the same site. These mutants were constitutively hemolytic and were both in a gene coding for peptide chain release factor 3. Release factors are proteins that allow for the termination of translation by recognizing the stop codon in an mRNA sequence (Zaher and Green 2011). They interact with the ribosome and stimulate the activities of other release factors (Petropoulos et al. 2014). A mutation that affects this process could affect the component responsible for hemolysis.

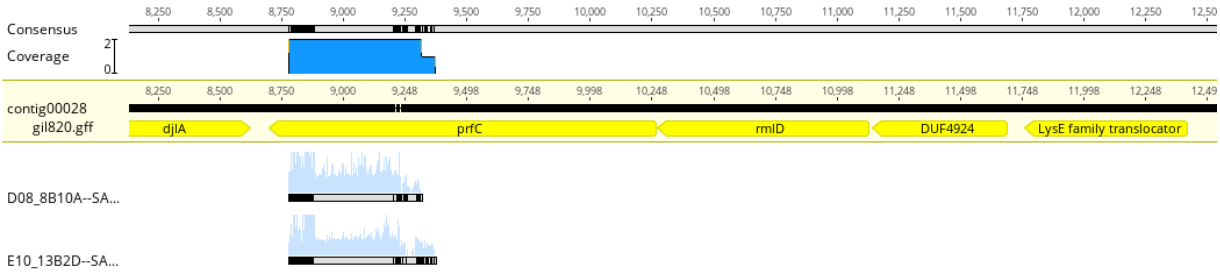

Figure 12. Genomic context for Hits 8B10A and 13B2D
Hit 8B10A and 13B2D mapped to reference genome GIL820. Both hits map to a gene coding for peptide chain release factor (*prfc*). The area in blue represents the coverage of the hits mapped back to the reference. *Geneious version 2019.2 created by Biomatters. Available from https://www.geneious.com*

4.2.4 Analysis of Hit 9B2C

Clostripain is a cysteine-activated protease found in *Clostridium* species. This hit was found in four different mutants but only one of the results was

successfully mapped back to a *Bacteroides vulgatus* genome. All four hits however, were loss of activation mutants. Clostripain is a protease that is specific to the carboxyl peptide bond of arginine. A similar protein fragipain, is found in *B. fragilis*. This protease is required for activation of BFT through removal of an auto-inhibitory prodomain. Structural analysis of fragipain revealed that similar proteins of C11 family cysteine proteases are conserved in multiple pathogenic *Bacteroides* and *Clostridium* species (Choi et al. 2016). It was found that removing fragipain from enterotoxigenic *B. fragilis* attenuated its ability to cause sepsis. In the context of hemolysis, it is possible that a protease is responsible for activating a toxin that causes hemolysis in a manner similar to the activation of BFT. Further characterization of the role of clostripain in *Bacteroides vulgatus* is required to understand how clostripain could be involved in this synergistic hemolysis phenotype.

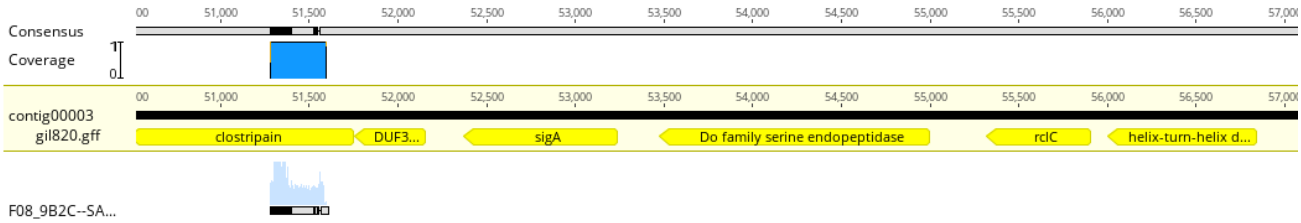

Figure 13. Genomic context for Hit 9B2C
Hit 9B2C mapped to reference genome GIL820. Hit mapped to gene coding for clostripain. The area in blue represents the coverage of the hit mapped back to the reference. *Geneious version 2019.2 created by Biomatters. Available from https://www.geneious.com*

4.2.5 Analysis of Hit 9B2E and 16A10B

Hits 9B2E and 16A10B were both mutants that resulted in loss of activation. These mutants had insertions in the same gene, which codes for 2,3-bisphosphoglycerate-independent phosphoglycerate mutase, an enzyme involved in glycolysis. Downstream of this gene is a phosphate-selective porin which could be affected by the mutation. These outer membrane proteins are anion-specific porins and have a binding site which has a high affinity for phosphate. The effects of a mutation in this region of the genome on the hemolysis phenotype is unclear.

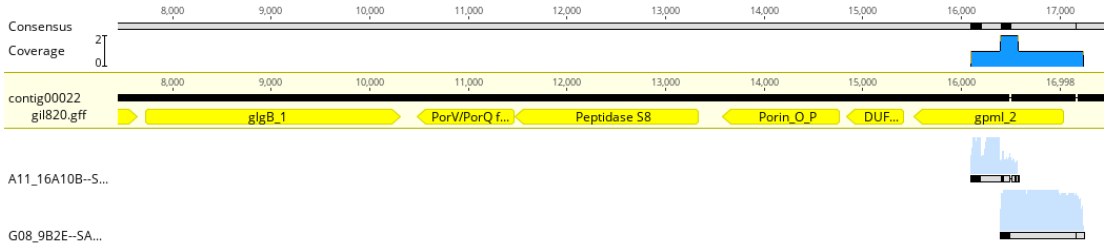

Figure 14. Genomic context for Hits 9B2E and 16A10B
Hit 9B2E and 16A10B mapped to reference genome. Both hits map to 2,3-bisphosphoglycerate-independent phosphoglycerate mutase. The area in blue represents the coverage of the hits mapped back to the reference. Areas where the blue bar is larger represents where the hits overlapped. *Geneious version 2019.2 created by Biomatters. Available from https://www.geneious.com*

4.2.6 Analysis of Hits 9B12C and 21B8G

Hits 9B12C and 12B8G both appear to be on the same operon and both have the loss of activation phenotype. The key features of this operon

include a CDP-alcohol phosphatidyltransferase family protein, a flippase-like domain-containing protein, HAD-family hydrolase and a sugar transporter.

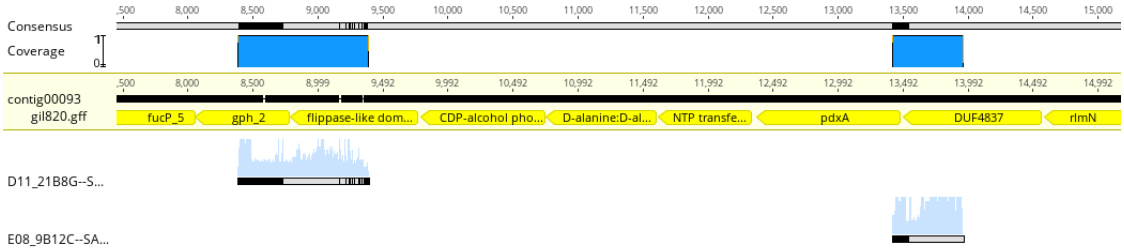

Figure 15. Genomic context for Hits 9B12C and 21B8G.
Hit 9B12C and 21B8G mapped to reference genome. Hit 21B8G is inserted in a flippase-like domain and (gph_2) Phosphoglycolate phosphatase. Hit 9B12C is in DUF4837. The area in blue represents the coverage of the hits mapped back to the reference *Geneious version 2019.2 created by Biomatters. Available from* https://www.geneious.com

Chapter 5: Discussion

Here, we have begun to characterize a unique synergistic hemolysis phenotype in *Bacteroides*, a genus of bacteria currently classified as nonhemolytic. Though there are many interactions between bacteria, especially in the gut, we have uncovered a unique hemolytic phenotype that has not been previously observed (Seth and Taga 2014). That 30% of our isolates become hemolytic in the presence of a single activator suggests that this phenomenon may actually be quite widespread. These isolates were from six different species. These findings suggests that hemolytic activity in the gut may be underestimated. Looking at similar screens with the other activators may reveal additional activatees in the collection. The role of hemolysis in the context of the human gut raises additional questions.

