Isolation, Culturing and Nutrient Analysis of *Candidatus* arthromitus

A Thesis
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Holly Ann Reiland

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

Advisor
Dr. David J. Baumler

August, 2016

Acknowledgements

There are many that I would like to thank for support and guidance throughout my graduate studies and the achievement of writing this document. First and foremost, I would like to thank my adviser, Dr. David Baumler, for the opportunity to study food science at the graduate level, for all of his guidance in both school and life in general, and for funding the majority of my Master's degree with his start-up funds. I also thank my peers Zachary Metz, Justin Wiertzema, Morrine Omolo, and Tong Ding as wells as undergraduate researchers Zen-Zi Wong, Eleni Beyene, Matthew (Fred) Frederickson, Nina Le, and Julianne Branca for all of the help with laboratory upkeep and maintenance. I would like to thank my friends on floor two of Andrew Boss Lab of Meat Science for keeping my spirits high. I would like to thank my mother for helping me "keep it together" when I was overwhelmed at times and for raising me to be the person that I am today. Thank you to my committee members Dr. Baraem Ismail and Dr. Timothy Johnson for supporting me in this process and thank you Dr. Johnson for the preliminary data provided at the start of my project. Thank you to Kyle Case for the many protocols and assistance with running through each of them for the first time. Thank you to Ana Guedes-Hasegawa and colleagues at Roche for the Lightcycler 96 qPCR machine and guidance through understanding how to run experiments and analyze data. Thank you to the University of Minnesota, the College of Food, Agricultural and Natural Resource Sciences, and the Department of Food Science and Nutrition for funding the remainder of my

degree through teaching assistantship funds. Thank you to the Skara Fellowship for funding my spring 2015 semester of school. Finally, thank you to my incredibly supportive and amazing fiancé, Maxwell Krajnik, for supporting me through this, I love you.

Dedication

- To my loving fiancé Maxwell Krajnik for the supporting me and believing me when I had trouble believing in myself
- To my mom for absolutely everything, I am a production of your success
- To my brother for making me smile and laugh effortlessly
- To the friends and family I've lost, I wouldn't be here without your influence

Abstract

Candidatus arthromitus (CA) is a Gram positive, spore-forming segmented filamentous bacteria known to be a commensal if not symbiotic organism residing in the gastrointestinal tract. *Candidatus* arthromitus was first isolated and grown using the methods explained in Schnupf *et al*, 2015 (3). Isolated cells were cultured using brain heart infusion broth (BHI) with additional carbohydrate sources predicted by systems biology and genome-scale metabolic modeling techniques to increase growth. Methods for culturing and computational predictions are described in the materials and methods found in chapter four of this thesis. Cultured CA was used in Biolog ™ assays to determine CA metabolic capabilities in comparison to *in-silico* predictions. Growth curves and dry cell weight experiments were completed on carbohydrate-spiked BHI broth to provide supporting evidence of CA external host viability and *in vitro* growth. This is the first effort to culture CA from commercial turkeys; the importance stems from an industry issue with commercial turkeys failing to reach full weight potential at an early age. This term describing this issue is Light Turkey Syndrome (LTS).

Table of Contents

List of Tables

List of Figures

1. General Introduction

One of the current challenges facing commercial turkey production in Minnesota is the difficulty obtaining flock average weights typical of the industry standard. A condition that contributes to this problem of obtaining average flock weights has been coined “Light turkey Syndrome” or LTS (Danzeisen et al., 2013). This condition has been identified in Minnesota turkey flocks for at least five years, and it has been observed that flock body weights never approach their growth potential (Danzeisen et al., 2013). A single causative agent responsible for these weight reductions has not been identified despite numerous research efforts (Danzeisen et al., 2013). It has been observed that the intestinal microbiome composition of under-performing turkeys contains fewer *Candidatus* arthromitus (CA) and they appear later in these birds than in high-performing flocks (Danzeisen et al., 2013).

The term Arthromitus was coined by Joseph Leidy in 1849 (Snel et al., 1995). The first portion of the word *Arthron* is Greek for “joint” and *mitos* meaning “thread” and describes what Leidy observed while examining animal intestinal microbiome communities under a microscope as well as in termites and arthropods (Snel et al., 1995, Schnupf et al., 2013). Molecular typing separated Segmented Filamentous Bacteria (SFB) into two separate subphyla of *Clostridiaceae* and *Lachnospiraceae.* SFB isolated from mice were first studied in 1962 and were classified under the *Clostridiaceae* subphylum (Snel *et al*., 1995). Vertebrate SFB show a high degree of similarity, such as the hold-fast

structure used to attach to ileum epithelial cells, and while *in situ* hybridization of the human microbiome under light microscopy revealed filaments, stool samples from humans revealed coccoid structures (Snel *et al*., 1995, Bolotin *et al*., 2014a, Schnupf *et al*., 2013).

Candidatus arthromitus is a Gram positive, segmented, filamentous bacterium (SFB) known to populate the gastrointestinal tract, primarily the lamina propria--of vertebrates and arthropods (Schnupf et al., 2013). The lamina propria is a constituent of the moist linings known as mucous membranes, or mucosa, which line various tubes in the body (Schnupf et al., 2013). According to previous research, vertical passage of SFB from mother to offspring is common (Schnupf et al., 2013, Snel et al., 1995). There has been no known horizontal passage of SFB from species to species observed indicating host-specificity; however, phylogenetic information has demonstrated that *Candidatus* species from different hosts are all similar to *Clostridia* species having close to 86% similarity as found by 16s DNA sequencing (Snel et al., 1995). *Candidatus* spp. have been linked to many functions including the priming and maintenance of healthy gut mucosal immune function (Schnupf et al., 2013). In one study, CA played a role in both pro and anti-inflammatory responses, having to do with the function of both B and helper T cells (Schnupf et al., 2013, Schnupf et al., 2015). Particularly in mice, SFB have been known to congregate on the epithelial cells of ileum tissue in the small intestine, preventing the colonization of deleterious bacteria (Schnupf et al., 2013, Schnupf et al., 2015). CA were first

monocolonized in mice in 1991. This was achieved by using homogenous suspensions of intestinal epithelial cells washed in chloroform and ethanol to inoculate germ-free mice via the intestinal tract. In 1995, monoclonization was independently achieved with success by a research group in Japan by Umesaki et al. who used 50 SFB isolated in chloroform and ethanol to administer to germ-free mice orally (Klaasen et al., 1992). Fecal suspension from these mice were then used to sub-culture SFB in new mice. When mice were monocolonized with SFB, an immune response of CD4+ helper T cells was produced (Schnupf et al., 2015). Also, colonization with SFB was correlated with an increase in gene expression of genes associated with inflammation, along with antimicrobial defense mechanisms (Schnupf et al., 2015). This demonstrates the ability of SFB to trigger innate immune response while causing no harm to the host.

SFB produce a large amount of proteins devoted to the development of cell cycle control and envelope biogenesis along with having a group of SFB-specific autolysins and dynamin-like proteins (Pamp et al., 2012). The presence of nucleotide polymorphisms also suggests the evolution of SFB-distinct lineages. A recent study attempted to relate five strains isolated with laser tweezers from a variety of vertebrate animal hosts including dog, monkey, mouse, rat, chicken, and trout to bacterial species within the human gastrointestinal tract (Pamp et al., 2012). Genome analysis was performed and approximately 8.6% of predicted proteins were newly discovered including several SFB-specific secreted proteins (Pamp et al., 2012). These proteins are believed to be involved in modulation of

host response and demonstrate the ability of SFB's adaptation to the mammalian intestinal tract (Pamp et al., 2012).

The objectives of this thesis were to isolate *Candidatus* arthromitus from extracted turkey ileums to determine nutrient utilization using experimental and *in silico* simulations in order to generate large volumes of CA culture for future turkey poult inoculation. The hypothesis is that by supplementing under-performing birds with CA, birds will overcome LTS and reach their full weight potential. Justification of the potential success of this project stems from initial rRNA metagenomic results found to show a higher prevalence of CA in heavy flocks in Minnesota versus lighter flocks (Figure 3).

2. Literature Review

2.1 Introduction. This chapter serves as a summary of knowledge of the microorganism *Candidatus* arthromitus (CA). The information in this literature review contains extensive information in categories such as metabolic analysis and host-cell interactions but is limited in *in vitro* experimental results as only one single lab has documented the culturing of this organism outside of a host.

2.2 Attachment. SFB were found to utilize brush-border epithelial attachment also referred to as the microvilli found on the surface of epithelial cells. (Ericsson et al., 2014a). Segmented and circular spores exist in the same area, and can exist in either dormant or vegetative forms. It is suggested that sporulation, which requires oxygen, occurs throughout the life cycle of SFB. The ileum however, does not contain much oxygen if any (Ericsson et al., 2014a). SFB attach at one of their ends that consists of a rounded, nipple-like structure called a hold-fast. The hold-fast attaches but does not penetrate the host's epithelial cell wall, only penetrating through the mucus membrane. Focal locations are disturbed when attachment of SFB occurs, which alters the electron density of both the host cell plasma membrane and apical cytoplasm. Once a SFB has attached itself via the hold-fast and is anchored, actin polymerization occurs directly underneath the hold-fast structure and creates a pedestal-like formation that is similar to the adherence of *Salmonella Typhimurium* and *Escherichia coli* (Jepson et al., 1993,

Ericsson et al., 2014a). SFB length increases upon maturity and distal segments are thought to release holdfasts and spores. Spread of SFB then occurs via vertical transmission i.e. from parents to offspring. This phenomenon is due to spore inoculation that can then become endemic. The first report of the habitat of what we now know as SFB was published in 1974. SFB are 0.7 microns to 1.8 microns in diameter and up to 80 to 100 microns in length when mature. Colonization occurs in mouse and rat hosts right before the weaning process and has been found to be the same in arbitrary human microbiome studies (Ericsson et al., 2014a). Studies performed on human subjects ages 0 months to 75 years old revealed that 25% of individuals carry SFB in their gut from ages 0-6 months, 75% carry SFB from ages 7-12 months, and only 6.2% carry SFB from ages 3-75 yrs (Yin et al., 2013).

2.3 Life Cycle*.* The life cycle of an SFB is postulated to be about 2-3 days based on the rapid shedding of the intestinal epithelial cells of rodents, and longitudinal studies have shown that SFB appear in juvenile mice that are around 20 days in age. At this age, SFB proliferate to become a dominant gut microbe and then recede in mature vertebrates to lower levels. During the early stages of colonization, SFB are transiently colonized with rod-shaped bacteria (Schnupf et al., 2013).

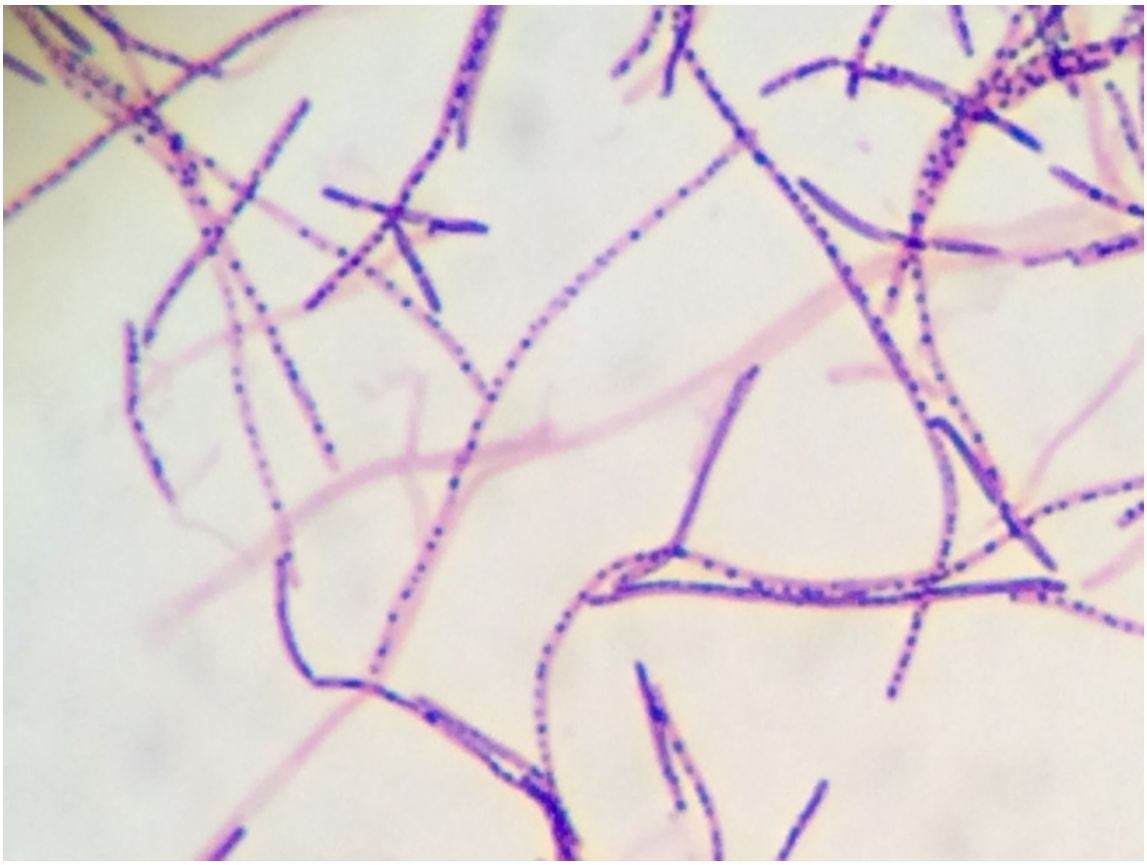

Figure 1: Gram Stain of *Candidatus* arthromitus from turkey ileum containing intrasegmental bodies.