These isolates may cooperating to release iron from erythrocytes in order to better proliferate during bloodstream infections. It is possible that hemolysis is a virulence factor that allows *Bacteroides* isolates to become pathogens when advantageous. As *Bacteroides* isolates are not routinely screened for, they are potentially being missed in polymicrobial infections. Though they may not be detected as pathogens themselves, certain isolates of *Bacteroides* might contribute to the virulence of more “conventional” pathogens. By lysing red blood cells, *Bacteroides* may be releasing iron which can regulate virulence

factors in pathogens (Skaar 2010). In fact, opportunistic pathogen *Bacteroides fragilis* requires both heme and free iron to support growth in extraintestinal infections (Rocha et al. 2019).

The differences in the shape of the hemolytic zone in the "CAMP" test done with the *Bacteroides* suggests that this interaction is different from synergistic hemolysis seen in the CAMP test between *Streptococcus* and *Staphylococcus* species. Imaging the plates after incubation showed a variety of different hemolytic patterns which warrants further investigation, especially the presence of multiple "activator" species that were able to activate hemolysis in other isolates, including GIL820. The timelapse analysis discussed in Chapter 3 suggests that this interaction occurs due to diffusible components. The zone of hemolysis may be a result of two components being secreted that interact to produce hemolysis. The interaction of two diffusible components would result in the formation of the zone of hemolysis in the middle of the two colonies. However, the clear directionality of this interaction suggests an interaction involving 3 components, one of which stimulates the production of a compound that interacts to produce hemolysis. The zone of hemolysis starting in the middle of the two colonies and proceeding towards the activated colony in a retrograde manner suggests that a "stimulating" factor may be present. This factor could be responsible for the production of another component by the activatee which interacts with an activator component to produce hemolysis. It is possible that more components are involved. If these molecules are costly to produce but are

intended to work together, an activator or activatee might only produce their contribution upon receipt of a signal.

Future work will need to focus on identifying the specific components of this interaction. We need to determine if these components are inducible under certain conditions or if they are constitutively expressed. Continuing to explore the physical characteristics of this interaction will provide more answers. It is important to determine the potential cytotoxic effects of this phenotype, as this interaction may not be specific to hemolysis. The only outcome of cell lysis measured in this work was the red blood cell lysis observed on the CBA plates. It is very possible that here we are using hemolysis as a marker for the activity of a cytolysin. A cytolysin could target many different types of cells, and could damage the intestinal epithelial cells with which *Bacteroides* are associated. Determining the cytotoxic properties of this interaction will provide better context for the function of this phenotype. Perhaps what we have discovered here are two bacteria that just need to be close enough for the two components to synergize, and that they interact in close proximity. This bacterial interaction could involve the intestinal epithelium and host response. Intestinal epithelial cells are in close proximity to luminal pathogenic and non-pathogenic bacteria. It is at these sites where the synergistic hemolysis phenotype may occur.

The complexity of these types of interactions make it clear that we are missing these phenotypes when studying bacterial species in isolation. Bacteria respond to stimuli in their environment, and the dynamics of these responses are

difficult to study in the lab. The dynamics of bacterial communities are influenced by interactions that occur between different species in the community. Though interactions may occur in only a subset of isolates at a time, the outcomes of these interactions may have effects on a larger scale. In the context of the gut, these microbial interactions can have a significant impact on host health.

My screen of the transposon mutant library revealed many hits that could be involved in the novel hemolysis phenotype. Hits from the mutant screen were in different parts of the *Bacteroides vulgatus* genome. Mutants with a loss of activation phenotype were sometimes in an operon with genes coding for a transmembrane or transport protein. This suggests that efflux of an intercellular component may be a part of the mechanism by which *B. vulgatus* GIL820 activates hemolysis in other isolates. Additionally, clostripain, a cysteine-protease like fragipain which is used to activate *Bacteroides fragilis* toxin, also came up in the screen multiple times, but not as duplicate insertions at the same site. This suggests that export of a protease might be part of the mechanism used in hemolysis. The constitutive hemolysis phenotype was seen less often in the screen, and two of the four hits were at the same site. The main finding in the investigation of this phenotype is that this interaction is perhaps more complex than we had initially imagined. Mutagenesis of the activatee GIL819 is necessary to further narrow down genes involved in this mechanism. By understanding which genes in the activatee *B. ovatus* GIL819 are required for

this phenotype, we can gain a better understanding of which components may be interacting.

The phenotype characterized here is an example of how microbes may be secreting chemical mediators to influence each other and shape the environment that they are in. By exploring interactions such as the one presented here, we can better understand the compounds bacteria secrete and how they are stimulated to do so. The role of these natural products secreted by bacteria may clarify how bacteria interact in different ecological niches.

Altering the availability of essential nutrients can also be a means to influence the behaviour of a microbial partner. Bacterial cross-feeding in the infant gut is a key example of metabolic interactions between different groups of bacteria sustaining communities (Seth and Taga 2014). Additionally, certain fungi are able to trigger a new phenotype in *Streptomycetes* by glucose depletion, allowing them to colonize new environments. Exploratory growth is also communicated by utilizing the volatile trimethylamine as a chemical mediator (Jones and Elliot 2017). The metabolic interactions taking place in the gut environment may be a part of the synergistic hemolysis phenotype presented here.

The discovery of this novel synergistic hemolytic phenotype in *Bacteroides* isolates warrants additional investigation of microbe-microbe interactions between *Bacteroides* isolates in the gut. The work presented here highlights the importance of studying bacteria in communities. Perhaps a lot of

what we hold to be true about bacteria, may only be correct in isolates and not communities. By understanding the interactions within these communities, we will be able to capture phenotypes that have yet to be seen.

Bibliography

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. J. Mol. Biol. **215**(3): 403–410. Doi:10.1016/S0022-2836(05)80360-2.

Andrews, S. 2015. FASTQC A quality control tool for high throughput sequence data. Babraham Inst.

Aziz, R.K., Bartels, D., Best, A.A., dejongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., mcneil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., and Zagnitko, O. 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics **9**(1): 75. Doi:10.1186/1471-2164-9-75.

Bessesen, D.H. 2007. Human gut microbes associated with obesity. Yearb. Endocrinol. **2007**: 163–165. Doi:10.1016/S0084-3741(08)70094-5.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T.G., Bertoni, M., Bordoli, L., and Schwede, T. 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. **42**(W1): W252–W258. Doi:10.1093/nar/gku340.

Bingle, L.E., Bailey, C.M., and Pallen, M.J. 2008. Type VI secretion: a beginner's guide. Curr. Opin. Microbiol. **11**(1): 3–8. Doi:10.1016/j.mib.2008.01.006.

Biomatters Limited. 2018. Geneious - Molecular Biology and NGS Analysis Tools.

BLAST. 2009. Nucleotide BLAST: Search nucleotide databases using a nucleotide query.

Bolger, A.M., Lohse, M., and Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30**(15): 2114–2120. Doi:10.1093/bioinformatics/btu170.

Browne, H.P., Forster, S.C., Anonye, B.O., Kumar, N., Neville, B.A., Stares, M.D., Goulding, D., and Lawley, T.D. 2016. Culturing of "unculturable" human microbiota reveals novel taxa and extensive sporulation. Nature. Doi:10.1038/nature17645.

Brynildsrud, O., Bohlin, J., Scheffer, L., and Eldholm, V. 2016. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. Genome Biol. **17**(1): 238. Doi:10.1186/s13059-016-1108-8.

Bunesova, V., Lacroix, C., and Schwab, C. 2018. Mucin cross-feeding of infant *Bifidobacteria* and *Eubacterium hallii.* Microb. Ecol. **75**(1): 228–238. Doi:10.1007/s00248-017-1037-4.

Buxton, R. 2005. Blood agar plates and hemolysis protocols. Am. Soc. Microbiol.

Cavalieri, S.J., and Snyder, I.S. 1982. Effect of *Escherichia coli* alpha-hemolysin on human peripheral leukocyte function in vitro. Infect. Immun. **37**(3): 966–974. Doi:10.1128/IAI.37.3.966-974.1982.