SFB spores transmitted to weaning offspring germinate in the gut to produce teardrop-shaped single-celled bacteria referred to as intracellular offspring (Schnupf et al., 2013). Intracellular offspring released upon germination may use flagella to reach the epithelial surface of the small intestine in the lamina propria. The analysis of a SFB genome revealed a full set of chemotaxis and flagellum biosynthesis genes (Schnupf et al., 2013). However, flagella have not been observed microscopically and only low levels of expressed genes associated with flagella could be detected in the small intestine where SFB are abundant, which may reflect the scarcity of spores found in this location. The nucleoid structure of

intracellular offspring released from filaments appears condensed, suggesting minimal transcriptional activity. This structure expands along with the growth of the offspring's filaments and upon attachment of intracellular offspring to the host (Martin et al., 2009). Intracellular offspring attach to absorptive epithelial cells via their holdfast and induce condensed actin rearrangements underneath the point of attachment while leaving the neighboring microvilli structures unaffected (Schnupf et al., 2013). Attachment of SFB appears to be species-specific. Transference of fecal matter between mice and rats failed to show any ileal attachment of SFB in the non-endogenous host of SFB sample origin (Schnupf et al., 2013). Following attachment in endogenous hosts the intracellular offspring increases in width and length, reaching up to 5 μm or more in length before bacterial division commences through transverse septum formation. Filaments continue to grow from their non-attached end, reaching 50–100 μm in length (Schnupf et al., 2013, Martin et al., 2009). A second round of segmentation then starts from the free end to divide each primary segment in half. Differentiation of these secondary segments is initiated, again at the free end, whereby individual segments divide unequally to form a mother cell that engulfs the smaller daughter cell. This daughter cell divides to form two intracellular offspring within the mother cell (Schnupf et al., 2013, Martin et al., 2009). Under continuous growth conditions in the gut, the septa of the mother cells that separate individual intracellular offspring compartments degenerate to form a hollow tube from which intracellular offspring are released at the distal end

(Martin et al., 2009, Schnupf et al., 2013). This completes the life cycle within the host as intracellular offspring are again free to attach to the host cell and undergo another round of filamentation and differentiation (Schnupf et al., 2013, Martin et al., 2009). When an unfavorable environment is presented, the two intracellular offspring become surrounded by one spore coat and mature into a full spore inside of the mother cell (Schnupf et al., 2013, Martin et al., 2009) . Spores are released from the filament and then shed in the feces for future transmission into a new host (horizontal transfer).

2.4 Host Specificity. Histological, post-mortum examination of the ileum contents requires a skilled pathologist to determine which microbes are SFB and which are other microorganisms. However, DNA primers targeting 16s rDNA usage is a non-invasive way of examining fecal samples of study subjects. Host specificity was discovered through studying SFB in these ways (Tannock et al., 1984). SFB distinguish hosts by the way in which attachment occurs to areas of the gut. In most rodent and pig hosts, SFB attach to follicle-associated epithelial cells (over Peyer's Patches) and absorptive villi (Tannock et al., 1984). In mice and horse hosts, attachment occurs primarily to follicle-associated epithelial cells. In rabbits, cows, and dogs, attachment occurs primarily to absorptive villi (Pamp et al., 2012).

There are several factors that have been discovered about SFB that explain their auxotrophic nature. The genome of SFB isolated from a rat host

(Rat-YIT) contains 28 putative genes predicted to encode for proteases and 53 for peptidases along with many other genes through to be involved with sporulation and germination (Prakash et al., 2011). Peroxidase and catalase genes were also found, which explains the potential for SFB to exist in microaerophilic environments (Prakash et al., 2011). Clustered Regularly Interspaced Palindromic Repeats (CRISPRS) were found in genomes of SFB sequenced from mice and rat hosts as well (which are generally a defense mechanism for prokaryotic DNA) meaning that SFB genomes may have had exposure to invading DNA. (Prakash et al., 2011). Flagellar, pilus, and chemotactic genes have been found in SFB genomes that suggest motility, which explains the organism's ability to penetrate the mucus layer lining of intestinal epithelial cells (transient motility) (Prakash et al., 2011, Schnupf et al., 2013).

2.5 Host Interactions. Several interactions between the host and SFB have been investigated, and, though both positive and negative associations have been found, only the rainbow trout host has been known to show a negative inflammatory disease response from SFB colonization (Del-Pozo et al., 2010). Monoassociated mice have shown increased numbers of the antibody IgA secreting cells and more IgA titers in their intestines and serum (Del-Pozo et al., 2010). Both specific and natural non-specific IgA production occurs with SFB monocolonization of mice (Schnupf et al., 2015). It is also this increased presence of IgA that has been found to limit the proliferation of SFB in adults. Also in monoassociated mice, SFB generated *in vitro* were found primarily in the

cecum while *in vivo* SFB proliferate in the ileum (Schnupf et al., 2015). This has an impact on the magnitude of the host's innate immune response, which relies on colonization of the ileum (Schnupf et al., 2015). Presence of SFB also have been linked to enhancing the development of CD4+ helper T cells in Peyer's Patches and are most effective on Th17 cells (Schnupf et al., 2015). Th17 cells are involved with protection against *Enterobacteriaceae* via protective effects from IL22--a cytokine that produces antimicrobial peptides that prevent pathogens from inducing attaching and effacing lesions (Schnupf et al., 2015). Also, flagella protein presence may be linked to Th17 cell induction and IgA production via TLR5 (toll-like receptor 5). Binding sites of flagella and TLR5 have been identified and the specific motif that codes for TLR5 is highly conserved in flagella proteins of SFB but is absent in all *Clostridium* spp. (Schnupf et al., 2015). Finally, sex-specific protection against diabetes in non-obese diabetic mice has been found, yet more research in this area must be done to support this hypothesis (Kriegel et al., 2011).

Overall, Ericsson *et al.* (2014) provided an explanation of characteristics that will aid in understanding future research of SFB (7). This review serves as a building block for future experiments such as designing and utilizing metabolic models in order to determine nutrient to develop an optimal culture medium that will allow for SFB to be cultured *in vitro* for further characterization. Ericsson *et al.* (2014) provides a basis for isolation of CA involving host epithelial cells (Ericsson et al., 2014b). Through the experimental methods in this paper, it has

been found that CA can be isolated using a combination of penicillin/streptomycin/amphotericin-B (PSA) antibiotic complex added into Dulbecco's Modified Eagle Medium (DMEM), high-glucose containing 4 mM L-glutamine, 4,500 mg/L glucose, and sodium pyruvate (catalog #SH30243.02). Dulbecco's Modified Eagle Medium (DMEM) containing antibiotics was used initially with the replacement of DMEM not containing antibiotics at days 7 and 14 (Ericsson et al., 2014b). A 1:2 dilution of PSA in microaerophilic conditions (2-16% O_2) optimized the isolation of CA. To obtain sample material, ileum sections were harvested from BALB/cAnHsd 2-3 days post weaning mice. Ileum samples were incised longitudinally and rinsed with sterile Phosphate-Buffered Saline (PBS) (Ericsson et al., 2014b). Mucosal scraping was performed by pooling four samples together to create an ileal mucosal scrape (IMS) material. Ciproflaxin (Ericsson et al., 2014b) and neomycin were also used in varying amounts; however, the addition of these antibiotics resulted in either contamination of IMS material or complete elimination of CA . Although efforts to culture CA were not successful, this article provides a cost-effective and simple method for CA isolation.