Chatzidaki-Livanis, M., Coyne, M.J., Roelofs, K.G., Gentyala, R.R., Caldwell, J.M., and Comstock, L.E. 2017. Gut symbiont *Bacteroides fragilis* secretes a eukaryotic-like ubiquitin protein that mediates intraspecies antagonism. Mbio **8**(6). Doi:10.1128/mbio.01902-17.

Cheng, Z., Huang, Y., Wie, W., Wang, Y., and Wang, Z. 2019. Bloodstream infection caused by *Bacteroides* caccae in a diabetic patient: a case report and review of the literature. Clin. Lab. **65**(12/2019). Doi:10.7754/Clin.Lab.2019.190534.

Cheung, S.G., Goldenthal, A.R., Uhlemann, A.-C., Mann, J.J., Miller, J.M., and Sublette, M.E. 2019. Systematic review of gut microbiota and major depression. Front. Psychiatry **10**. Doi:10.3389/fpsyt.2019.00034.

Choi, V.M., Herrou, J., Hecht, A.L., Teoh, W.P., Turner, J.R., Crosson, S., and Wardenburg, J.B. 2016. Activation of *Bacteroides fragilis* toxin by a novel bacterial protease contributes to anaerobic sepsis in mice. Nat. Med. **22**(5): 563–567. Doi:10.1038/nm.4077.

Clemente, J.C., Ursell, L.K., Parfrey, L.W., and Knight, R. 2012. The impact of the gut microbiota on human health: an integrative view. Cell **148**(6): 1258–1270. Doi:10.1016/j.cell.2012.01.035.

Coyne, M.J., and Comstock, L.E. 2019. Type VI secretion systems and the gut microbiota. Microbiol. Spectr. **7**(2). Doi:10.1128/microbiolspec.PSIB-0009-2018.

Coyne, M.J., Roelofs, K.G., and Comstock, L.E. 2016. Type VI secretion systems of human gut *Bacteroidales* segregate into three genetic architectures, two of which are contained on mobile genetic elements. BMC Genomics **17**(1): 58. Doi:10.1186/s12864-016-2377-z.

De Crécy-Lagard, V., and Hanson, A. 2013. Comparative Genomics. *In* Brenner's Encyclopedia of Genetics: Second Edition. Doi:10.1016/B978-

0-12-374984-0.00299-0.

Donohoe, D.R., Garge, N., Zhang, X., Sun, W., O'Connell, T.M., Bunger, M.K., and Bultman, S.J. 2011. The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. Cell Metab. **13**(5): 517–526. Doi:10.1016/j.cmet.2011.02.018.

Fallingborg, J. 1999. Intraluminal pH of the human gastrointestinal tract. Dan. Med. Bull. **46**(3): 183–96. Available from http://www.ncbi.nlm.nih.gov/pubmed/10421978.

Fischetti, V.A., and Ryan, P. 2015. Practical handbook of microbiology. *In* Practical Handbook of Microbiology, Third Edition. *Edited by*e. Goldman and L.H. Green. CRC Press. Doi:10.1201/b17871.

Franco, A.A. 2004. The *Bacteroides fragilis* pathogenicity island is contained in a putative novel conjugative transposon. J. Bacteriol. Doi:10.1128/JB.186.18.6077-6092.2004.

Garrett, W.S., and Onderdonk, A.B. 2014. *Bacteroides, Prevotella, Porphyromonas*, and *Fusobacterium* species (and other medically important anaerobic gram-negative bacilli). *In* Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases. Doi:10.1016/B978-1-4557-4801-3.00249-6.

Gibson, F.C., Tzianabos, A.O., and Onderdonk, A.B. 1996. The capsular polysaccharide complex of *Bacteroides fragilis i*nduces cytokine production from human and murine phagocytic cells. Infect. Immun. **64**(3): 1065–1069. Doi:10.1128/IAI.64.3.1065-1069.1996.

Goodman, A.L., mcnulty, N.P., Zhao, Y., Leip, D., Mitra, R.D., Lozupone, C.A., Knight, R., and Gordon, J.I. 2009. Identifying genetic determinants needed to establish a human gut symbiont in its habitat. Cell Host Microbe **6**(3): 279–289. Doi:10.1016/j.chom.2009.08.003.

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics **29**(8): 1072–1075. Doi:10.1093/bioinformatics/btt086.

Hansen, M. V., and Elliott, L.P. 1980. New presumptive identification test for *Clostridium perfringens:* reverse CAMP test. J. Clin. Microbiol. **12**(4): 617–619. Doi:10.1128/JCM.12.4.617-619.1980.

Hardison, R.C. 2003. Comparative genomics. Plos Biol. **1**(2): e58. Doi:10.1371/journal.pbio.0000058.

Hayes, F. 2003. Transposon-based strategies for microbial functional genomics and proteomics. Annu. Rev. Genet. **37**(1): 3–29. Doi:10.1146/annurev.genet.37.110801.142807.

Hebert, G.A., and Hancock, G.A. 1985. Synergistic hemolysis exhibited by species of staphylococci. J. Clin. Microbiol. **22**(3): 409–415.

Herrou, J., Choi, V.M., Bubeck Wardenburg, J., and Crosson, S. 2016. Activation mechanism of the *Bacteroides fragilis* cysteine peptidase, Fragipain. Biochemistry **55**(29): 4077–4084. Doi:10.1021/acs.biochem.6b00546.

Holth-Haug, R., and Olsvik, O. 1981. The CAMP test performed by using *Staphylococcus aureus* sphingomyelinase (beta-haemolysin) and *Clostridium perfringens* lecithinase (alpha-toxin). NIPH Ann. **4**(2): 43–48.

Hooper, L. V., Stappenbeck, T.S., Hong, C. V., and Gordon, J.I. 2003. Angiogenins: a new class of microbicidal proteins involved in innate immunity. Nat. Immunol. **4**(3): 269–273. Doi:10.1038/ni888.

Issartel, J.P., Koronakis, V., and Hughes, C. 1991. Activation of *Escherichia coli* prohaemolysin to the mature toxin by acyl carrier protein-dependent fatty acylation. Nature. Doi:10.1038/351759a0.

Jacobson, A., Lam, L., Rajendram, M., Tamburini, F., Honeycutt, J., Pham, T., Van Treuren, W., Pruss, K., Stabler, S.R., Lugo, K., Bouley, D.M., Vilches-Moure, J.G., Smith, M., Sonnenburg, J.L., Bhatt, A.S., Huang, K.C., and Monack, D. 2018. A gut commensal-produced metabolite mediates colonization resistance to *Salmonella* infection. Cell Host Microbe **24**(2): 296-307.e7. Doi:10.1016/j.chom.2018.07.002.

Jones, S.E., and Elliot, M.A. 2017. *Streptomyces* exploration: Competition, volatile communication and new bacterial behaviours. Trends Microbiol. **25**(7): 522–531. Doi:10.1016/j.tim.2017.02.001.

Kalfa, T.A. 2014. Hemolytic anemias. *In* Pathobiology of Human Disease. Elsevier. Pp. 1532–1543. Doi:10.1016/B978-0-12-386456-7.07907-7.

Karp, P.D., Billington, R., Caspi, R., Fulcher, C.A., Latendresse, M., Kothari, A., Keseler, I.M., Krummenacker, M., Midford, P.E., Ong, Q., Ong, W.K., Paley, S.M., and Subhraveti, P. 2019. The biocyc collection of microbial genomes and metabolic pathways. Brief. Bioinform. **20**(4): 1085–1093. Doi:10.1093/bib/bbx085.

Kastl, A.J., Terry, N.A., Wu, G.D., and Albenberg, L.G. 2020. The structure and function of the human small intestinal microbiota: current understanding and future directions. Cell. Mol. Gastroenterol. Hepatol. **9**(1): 33–45. Doi:10.1016/j.jcmgh.2019.07.006.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. 2015. The Phyre2 web portal for protein modeling, prediction and

analysis. Nat. Protoc. **10**(6): 845–858. Doi:10.1038/nprot.2015.053.

Krinos, C.M., Coyne, M.J., Weinacht, K.G., Tzianabos, A.O., Kasper, D.L., and Comstock, L.E. 2001. Extensive surface diversity of a commensal microorganism by multiple DNA inversions. Nature **414**(6863): 555–558. Doi:10.1038/35107092.

Krummenacker, M., Paley, S., Mueller, L., Yan, T., and Karp, P.D. 2005. Querying and computing with biocyc databases. Bioinformatics **21**(16): 3454–3455. Doi:10.1093/bioinformatics/bti546.

Los, F.C.O., Randis, T.M., Aroian, R. V., and Ratner, A.J. 2013. Role of pore-forming toxins in bacterial infectious diseases. Microbiol. Mol. Biol. Rev. **77**(2): 173–207. Doi:10.1128/MMBR.00052-12.

Lozupone, C.A., Stombaugh, J.I., Gordon, J.I., Jansson, J.K., and Knight, R. 2012. Diversity, stability and resilience of the human gut microbiota. Nature **489**(7415): 220–230. Doi:10.1038/nature11550.

Thompson, H. 1995. Gram positive pyogenic cocci. Available from https://www.medschool.lsuhsc.edu/Microbiology/DMIP/dmex16.htm [accessed 10 March 2020].

Manaargadoo-Catin, M., Ali-Cherif, A., Pougnas, J.-L., and Perrin, C. 2016. Hemolysis by surfactants — A review. Adv. Colloid Interface Sci. **228**: 1–16. Doi:10.1016/j.cis.2015.10.011.

Mckellar, R.C. 1994. Use of the CAMP test for identification of *Listeria monocytogenes*. Appl. Environ. Microbiol. **60**(12): 4219–4225. Doi:10.1128/AEM.60.12.4219-4225.1994.

Miller, M.B., and Bassler, B.L. 2001. Quorum sensing in bacteria. Annu. Rev. Microbiol. **55**(1): 165–199. Doi:10.1146/annurev.micro.55.1.165.

Morris, J.J., Kirkegaard, R., Szul, M.J., Johnson, Z.I., and Zinser, E.R. 2008. Facilitation of robust growth of *Prochlorococcus* colonies and dilute liquid cultures by "helper" heterotrophic bacteria. Appl. Environ. Microbiol. **74**(14): 4530–4534. Doi:10.1128/AEM.02479-07.

Munch-Petersen, E., Christie, R., Simmons, R., and Beddome, H. 1945. Further notes on a lytic phenomenon shown by group B streptococci. Aust. J. Exp. Biol. Med. Sci. **23**(3): 193–195. Doi:10.1038/icb.1945.30.

Muncie, H.L., and Campbell, J. 2009. Alpha and beta thalassemia. Am. Fam. Physician **80**(4): 339–44. Available from http://www.ncbi.nlm.nih.gov/pubmed/19678601.

Munoz-Lopez, M., and Garcia-Perez, J. 2010. DNA transposons: nature and applications in genomics. Curr. Genomics **11**(2): 115–128.

Doi:10.2174/138920210790886871.

Onderdonk, A.B., Cisneros, R.L., and Bronson, R. 1983. Enhancement of experimental ulcerative colitis by immunization with *Bacteroides vulgatus*. Infect. Immun. Doi:10.1128/iai.42.2.783-788.1983.

Page, A.J., Cummins, C.A., Hunt, M., Wong, V.K., Reuter, S., Holden, M.T.G., Fookes, M., Falush, D., Keane, J.A., and Parkhill, J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics **31**(22): 3691–3693. Doi:10.1093/bioinformatics/btv421.

Petropoulos, A.D., mcdonald, M.E., Green, R., and Zaher, H.S. 2014. Distinct roles for release factor 1 and release factor 2 in translational quality control. J. Biol. Chem. **289**(25): 17589–17596. Doi:10.1074/jbc.M114.564989.

Rios-Covian, D., Salazar, N., Gueimonde, M., and de los Reyes-Gavilan, C.G. 2017. Shaping the metabolism of intestinal *Bacteroides* population through diet to improve human health. Front. Microbiol. **8**. Doi:10.3389/fmicb.2017.00376.

Robert, R., deraignac, A., Moal, G., Ragot, S., and Grollier, G. 2008. Prognostic factors and impact of antibiotherapy in 117 cases of anaerobic bacteraemia. Eur. J. Clin. Microbiol. Infect. Dis. **27**(8): 671–678. Doi:10.1007/s10096-008-0487-5.

Robertson, K.P., Smith, C.J., Gough, A.M., and Rocha, E.R. 2006. Characterization of *Bacteroides fragilis* hemolysins and regulation and synergistic interactions of HlyA and HlyB. Infect. Immun. **74**(4): 2304–2316. Doi:10.1128/IAI.74.4.2304-2316.2006.

Rocha, E.R., Bergonia, H.A., Gerdes, S., and Jeffrey Smith, C. 2019. *Bacteroides fragilis* requires the ferrous-iron transporter Feo AB and the cobn-like proteins btus1 and btus2 for assimilation of iron released from heme. Microbiologyopen **8**(4): e00669. Doi:10.1002/mbo3.669.

Rodriguez-Castaño, G.P., Dorris, M.R., Liu, X., Bolling, B.W., Acosta-Gonzalez, A., and Rey, F.E. 2019. *Bacteroides thetaiotaomicron* starch utilization promotes quercetin degradation and butyrate production by *Eubacterium ramulus*. Front. Microbiol. **10**. Doi:10.3389/fmicb.2019.01145.

Saffarian, A., Mulet, C., Regnault, B., Amiot, A., Tran-Van-Nhieu, J., Ravel, J., Sobhani, I., Sansonetti, P.J., and Pédron, T. 2019. Crypt- and mucosa-associated core microbiotas in humans and their alteration in colon cancer patients. Mbio **10**(4). Doi:10.1128/mbio.01315-19.

Sagar, A., Klemm, C., Hartjes, L., Mauerer, S., van Zandbergen, G., and

Spellerberg, B. 2013. The β-Hemolysin and intracellular survival of *Streptococcus agalactiae* in human macrophages. Plos One **8**(4): e60160. Doi:10.1371/journal.pone.0060160.

Schink, B. 2002. Synergistic interactions in the microbial world. Antonie Van Leeuwenhoek **81**(1–4): 257–61. Doi:10.1023/a:1020579004534.

Schwede, T. 2003. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. **31**(13): 3381–3385. Doi:10.1093/nar/gkg520.

Seemann, T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics **30**(14): 2068–2069. Doi:10.1093/bioinformatics/btu153.

Seth, E.C., and Taga, M.E. 2014. Nutrient cross-feeding in the microbial world. Front. Microbiol. **5**. Doi:10.3389/fmicb.2014.00350.

Singer-Englar, T., Barlow, G., and Mathur, R. 2019. Obesity, diabetes, and the gut microbiome: an updated review. Expert Rev. Gastroenterol. Hepatol. **13**(1): 3–15. Doi:10.1080/17474124.2019.1543023.

Skaar, E.P. 2010. The battle for iron between bacterial pathogens and their vertebrate hosts. Plos Pathog. **6**(8): e1000949. Doi:10.1371/journal.ppat.1000949.

Smith, D.C., Foltz, V.D., and Lord, T.H. 1964. Demonstration of induced synergistic hemolysis by "nonhemolytic" *Staphylococcus* species. J. Bacteriol. **87**(1): 188–195. Doi:10.1128/JB.87.1.188-195.1964.

Stappenbeck, T.S., Hooper, L. V, and Gordon, J.I. 2002. Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells. Proc. Natl. Acad. Sci. U. S. A. **99**(24): 15451–5. Doi:10.1073/pnas.202604299.

Stone H.H., Kolb, L.D., and Geheber, C.E. 1975. Incidence and significance of intraperitoneal anaerobic bacteria. Ann. Surg. **181**(5): 705–715. Doi:10.1097/00000658-197505000-00027.

Swidsinski, A., Weber, J., Loening-Baucke, V., Hale, L.P., and Lochs, H. 2005. Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. J. Clin. Microbiol. Doi:10.1128/JCM.43.7.3380-3389.2005.

Tamboli, C.P. 2004. Dysbiosis in inflammatory bowel disease. Gut **53**(1): 1–4. Doi:10.1136/gut.53.1.1.