SFB have been linked to gastroenteritis in rainbow trout (Del-Pozo et al., 2010). SFB inhabitance in this species occurs in the distal intestine as with most other species; however, instead of attaching to enterocytes that remain attached to the intestinal walls, the enterocytes detach from those walls (Del-Pozo et al., 2010). This detachment is causative of the subsequent congestion of the lamina

propria. SFB were not always adjacent to the affected areas, so extracellular influences have been assumed to play a role in the occurrence of gastroenteritis in rainbow trout (Del-Pozo et al., 2010).

2.6 Characteristics. The genome of a rat isolated SFB and a number of mouse SFB isolates have recently been sequenced and published (Sczesnak et al., Pamp et al., 2012, Prakash et al., 2011, Kuwahara et al., 2011). The SFB genomes from both rat and mouse hosts are highly similar but do contain several species-specific genes of unknown function that may be involved in the species-specificity of SFB colonization (Prakash et al., 2011). All SFB have a small genome size of 1.5–1.62 Mbp in comparison to most free living bacteria, a low GC-content (27.9%), and relatively small number of genes (1,420–1,534 for mouse SFB) (Prakash et al., 2011). This reduced genome indicates that SFB have highly auxotrophic needs and may require certain metabolites from hosts. The biosynthetic pathways of most amino acids, vitamins and cofactors (such as B1, B2 and B12, pyridoxine, nicotinamide, pantothenate and biotin) are incomplete or absent altogether in SFB genomes (Sczesnak et al., Pamp et al., 2012, Kuwahara et al., 2011). SFB are also unable to synthesize nucleotides independently; instead they utilize alternative pathways that rely on the uptake of nucleotide bases (Prakash et al., 2011). To obtain nucleotides, amino acids, and peptides from the environment, SFB genomes contains genes encoding two extracellular nucleases as well as a list of proteases and peptidases, 20 of which are membrane associated and four to six that are thought to be secreted

(Kuwahara et al., 2011, Sczesnak et al.). In addition, SFB genomes contain numerous ORFs thought to encode a large number of transporters and permeases for small molecules and ions (such as amino acids, oligopeptide, dipeptides, manganese, zinc, iron and phosphate) compared to other organisms with small genomes (Sczesnak et al.). A particularly strong requirement for iron uptake was noted by Sczesnak et al., since six different ORFs for iron transporters are found in the mouse genome as well as three ORFs for ferric iron regulator family proteins (Sczesnak et al.). SFB also have several ORFs for phosphotransferase systems predicted for uptake of sugars such as mannose, cellobiose, mannitol and fructose as well as for ascorbate (Sczesnak et al.) (Prakash et al., 2011). Finally, the SFB genomes contain genes for the nonoxidative pentose phosphate pathway and a complete glycolysis pathway to convert glucose to pyruvate but is deficient for genes encoding almost all components of the Krebs cycle, which is required for aerobic respiration (Prakash et al., 2011, Sczesnak et al.). However, SFB can tolerate some oxygen and counteract oxidative stress, since its genome contains genes predicted for two catalases, a peroxidase (rubrerythrin), and an arginase, which might limit nitric oxide production through catabolism of arginine (Kuwahara et al., 2011) (Pamp et al., 2012). These protective mechanisms are likely essential, given the replicative niche of SFB at the surface of the small intestinal epithelium where the oxygen tension is estimated to be around 1.4% (He et al., 1999). Research environments, such as an anaerobic hood, can provide microaerophilic

atmospheric conditions to aid researching *Candidatus* arthromitus directly outside of hosts.

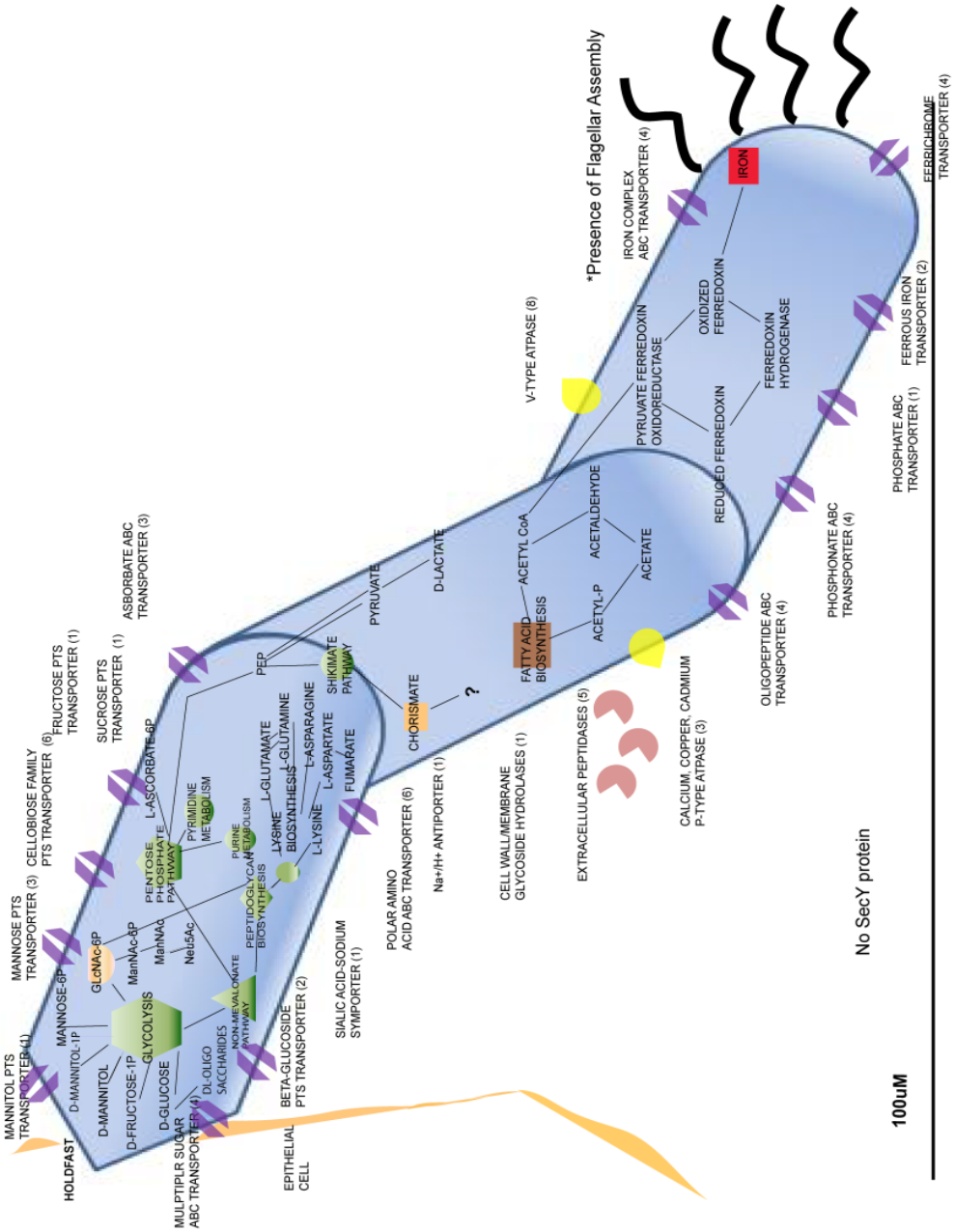

Figure 2: Metabolic Features of *Candidatus* arthromitus through inference of genome contents and examination using RAST