Teravest, M.A., He, Z., Rosenbaum, M.A., Martens, E.C., Cotta, M.A., Gordon, J.I., and Angenent, L.T. 2014. Regulated expression of polysaccharide utilization and capsular biosynthesis loci in biofilm and

planktonic *Bacteroides thetaiotaomicron* during growth in chemostats. Biotechnol. Bioeng. **111**(1): 165–173. Doi:10.1002/bit.24994.

Titball, R.W., Naylor, C.E., and Basak, A.K. 1999. The *Clostridium perfringen s*α-toxin. Anaerobe **5**(2): 51–64. Doi:10.1006/anae.1999.0191.

Trent, M.S., Worsham, L.M.S., and Ernst-Fonberg, M. Lou. 1999. HlyC, the internal protein acyltransferase that activates hemolysin toxin: role of conserved histidine, serine, and cysteine residues in enzymatic activity as probed by chemical modification and site-directed mutagenesis Biochemistry **38**(11): 3433–3439. Doi:10.1021/bi982491u.

Uchiyama, K., Naito, Y., and Takagi, T. 2019. Intestinal microbiome as a novel therapeutic target for local and systemic inflammation. Pharmacol. Ther. **199**: 164–172. Doi:10.1016/j.pharmthera.2019.03.006.

Vartoukian, S.R., Palmer, R.M., and Wade, W.G. 2010. Strategies for culture of 'unculturable' bacteria. FEMS Microbiol. Lett.: Doi:10.1111/j.1574-6968.2010.02000.x.

Veeranagouda, Y., Husain, F., and Wexler, H.M. 2012. Transposon mutagenesis of the anaerobic commensal, *Bacteroides fragilis,* using the EZ::TN5 transposome. FEMS Microbiol. Lett. **333**(2): 94–100. Doi:10.1111/j.1574-6968.2012.02602.x.

Veeranagouda, Y., Husain, F., and Wexler, H.M. 2019. Transposon mutagenesis of *Bacteroides fragili*s. Pp. 105–116. Doi:10.1007/978-1-4939-9570-7_10.

De Vos, P., Garrity, G.M., Jones, D., Krieg, N.R., Ludwig, W., Rainey, F.A., Karl-Heinz, S., and Whitman, W.B. 2009. Bergey's Manual of Systematic Bacteriology - Vol 3: The Firmicutes. *In* Springer-Verlag New York Inc. Doi:10.1007/b92997.

Welch, R.A. 1991. Pore-forming cytolysins of Gram-negative bacteria. Mol. Microbiol. **5**(3): 521–528. Doi:10.1111/j.1365-2958.1991.tb00723.x.

Wells, C.L., Maddaus, M.A., Jechorek, R.P., and Simmons, R.L. 1988. Role of intestinal anaerobic bacteria in colonization resistance. Eur. J. Clin. Microbiol. Infect. Dis. **7**(1): 107–113. Doi:10.1007/BF01962194.

Wexler, A.G., and Goodman, A.L. 2017. An insider's perspective: *Bacteroides* as a window into the microbiome. Nat. Microbiol. **2**(5): 17026. Doi:10.1038/nmicrobiol.2017.26.

Wexler, H.M. 2007. *Bacteroides*: The good, the bad, and the nitty-gritty. Clin. Microbiol. Rev. **20**(4): 593–621. Doi:10.1128/CMR.00008-07.

Wick, E.C., and Sears, C.L. 2010. *Bacteroides* spp. And diarrhea. Curr. Opin.

Infect. Dis. **23**(5): 470–474. Doi:10.1097/QCO.0b013e32833da1eb.

Wick, R.R., Judd, L.M., Gorrie, C.L., and Holt, K.E. 2017. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLOS Comput. Biol. **13**(6): e1005595. Doi:10.1371/journal.pcbi.1005595.

Wilkinson, H.W. 1977. CAMP-disk test for presumptive identification of group B streptococci. J. Clin. Microbiol. **6**(1): 42–5. Available from http://www.ncbi.nlm.nih.gov/pubmed/328534.

Wu, G.D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.-Y., Keilbaugh, S.A., Bewtra, M., Knights, D., Walters, W.A., Knight, R., Sinha, R., Gilroy, E., Gupta, K., Baldassano, R., Nessel, L., Li, H., Bushman, F.D., and Lewis, J.D. 2011. Linking long-term dietary patterns with gut microbial enterotypes. Science . **334**(6052): 105–108. Doi:10.1126/science.1208344.

Yoon, Y.M., Lee, J.Y., Yoo, D., Sim, Y.S., Kim, Y.J., Oh, Y.K., Kang, J.S., Kim, S., Kim, J.S., and Kim, J.M. 2010. *Bacteroides fragilis* enterotoxin induces human β-defensin-2 expression in intestinal epithelial cells via a mitogen-activated protein kinase/iκb kinase/NF-κb-dependent pathway. Infect. Immun. Doi:10.1128/IAI.00118-10.

Yoshida, N., Emoto, T., Yamashita, T., Watanabe, H., Hayashi, T., Tabata, T., Hoshi, N., Hatano, N., Ozawa, G., Sasaki, N., Mizoguchi, T., Amin, H.Z., Hirota, Y., Ogawa, W., Yamada, T., and Hirata, K. 2018. *Bacteroides vulgatus* and *Bacteroides dorei* reduce gut microbial lipopolysaccharide production and inhibit atherosclerosis. Circulation **138**(22): 2486–2498. Doi:10.1161/CIRCULATIONAHA.118.033714.

Zaher, H.S., and Green, R. 2011. A primary role for release factor 3 in quality control during translation elongation in *Escherichia coli.* Cell **147**(2): 396–408. Doi:10.1016/j.cell.2011.08.045.

Zakharzhevskaya, N.B., Tsvetkov, V.B., Vanyushkina, A.A., Varizhuk, A.M., Rakitina, D. V., Podgorsky, V. V., Vishnyakov, I.E., Kharlampieva, D.D., Manuvera, V.A., Lisitsyn, F. V., Gushina, E.A., Lazarev, V.N., and Govorun, V.M. 2017. Interaction of *Bacteroides fragilis* toxin with outer membrane vesicles reveals new mechanism of its secretion and delivery. Front. Cell. Infect. Microbiol. **7**. Doi:10.3389/fcimb.2017.00002.

Zheng, J., Gänzle, M.G., Lin, X.B., Ruan, L., and Sun, M. 2015. Diversity and dynamics of bacteriocins from human microbiome. Environ. Microbiol. **17**(6): 2133–2143. Doi:10.1111/1462-2920.12662.

Zhou, Y., and Zhi, F. 2016. Lower level of *Bacteroides i*n the gut microbiota is associated with inflammatory bowel disease: A Meta-Analysis.

Doi:10.1155/2016/5828959.

Zitomersky, N.L., Coyne, M.J., and Comstock, L.E. 2011. Longitudinal analysis of the prevalence, maintenance, and IgA response to species of the order *Bacteroidales* in the human gut. Infect. Immun. **79**(5): 2012–2020. doi:10.1128/IAI.01348-10.