2.7 Medicinal properties. In mice, SFB are responsible for postnatal maturation of gut immune functions as well as producing an IgA response that recruits CD8+ T lymphocytes and lamina propria CD4+ T cells (Schnupf et al., 2013, Schnupf et al., 2015). Also, SFB are thought to be responsible for inducing the histocompatibility complex class II molecules in host intestinal epithelial cells (Ferguson and Birch-Andersen, 1979). An electron morphological experiment was designed in order to study the host-species interaction to provide an understanding of the relationship between the inductions occurring due to attachment to host epithelial cells (Ferguson and Birch-Andersen, 1979). Phagocytosis, endocytosis, and transcytosis that characterize the follicular epithelium are involved in the interaction between SFB and host cells (Ferguson and Birch-Andersen, 1979).

SFB induce Th17 cells, meaning that SFB increase Th17 presence and differentiation. Th17 cells are responsible for inflammatory immune responses and the SFB attach to the epithelial cells and prevent pathogenic bacteria from infecting the host (Schnupf et al., 2013) *Citrobacter rodentium* bacteria were used as a model to demonstrate this property of SFB. Previous experiments have shown SFB-host resistance to bacterial pathogens *Listeria monocytogenes*, *Escherichia coli*, *Salmonellla Enteriditis* and *Citrobacter rodentium* (Caselli et al., 2010). Both antibiotics and probiotics have been proven to be active against SFB in mice (Caselli et al., 2010). Penicillin has been found to remove SFB from a mouse ileum, but SFB restored themselves in this host niche several weeks after

the treatment was halted (Caselli et al., 2010). A *Lactobacillus plantarum* strain introduced into immunosuppressed mice increased/restored the number of SFB in the mouse ileum (Caselli et al., 2010).

2.8 Colonization*.* The host diet may also influence SFB colonization. SFB abundance significantly decreased in mice deficient in dietary retinoic acid (Schnupf et al., 2015). Also, an unusual relationship of SFB with *Lactobacillus* spp. has been suggested. Murine and chicken SFB can be found sub-colonized with bacteria resembling cells of *Lactobacillus* (Klaasen et al., 1992) and (Koopman et al., 1987), and Ivanov *et al.* correlated a greater abundance of *Lactobacillus murinus* with the presence of SFB. Koopman, et al (*Koopman et al., 1987)* describes a loose, fuzzy outer layer on the cell surface of SFB that may play a role in rod-shaped bacteria proposed to be *Lactobacillus* spp. attaching to SFB cells.

SFB colonization on intestinal epithelial cells causes a positive response from the host's immune system where as SFB colonization on mucosal cells does not occur (Schnupf et al., 2015). SFB are an important trigger of the post-natal maturation of the gut immune system in mice, but that synergy between SFB and other members of the microbiome is necessary to fully stimulate homeostatic gut immune responses (Schnupf et al., 2015). Stimulating the host's innate and adaptive immune responses creates a barrier effect inside of the host. This demonstrates the ability of SFB to strengthen host epithelial cell walls (Schnupf et al., 2013). Similar to the distantly related bacterium *Clostridium*, SFB might

notably produce peptidoglycan, lipoproteins, teichoic acid, and lipoteichoic acid, which may all act as microbial associated patterns to stimulate a host immune reaction (Schnupf et al., 2013).

Overall, studies indicate that in immunocompetent hosts, SFB induce a self-limited immune response that does not impair, but rather fosters intestinal homeostasis (Schnupf et al., 2013).

Mice that are colonized by microbiota that lack SFB have weaker IgA antibody responses and much less intestinal T cell responses compared with mice that are colonized with SFB (Caselli et al., 2010). Notably, mice that are colonized by an SFB-deficient microbiota lack mucosal Th17 cells (Caselli et al., 2010). Furthermore, these animals cannot control colonization by the invasive pathogen *Citrobacter rodentium*, which suggests that microbiota-induced immune responses participate in the barrier function of the flora (Cerf-Bensussan and Gaboriau-Routhiau, 2010).

A unifying feature of all SFB is the host interaction involved in colonization (Snel et al., 1995). SFB length ranges in length from 50-1,000 microns. They can also exist as single “hold-fasts” or spores (Snel et al., 1995). Both spores and hold-fasts attach to epithelial cells in the ileum and produce segment cells through continuous growth and septa formation (Snel et al., 1995). Host-specificity has been identified as many attempts to implant SFB from one host species into a germ-free environment of another host type have failed (Snel et al., 1995).

In one study, complete genome sequencing and comparative analysis to determine conservation between Rat and Mouse SFB (YIT) strains was performed (Prakash et al., 2011). Fecal samples collected from the mice were compared to information gathered from rat ceca in order to compare the two sources to determine degrees of similarity and host-specificity (Prakash et al., 2011). DNA was isolated using the lysis method using both lysozyme and purified achromopeptidase (Prakash et al., 2011). Whole genome shotgun and assembly were performed using a combination of Sanger and 454 pyrosequencing platforms: 46 contigs were assembled for the rat genome and 18 contigs for the mouse genome (16). These contigs were then reassembled using gap filling by Sanger sequencing (Tannock et al., 1984). Computational analysis was performed using the iMetaSys pipeline that uses both Glimmer and MetaGene software (Noguchi et al., 2006). MetaGene was also used to generate Clusters of Orthologous Genes (COGS) and cellular metabolic pathway maps (Noguchi et al., 2006).

The SFB isolated from the mouse host (Japan strain) were found to promote helper -T cell (Th) 17 response in the intestinal lamina propria while a newly sequenced mouse isolated SFB strain from a research group in the Netherlands has shown the ability to induce not only Th17 response but also Th1 and Th2 responses (Bolotin et al., 2014b).

Research suggests that though SFB are host-specific, the microbial communities of the guts of hosts they inhabit are similar (Gong et al., 2007). Host

specificity was suggested after studying the microbiome of broiler chickens using culture-independent methods of 16s rRNA gene sequencing and comparing results with data generated from similar methods from ruminants, pigs, rodents, and humans hosts (Gong et al., 2007). In this study, 16s rRNA gene sequencing revealed a large number of clones identified as *Candidatus* arthromitus with 97-100% similarity using BLASTN nucleotide queries. *Candidatus* arthromitus represent a subline within the *Clostridium* subphylum. Also prevalent were *Lactobacilli* spp. strains with 98 clones ranging from 92-100% similarity using BLASTN nucleotide queries (Gong et al., 2007).

SFB primarily inhabit young animal hosts close to the weaning stage of life also known as the time when offspring no longer rely on parents for food but begin to eat on their own. SFB population can be significantly altered both by diet and age of the animal host (Gong et al., 2007).

2.9 LTS. Light Turkey Syndrome (LTS) is similar to Poult Enteritis Complex/Syndrome (PEC/PES) in that birds experience weight loss however it is dissimilar in that birds do not experience watery and pale intestinal contents, distended ceca, diarrhea, lethargy, and depression as in PEC/PES (Morishita et al., 1992). Both conditions are not related to mortality as can be seen in Poult Enteritis and Mortality Syndrome (PEMS). Typically, LTS/PES affects birds less than three weeks of age (Morishita et al., 1992). A higher number of different pathogenic organisms are found in these younger birds than in birds aged four to nine weeks. In addition to the aforementioned symptoms, other pathohistological

symptoms include the presence of inflammatory infiltrates, plasmocytic, heterophilic, and lymphocytic cells with necrosis and clumping seen in the distal villi tips (Morishita et al., 1992). Virus strains such as astrovirus, reovirus, and rotavirus types were also found, but are not associated with mortality (Morishita et al., 1992). Coronavirus, which is associated with mortality, was not found (Morishita et al., 1992). A higher rate of PES was found in commercial flocks in autumn which is thought to be associated with increased bacterial proliferation towards the end of summer (Danzeisen et al., 2013). Bird weight is highly correlated with intestinal weight (Danzeisen et al., 2013). As identified by Operational Taxonomic Units (OTU's), *Candidatus* arthromitus emergence is thought to occur in weeks 5-7 of the turkey lifetime (Morishita et al., 1992).

Clostridium bartletti and *Lactobacillus avarius* are two strict anaerobes that play a role in the development of the turkey gut microbiome (Danzeisen et al., 2013). There is a possible research bias when studying turkeys, both commercially and in a research setting, since research flocks experienced a shift in microbiome content at a faster rate and earlier on than commercial flocks (Danzeisen et al., 2013). Heavy versus light flocks were compared to assess this difference, and it was seen that heavy research flocks containing the unknown but later identified *Candidatus* arthromitus bacteria made this microbiome switch fastest and earliest (Danzeisen et al., 2013).

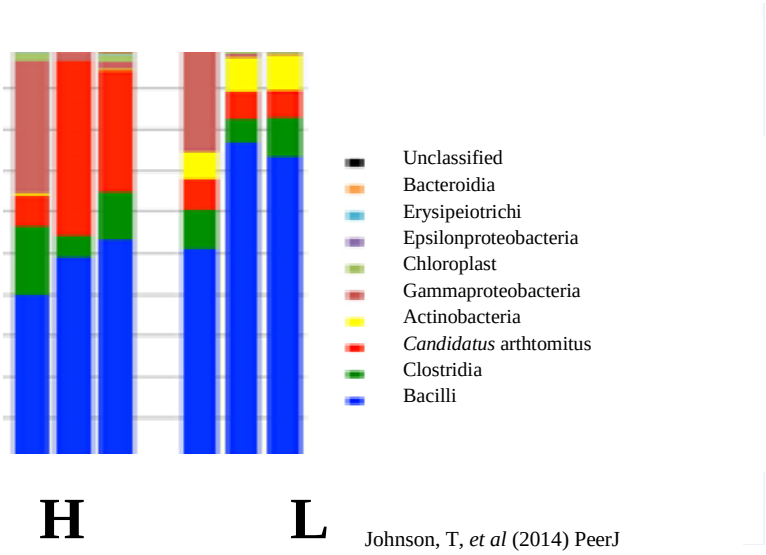

Figure 3: **16S rRNA Amplicon Profiling Results of Heavy (H) and Light (L) Birds**

2.10 Systems Biology and Genomics of Candidatus arthromitus. As explained above, several genomes of *Candidatus* arthromitus have been sequenced using various methods and are publicly available (Sczesnak et al., Pamp et al., 2012, Prakash et al., 2011, Kuwahara et al., 2011). The main aim of this project was to provide an understanding of the metabolic network of *Candidatus* arthromitus strains to allow for optimized culturing medium for CA *in vitro*. After genomes have been sequenced, these genomes were then uploaded to KBase, an online genome-scale metabolic reconstruction pipeline sponsored by the U.S. Department of Energy, and draft genome-scale metabolic models (GEMs) were created using the semi-automated tools provided by KBase that is built with the same bioinformatics pipeline from a ModelSEED (Devoid et al., 2013). In the first step of the ModelSEED pipeline implemented in K-Base, the

assembled genome sequence is annotated by the RAST server (Glass et al., 2010) and imported into the SEED analysis system. Next, a preliminary model is generated consisting of intracellular and transport reactions associated with genes on the basis of RAST annotations, spontaneous reactions and an organism-specific biomass reaction. In the auto-completion step of the pipeline, additional intracellular and transport reactions are added to create an analysis-ready model capable of simulating biomass production using only transportable nutrients. Flux Balance Analysis is then used to generate phenotype predictions in the model analysis step. The final three steps of the pipeline involve the removal and addition of reactions from the model to fit Biolog ™ phenotyping array data (when available) to produce an optimized model. Figure 4 below serves as a visual to emulate the steps outlined above. Figure style was modeled after Figure 1 in Henry et al, 2010 (Henry et al., 2010).

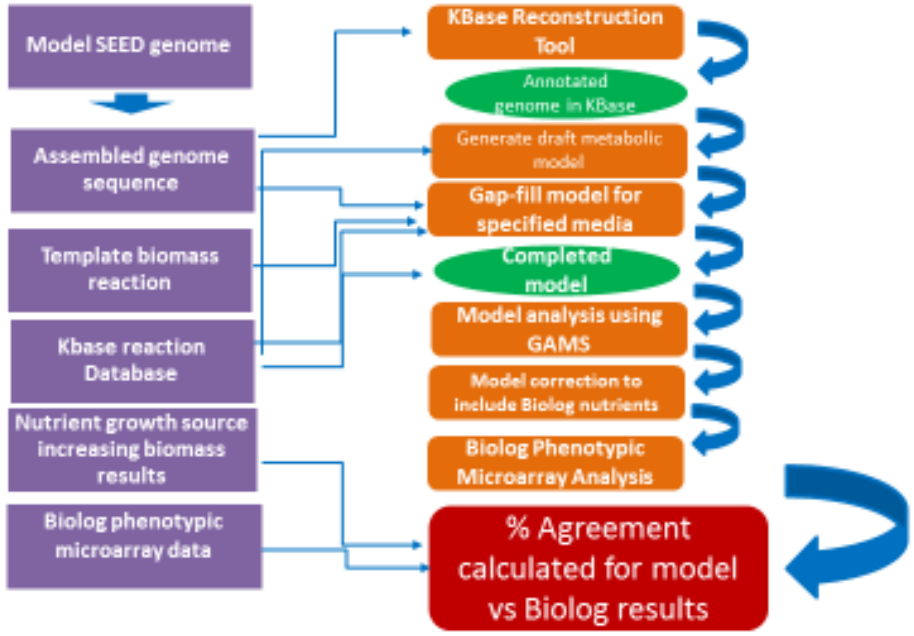

Figure 4: Pipeline outlining construction & validation of genome-scale metabolic models in KBase

These GEMs can be used for metabolic inquiry using the constraint based optimization methods outlined in the COBRA toolbox (Schellenberger et al., 2011) and implemented in optimization software applications like General Algebraic Modeling System (GAMS) which uses algorithms such as Flux Balance Analysis (FBA) to simulate growth conditions utilizing possible nutrients that may be predicted to lead to increased cellular biomass production. In FBA, reactions-related fluxes predicted in a metabolic network are manipulated by constraining fluxes. Flux constraints take the form of altering steady state mass balances,

reaction directionality, and metabolite availability (Ding et al., 2016). An overall objective function determines the flux distribution in order to either maximize or minimize the outcome of this function. For the purpose of this project, the objective function is the biomass production. FBA was conducted in a software platform discussed in the next section of this chapter. This platform utilizes the metabolites (rows) and reactions (columns) to create a stoichiometric matrix (S-matrix) that corresponds to genes in an organism (Ding et al., 2016). See Figure 5 below for an example of how a solution space for a metabolic model is created using a reaction network converted into a stoichiometric matrix.