Appendices

Table S1: Bacterial Strains

Strain List	Properties	Source
Bacteroides xylanisolvens GIL41		Surette Lab, GIL
Bacteroides uniformis GIL45		Surette Lab, GIL
Bacteroides intestinalis GIL81	Activatee	Surette Lab, GIL
Bacteroides fragilis GIL83		Surette Lab, GIL
Bacteroides uniformis GIL89		Surette Lab, GIL
Bacteroides fragilis GIL92		Surette Lab, GIL
Bacteroides fragilis GIL93		Surette Lab, GIL
Bacteroides vulgatus GIL95	Activatee	Surette Lab, GIL
Bacteroides vulgatus GIL97	Activatee/Activator	Surette Lab, GIL
Bacteroides intestinalis GIL98	Activatee	Surette Lab, GIL
Bacteroides fragilis GIL100		Surette Lab, GIL
Bacteroides vulgatus GIL102	Activatee	Surette Lab, GIL
Bacteroides fragilis GIL120		Surette Lab, GIL
Bacteroides xylanisolvens GIL122	Activatee	Surette Lab, GIL
Bacteroides vulgatus GIL123	Activator	Surette Lab, GIL
Bacteroides xylanisolvens GIL136	Activatee	Surette Lab, GIL
Bacteroides ovatus GIL137		Surette Lab, GIL
Bacteroides fragilis GIL162		Surette Lab, GIL
Bacteroides intestinalis GIL165	Activatee	Surette Lab, GIL
Bacteroides cellulosilyticus GIL175	Activatee	Surette Lab, GIL
Bacteroides intestinalis GIL176	Activatee	Surette Lab, GIL
Bacteroides vulgatus GIL177	Activatee	Surette Lab, GIL
Bacteroides intestinalis GIL178		Surette Lab, GIL
Bacteroides thetaiotaomicron GIL179		Surette Lab, GIL
Bacteroides fragilis GIL180		Surette Lab, GIL
Bacteroides thetaiotaomicron GIL181		Surette Lab, GIL
Bacteroides xylanisolvens GIL182	Constitutive Hemolysis	Surette Lab, GIL
Bacteroides xylanisolvens GIL183	Activatee	Surette Lab, GIL
Bacteroides vulgatus GIL190		Surette Lab, GIL
Bacteroides vulgatus GIL194		Surette Lab, GIL
Bacteroides stercoris GIL204		Surette Lab, GIL
Bacteroides vulgatus GIL206		Surette Lab, GIL
Bacteroides uniformis GIL213	Activator	Surette Lab, GIL
Bacteroides uniformis GIL214		Surette Lab, GIL
Bacteroides fragilis GIL220		Surette Lab, GIL
Bacteroides fragilis GIL231		Surette Lab, GIL
Bacteroides xylanisolvens GIL238		Surette Lab, GIL
Bacteroides xylanisolvens GIL239	Activatee	Surette Lab, GIL
Bacteroides cellulosilyticus GIL243	Activator/Activatee	Surette Lab, GIL
Bacteroides fragilis GIL252		Surette Lab, GIL
Bacteroides fragilis GIL253		Surette Lab, GIL

Strain List	Properties	Source
Bacteroides fragilis GIL280		Surette Lab, GIL
Bacteroides xylanisolvens GIL282	Activatee	Surette Lab, GIL
Bacteroides xylanisolvens GIL284	Activatee	Surette Lab, GIL
Bacteroides xylanisolvens GIL293		Surette Lab, GIL
Bacteroides vulgatus GIL313		Surette Lab, GIL
Bacteroides xylanisolvens GIL314		Surette Lab, GIL
Bacteroides vulgatus GIL331		Surette Lab, GIL
Bacteroides fragilis GIL391		Surette Lab, GIL
Bacteroides vulgatus GIL418	Activatee	Surette Lab, GIL
Bacteroides vulgatus GIL419	Activatee	Surette Lab, GIL
Bacteroides fragilis GIL420		Surette Lab, GIL
Bacteroides fragilis GIL421		Surette Lab, GIL
Bacteroides fragilis GIL422		Surette Lab, GIL
Bacteroides vulgatus GIL423		Surette Lab, GIL
Bacteroides fragilis GIL425		Surette Lab, GIL
Bacteroides vulgatus GIL460		Surette Lab, GIL
Bacteroides uniformis GIL461		Surette Lab, GIL
Bacteroides vulgatus GIL480		Surette Lab, GIL
Bacteroides vulgatus GIL485		Surette Lab, GIL
Bacteroides vulgatus GIL504	Activatee	Surette Lab, GIL
Bacteroides vulgatus GIL505	Activatee	Surette Lab, GIL
Bacteroides vulgatus GIL506	Activatee	Surette Lab, GIL
Bacteroides fragilis GIL508		Surette Lab, GIL
Bacteroides vulgatus GIL510		Surette Lab, GIL
Bacteroides vulgatus GIL511	Activator	Surette Lab, GIL
Bacteroides fragilis GIL512		Surette Lab, GIL
Bacteroides fragilis GIL514		Surette Lab, GIL
Bacteroides fragilis GIL516		Surette Lab, GIL
Bacteroides fragilis GIL517		Surette Lab, GIL
Bacteroides vulgatus GIL552	Activatee	Surette Lab, GIL
Bacteroides vulgatus GIL554	Activatee/Activator	Surette Lab, GIL
Bacteroides ovatus GIL568	Activatee	Surette Lab, GIL
Bacteroides caccae GIL574	Activatee	Surette Lab, GIL
Bacteroides caccae GIL620	Activatee	Surette Lab, GIL
Bacteroides fragilis GIL625		Surette Lab, GIL
Bacteroides fragilis GIL626		Surette Lab, GIL
Bacteroides fragilis GIL627		Surette Lab, GIL
Bacteroides vulgatus GIL628		Surette Lab, GIL
Bacteroides vulgatus GIL629	Activator/Activatee	Surette Lab, GIL
Bacteroides uniformis GIL720		Surette Lab, GIL
Bacteroides uniformis GIL721		Surette Lab, GIL
Bacteroides vulgatus GIL726	Activator	Surette Lab, GIL
Bacteroides vulgatus GIL728	Activator	Surette Lab, GIL
Bacteroides vulgatus GIL731	Activator	Surette Lab, GIL
Bacteroides vulgatus GIL732		Surette Lab, GIL
Bacteroides vulgatus GIL733	Activatee	Surette Lab, GIL
Bacteroides ovatus GIL819	Activatee	Surette Lab, GIL
Bacteroides vulgatus GIL820	Activatee	Surette Lab, GIL
Bacteroides vulgatus GIL827	Activator	Surette Lab, GIL
Bacteroides caccae GIL829	Activator	Surette Lab, GIL
Bacteroides caccae GIL834	Activatee	Surette Lab, GIL
Bacteroides vulgatus GIL849	Activator/Activatee	Surette Lab, GIL
Bacteroides vulgatus GIL857	Activator	Surette Lab, GIL

Strain List	**Properties**	**Source**
Bacteroides uniformis GIL858		Surette Lab, GIL
Bacteroides vulgatus GIL863	Activator	Surette Lab, GIL
Bacteroides fragilis GIL895		Surette Lab, GIL
Bacteroides fragilis GIL898		Surette Lab, GIL
Bacteroides fragilis GIL904		Surette Lab, GIL
Bacteroides vulgatus GIL907		Surette Lab, GIL
Bacteroides vulgatus GIL908		Surette Lab, GIL
Bacteroides thetaiotamicron GIL913	Activator	Surette Lab, GIL
Bacteroides vulgatus GIL914		Surette Lab, GIL
Bacteroides fragilis GIL918		Surette Lab, GIL
Bacteroides fragilis GIL921		Surette Lab, GIL
Bacteroides fragilis GIL923		Surette Lab, GIL
Bacteroides fragilis GIL940		Surette Lab, GIL
Bacteroides fragilis GIL941		Surette Lab, GIL
Bacteroides fragilis GIL942		Surette Lab, GIL
Bacteroides fragilis GIL943		Surette Lab, GIL
Bacteroides fragilis GIL944		Surette Lab, GIL
Bacteroides uniformis GIL945		Surette Lab, GIL
Bacteroides uniformis GIL947		Surette Lab, GIL
Bacteroides uniformis GIL950		Surette Lab, GIL
Bacteroides caccae GIL1019		Surette Lab, GIL
Bacteroides caccae GIL1022		Surette Lab, GIL
Bacteroides vulgatus GIL1026		Surette Lab, GIL
Bacteroides vulgatus GC186	Activatee	Surette Lab, GC
Bacteroides vulgatus GC195	Activatee	Surette Lab, GC
Bacteroides vulgatus GC196	Activatee	Surette Lab, GC
Bacteroides vulgatus GC197	Activatee	Surette Lab, GC
Bacteroides vulgatus GC397	Activatee	Surette Lab, GC
E. coli S-17 λ pir-pSAM-Bt	TpR SmR recA, thi, pro, hsdR-M+RP4: 2-Tc:Mu: Km Tn7 λpir	Yost Lab

Table S2: List of Sequenced Isolates with Quality Metrics.