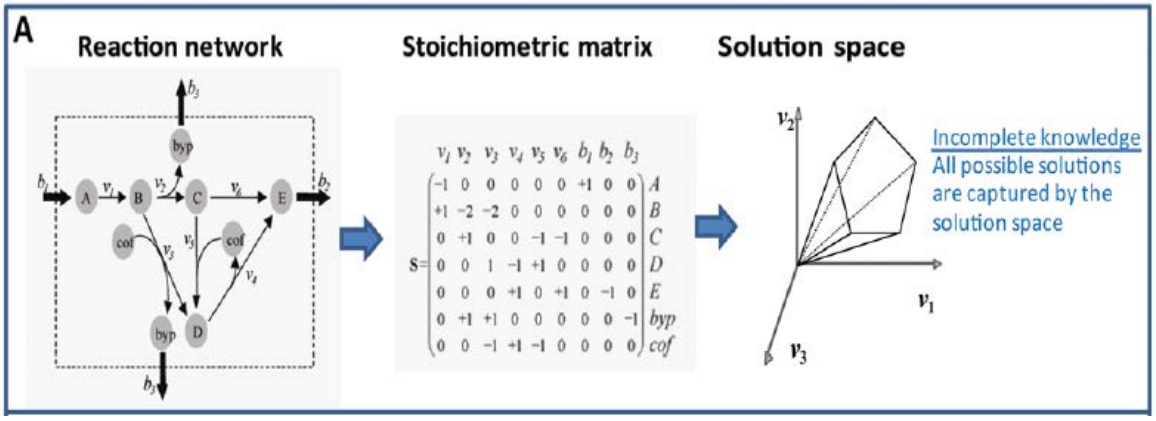

Figure 5: Metabolic model construction: from reaction network to stoichiometric matrix to solution space

Some environmental factors can be constrained to simulate environments such as anaerobic vs aerobic atmosphere however environmental influences such as temperature and mechanical movement while incubating cannot be inferred using these methods. GEMS representing numerous strains of *Candidatus* arthromitus may be used to identify nutrients such as carbon, nitrogen, phosphorous, and iron sources that may lead to increased biomass predictions

for future development of an optimized culture medium for *in vitro* growth of *Candidatus* arthromitus.

2.11 Microbiome Metabolic Models. Several GEMs have been created in order to simulate and/or predict the products metabolized and the interactions between different microbes in a given system. An article that encompasses computational modeling of the entire human microbiome has been published (Borenstein, 2012). The methods focus primarily on topological and constraint-based models (Borenstein, 2012). Topology-based models require a list of enzymatic genes contained within a given organism's genome and a set of biochemical reactions that an organism may catalyze. The topology of the reconstructed pathways is observed and related to metabolic phenotypes. These types of models are simplified versions of what exists in reality, since stoichiometry, rate, and temperature are not taken into consideration. Though these components are missing, large insights into the metabolic capacities of organisms have been made to understand the evolution of pathogenic and non-pathogenic *Escherichia coli* strains (Baumler et al., 2011).

GEMs can use optimization methods such as FBA to define the constraints on the metabolic fluxes of enzymatic reactions that occur in the organism's metabolism. Mass balance, capacity limitations, and reversibility are examples of constraints (Borenstein, 2012). There are four parts to constructing a genome-scale metabolic model: metabolic gene annotation, manual curation and conversion to a mathematical model, experimental validation, and revision

and improvement through computation and experimental cycles (Borenstein, 2012). These mathematical models provide explanation for the linkage between phenotype and genotype of organisms (Greenblum et al., 2012). High throughput "omics" technologies such as metabolomics, metagenomics, metaproteomics, and metatranscriptomics have revealed the state of ecological systems, e.g. the relationship between gut microbiota and its host (Borenstein, 2012). The data coming from these technologies is used in constructing mathematical models found in the field of systems biology in order to understand the mechanisms and importance of different reactions within an organism as well as the importance to a given ecosystem (Borenstein, 2012).

GEMs can be used to examine differentiating metabolic capacities of numerous bacteria, and also used to elucidate phenotype-genotype relationships, biological inquiry, to identification of new control targets for future therapeutic intervention (Baumler et al., 2011, Borenstein, 2012). GEMs can also be used to serve as tools when interpreting global data, since the connectivity enhances the data's statistical power and therefore can reveal novel casual relationships (Greenblum et al., 2012, Borenstein, 2012, Klitgord and Segrè, 2011, Karlsson et al., 2011, Shoaie et al., 2013).

2.12 Intestinal GEMs*.* Numerous GEMs have been constructed in order to generate *in silico* experimental results about microbiome members of the human gut. Until recently, however, the ability to model co-species interactions as well as entire environments had not been attempted. To compliment these *in silico*

predictions, many methods are utilized. *In vitro* cell culture methods, *in vitro* gut models, *ex vivo* organ models, animal models, and human in-patient studies all serve as a basis to check the results of the “omics” modeling networks (Thiele et al., 2013). The first manually curated reconstruction of human metabolism was completed in 2007, entitled Recon 1. Recon 1 encompassed many metabolic functions of any general human cell. (Thiele et al., 2013). Klitgord and Segre put seven published singular reconstructions into an *in silico* environment and investigated co-species interactions cross feeding in several different medium recipes (Klitgord and Segre, 2010). Freilich et al. analyzed the interactions between 118 different microbes’ reconstructions and observed that most of the cooperative interactions were unidirectional meaning that one microbe benefitted while the other remained unaffected (Freilich et al., 2011). A constraint-based model of *Mycobacterium tuberculosis* and human aveolar macrophages was constructed in order to simulate the pathogen-host interactions occurring during intracellular infection. The constraint-based reconstruction and analysis (COBRA) approach was also explored in order to create host-microbe models as well as microbe-microbe models (Schellenberger et al., 2011). This particular approach was used to successfully determine metabolic dependency between a commensal gut microbe and a murine host; mutually beneficial cross-feeding between host and microbe *Bacteroides thetaiotaomicron* were observed and related to host diet (Heinken et al., 2013).