Isolate	*Total length*	*# contigs*	*N50*	*L50*	*Largest contig*	*GC (%)*
B. vulgatus GIL97	5064995	175	97158	17	426251	42.33
B. vulgatus GIL102	5063128	154	102888	16	239320	42.32
B. vulgatus Gll504	5807332	253	5738	34	210890	43.03
B. vulgatus GIL552	5807332	253	57358	34	210890	43.03
B. vulgatus GIL629	5719593	296	49725	37	164213	43.12
B. vulgatus GIL726	5722810	837	50850	31	199946	41.74
B. vulgatus GIL849	5238376	109	110825	16	426035	42.30
B. vulgatus GIL914	5229041	160	69152	25	211303	42.31
B. ovatus GIL819	6829030	311	49554	41	313388	41.94
B. vulgatus GIL820	5183595	267	43079	41	138176	42.41

Table S3: Primers Used in This Study

Primer name	Sequence	Purpose
27F	AGAGTTTGATCCTGGCTCAG	16S Identification
1492R	GGTTACCTTGTTACGACTT	16S Identification
SAMSeq1	ACGTACTCATGGTTCATCCCGATA	SRP-PCR 1
SAMSeq2	GCGTATCGGTCTGTATATCAGCAA	SRP-PCR 2
SAMSeq3	TCTATTCTCATCTTTCTGAGTCCAC	Sanger Sequencing of TIS
SRP1	GGCCACGCGTCGACTAGTACNNNNNNNNNGATAT	SRP-PCR 1
SRP2	GGCCACGCGTCGACTAGTAC	SRP-PCR 2

Table S4: SRP-PCR Step 1

	TEMPERATURE (°C)	TIME
HOLD	95	2 min
6 CYCLES	95	30 s
	30	30 s
	68	1.5 min[1]
30 CYCLES	95	30 s
	45	30 s
	68	2 min[1]
FINAL EXTENSION	68	5 min
HOLD	4	

[1]Increase by 5 s increment each cycle

Table S5: SRP-PCR Step 2

	TEMPERATURE (°C)	TIME
HOLD	95	10 min
35 CYCLES	95 55 68	45 s 45 s 1.5 min[1]
FINAL EXTENSION	68	10 min
HOLD	4	-

[1]Increase by 5 s increment each cycle

Table S6: Results generated for hypothetical protein in Hit 6A3C by SWISS-MODEL

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
4k3u.1.A	21.78	monomer	-	HH-blits	X-ray	2.16Å	0.31	0.77	GDSL-like Lipase/Acylhydrolase family protein
4k3u.1.A	24.36	monomer	-	BLAST	X-ray	2.16Å	0.33	0.78	GDSL-like Lipase/Acylhydrolase family protein
3bzw.1.A	18.99	homo-hexamer	-	HH-blits	X-ray	1.87Å	0.29	0.35	Putative lipase
3dc7.1.A	18.35	homo-trimer	-	HH-blits	X-ray	2.12Å	0.28	0.35	Putative uncharacterized protein lp_3323
4jj6.1.A	17.53	homo-octamer	-	HH-blits	X-ray	1.80Å	0.28	0.34	Acetyl xylan esterase
3skv.1.A	17.53	monomer	-	HH-blits	X-ray	2.49Å	0.29	0.34	SsfX3
4jko.1.A	17.53	homo-octamer	-	HH-blits	X-ray	1.90Å	0.28	0.34	Acetyl xylan esterase
5bn1.1.A	18.18	homo-octamer	-	HH-blits	X-ray	1.60Å	0.28	0.34	Acetyl xylan esterase
4jj4.1.A	16.88	homo-octamer	-	HH-blits	X-ray	2.13Å	0.27	0.34	Acetyl xylan esterase
3w7v.1.H	18.18	homo-octamer	-	HH-blits	X-ray	1.85Å	0.28	0.34	Acetyl xylan esterase
2o14.1.A	14.84	monomer	-	HH-blits	X-ray	2.10Å	0.27	0.34	Hypothetical protein yxiM
5w7c.1.B	16.13	monomer	-	HH-blits	X-ray	2.23Å	0.29	0.34	Acyloxyacyl hydrolase
4iyj.1.A	16.45	homo-dimer	-	HH-blits	X-ray	1.37Å	0.29	0.33	GDSL-like protein
3skv.2.A	17.53	monomer	-	HH-blits	X-ray	2.49Å	0.29	0.34	SsfX3
4ppy.1.A	14.29	homo-dimer	-	HH-blits	X-ray	2.00Å	0.28	0.34	Putative acylhydrolase
5w78.1.B	16.13	monomer	-	HH-blits	X-ray	2.27Å	0.29	0.34	Acyloxyacyl hydrolase
6jl2.1.A	15.85	monomer	-	HH-blits	X-ray	2.30Å	0.28	0.36	Thermolabile hemolysin
4q9a.1.A	12.74	homo-dimer	-	HH-blits	X-ray	2.86Å	0.27	0.35	Tat pathway signal sequence domain protein
4oap.1.A	18.18	homo-octamer	-	HH-blits	X-ray	1.93Å	0.28	0.34	Acetyl xylan esterase
3p94.1.A	13.64	monomer	-	HH-blits	X-ray	1.93Å	0.28	0.34	GDSL-like Lipase

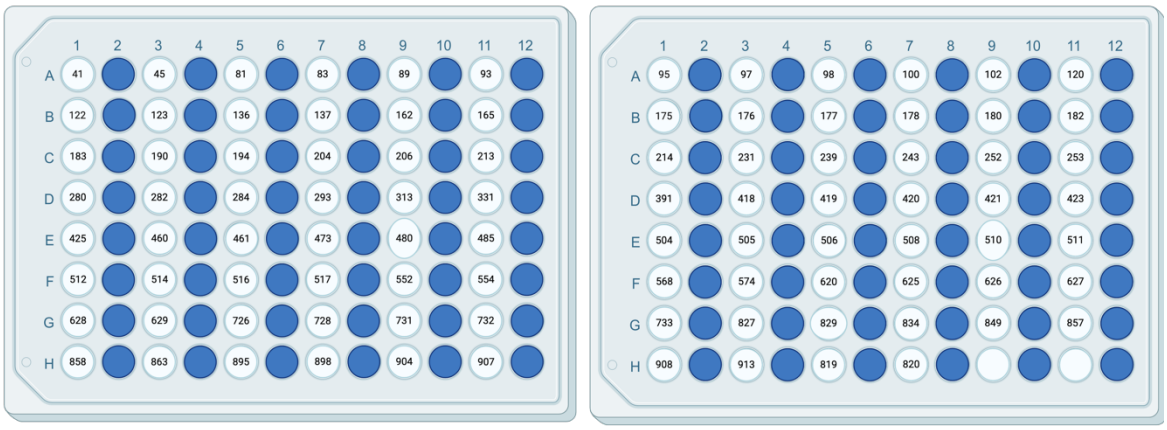

Figure S1. GIL820 screen for activation plate layout.
Blue circles represent GIL820, numbers indicate GIL isolate number.

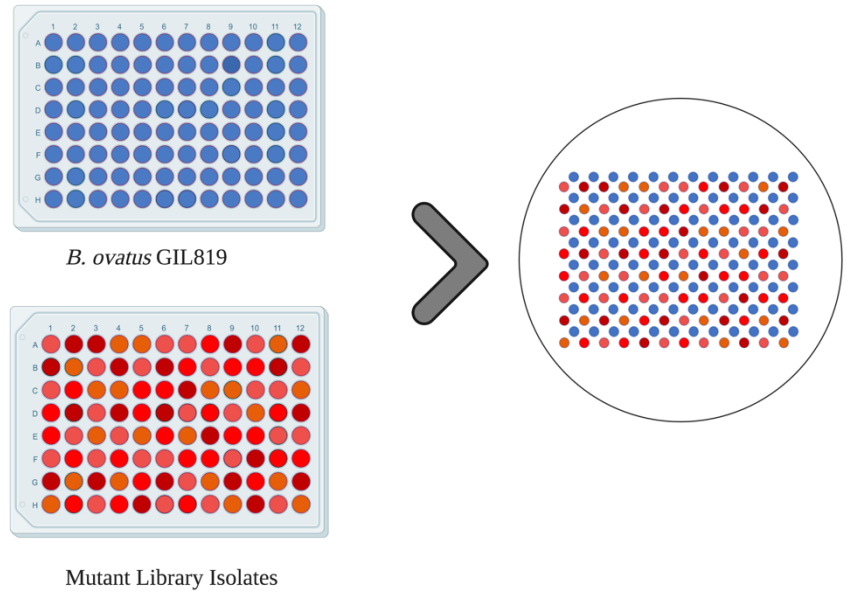

Figure S2. Initial mutant library screening protocol.

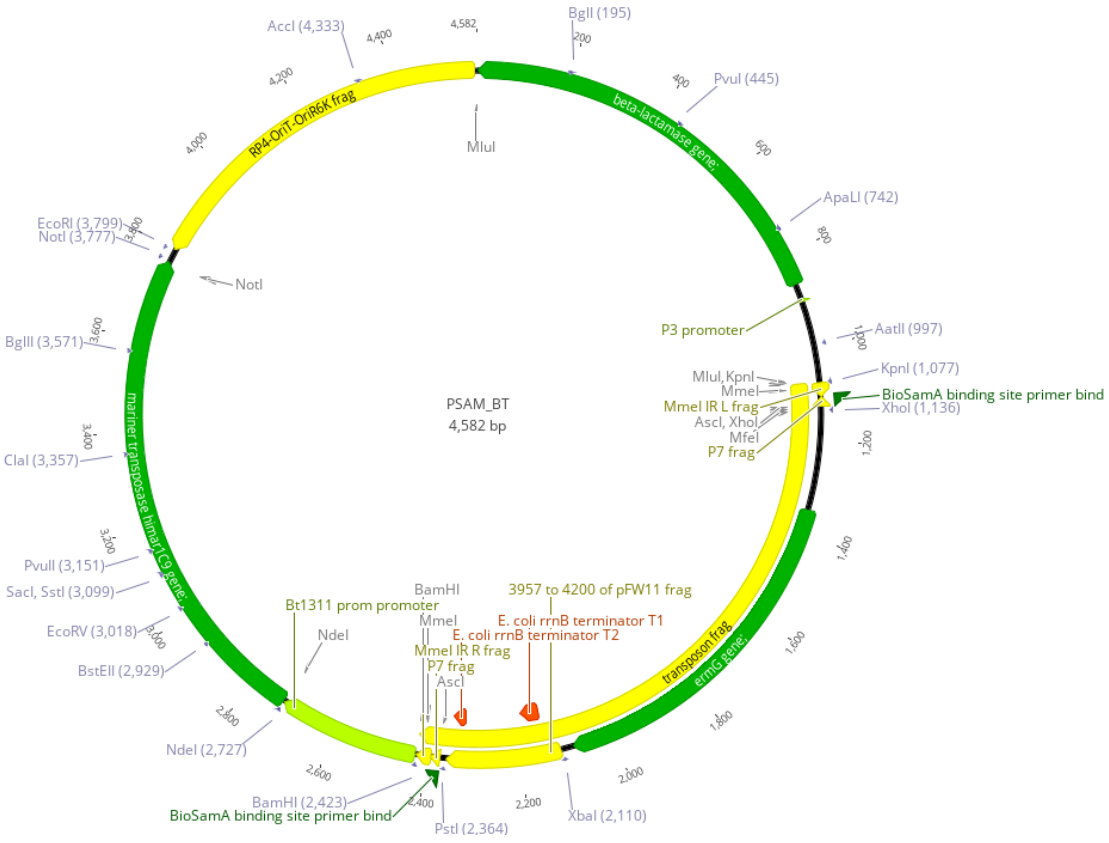

Figure S3. pSAM_Bt plasmid map. (Obtained from Addgene)

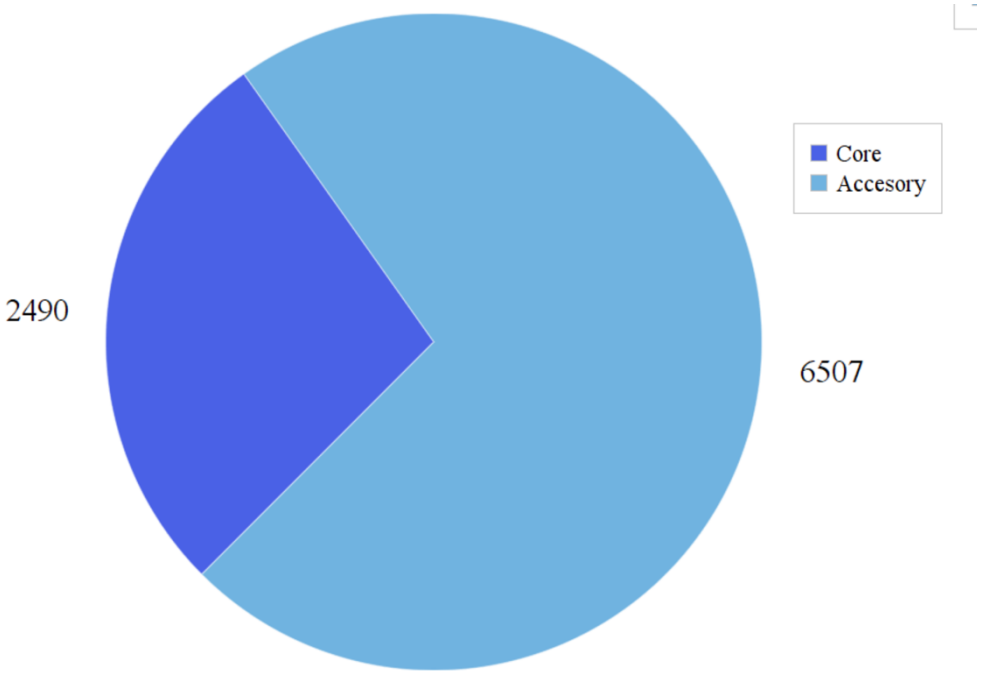

Figure S4. The predicted gene composition of the pangenome of B. *vulgatus* isolates.
Genes of the pangenome are categorized by the number of isolates that share a gene. 2490 genes total, or 28% of genes are shared by all 15 isolates. 6507 genes differ amongst all the isolates.

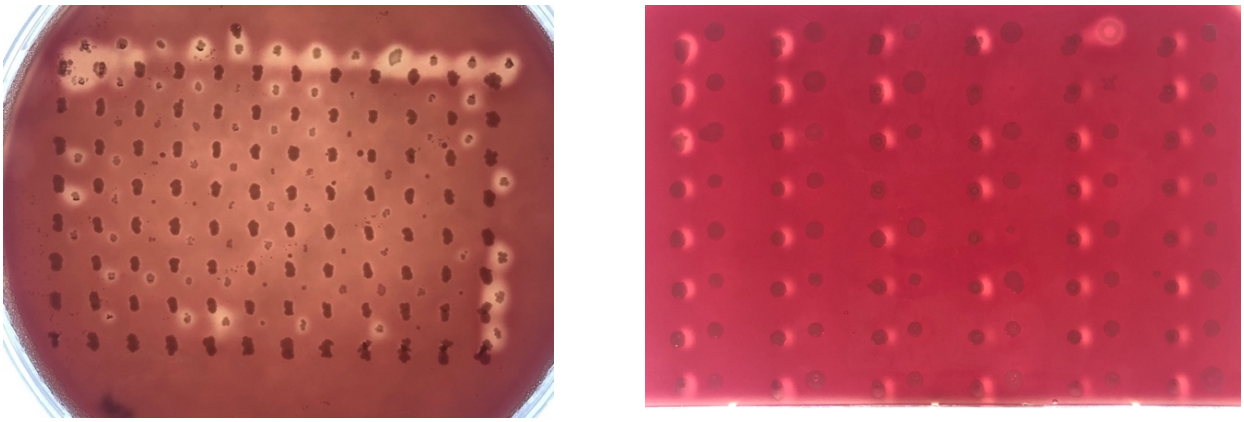

Figure S5. Mutant library screening before (left) and after (right) optimization.