One study focusing on carbohydrate metabolism summarized the bacterial flora in the gut into three phyla: *Bacteroidetes*, *Actinobacteria*, and *Firmicutes*. Of the three, *Bacteroidetes* is Gram-negative while the latter two are Gram-positive (Musso et al., 2011). Cross-feeding between host and members of the microbiome occurs with 64 total enzymes encoding 87 different metabolic reactions. Of those reactions, seven are exclusive to the host and 27 are exclusive to the microbial flora. This work demonstrates the definitive host-interaction and warrants the importance of creating GEMs to determine the relationship between the microbiome and host-interactions in reference to health status and onset of disease.

Of the three phyla, a higher proportion of *Firmicutes* is associated with higher fat and carbohydrate consumption (Ibrahim and Anishetty, 2012). Inflated *Firmicutes* presence amplifies the number of enzymes responsible for carbohydrate breakdown (Ibrahim and Anishetty, 2012). Several metabolic pathways are therefore enhanced including: the phosphotransferase system (PTS), fructose and mannose metabolism, glycolysis and gluconeogenesis cycles. As these pathways are enhanced, more overall carbohydrate is metabolized in comparison to *Bacteroidetes*, which contains bacteria such as *B. thetaiotaomicron* that lack a complete PTS (Ibrahim and Anishetty, 2012).

Ibrahim *et al.* (Ibrahim and Anishetty, 2012) also describes lactose metabolism within the metabolome network of carbohydrate metabolism. By members of the phyla *Firmicutes*, lactose is broken down into tagatose which

may have an anti-hyperglycemic effect which may be effective in controlling type 2 diabetes (Lu et al., 2008). Tagatose is then converted to short-chain fatty acids (SCFA) by microbes in the gut to aid in bodily absorption (Bertelsen et al., 2011).

Additionally, there are several enzymatic reactions occurring directly from the host that mirror reactions carried out by gut microbes. These reactions primarily produce SCFA and the merriment of both host and microbiome reactions escalates the production of these SCFA (Ibrahim and Anishetty, 2012). Although an excess in carbohydrate metabolism has been linked to autoimmune diseases like obesity and diabetes, a lack in the ability to metabolize carbohydrates also is linked to other ailments such as Fabry disease (Ibrahim and Anishetty, 2012). Enzyme alpha galactosidase is responsible for the catalysis of various carbohydrate metabolic reactions. Without this enzyme, many of those reactions would not occur. 24 total bacteria in the three phyla mentioned above contain this enzyme and may be used as a probiotic in individuals lacking this enzyme to aid in carbohydrate digestion (Ibrahim and Anishetty, 2012). Overall, Ibrahim *et al.* (Ibrahim and Anishetty, 2012) provided a network that encompasses both host and microbe enzymatic pathways to coincide with the systems biology top-down approach. The potential for this type of analysis may benefit humans by generating new solutions to aiding the metabolic process (Ibrahim and Anishetty, 2012).

Krishnan *et al*. explains the necessity of the gut microbiome in order to mediate interactions and connect the intestines with the rest of the body,

especially the liver and brain (Krishnan et al., 2015). To compliment, the term dysbiosis was defined and studied by the Gordon Laboratory. Dysbiosis is the state of having a greater capacity for energy storage (Turnbaugh et al., 2006). To understand the gut microbiome and its interactions with the host, a dual intersecting reaction network was constructed with added search methods to find aromatic amino acid catabolism (AAA) and differentiate those reactions between microbial community and host (Sridharan et al., 2014). Through this host-microbiome interaction network in combination with experimental studies, it was found the serotonin levels in germ free (GF) mice is lower than when *Clostridium* species populate the GI tract (Krishnan et al., 2015). Also, specifically *Clostridium scindens* aids in the dehydroxylation step for producing secondary bile acids as well as acting as an inhibitory element in the colonization of *Clostridium difficle* in a bile acids dependent manner (Krishnan et al., 2015). Recently an analysis conducted at the Fischbach laboratory revealed over 3,000 small molecule biosynthetic gene clusters specifically in bacterial genomes associated with humans (Krishnan et al., 2015, Donia et al., 2014). This emphasizes the need to continue exploration in the meta-metabolome of the gut microbiome.

Greenblum *et al.* highlights the importance of understanding co-species microbial interactions and can be studied through co-culture growth assays (Greenblum et al., 2013). The coupling of these studies along with phylogenetic information shows the occurrence of horizontal gene transfer and intestinal

microbe-microbe interactions (Greenblum et al., 2013). Two methods were used in order to account for the species' habitat and metabolic capacity. A graph theory-based algorithm is used to determine metabolites from the environment which serves as a model for the biological environment of the species. A network expansion algorithm is then used to determine additional external metabolites utilized and produced under given conditions. Coupled together, these two methods serve as a predictive system to determine the outcome of metabolic competition. To determine its effectiveness, this dual method framework was attempted on literature-cited co-occurring species with results that showed higher competition results in lower mean growth (Greenblum et al., 2013).

In addition, an extension was added to quantify metabolic capacity two organisms achieves through cooperation termed metabolic synergy. By determining metabolic synergy, It was determined that to obtain an optimal cooperative network the organisms needed to be neither too similar nor dissimilar (Greenblum et al., 2013). In order to determine the effects of the entire microbiome on the host, Greenblum *et al.* state that boundaries between species need to be ignored to create a "supra-organism" in which pathways function at the community level (Greenblum et al., 2013). Emulating the "supra-organismal" approach, the gut microbiome was examined to determine its role in obesity and irritable bowel disease (IBD). Shotgun genomic data was used for reconstruction of a community-level metabolic network as well as the placement determination of disease-associated enzymatic genes. Enzymatic reactions pertaining to host

health tend to reside on the outskirts of the network (Greenblum et al., 2013). These results show that different microbiome compositions from a medical standpoint are associated with systems-level organization (Greenblum et al., 2013). Constraint-based modeling has been used to study host-gut microbe interactions singularly. In order to analyze changes dependent of dietary differences, secretions and uptakes were monitored for change and successfully predicted growth dependence and cross feeding (Greenblum et al., 2013).

2.13 Conclusion*.* The gut has become more and more of a focal point when accessing the prevalence and cause of disease for both vertebrate animals and humans. As those in the medical field begin an internal approach to curing disease, additional knowledge on specific members of the gut microbiome is needed. *Candidatus* arthromitus was formerly an unculturable gut microbe however there are new methods to both keep cells viable (Schnupf et al., 2015) and grow cells *in vitro*.

For *in vitro* growth of *Candidatus* arthromitus , bacterial isolation was performed using several methods (Schnupf et al., 2015, Pamp et al., 2012, Snel et al., 1995, Ericsson et al., 2014b). Only one method achieved viability outside of the host. The use of Nycodenz bacterial layer isolation by density with centrifugation steps and filtration through 70 micrometer mesh filters allowed the isolated cells to maintain viability once placed into culture medium (Schnupf et al., 2015). The experiments that led to determining the media recipe recently used (Schnupf et al., 2015) are not described however the use of trans-well

plates and live cells for growth offered the most optimal results. Explanation of how each media ingredient was determined as a significant ingredient to enhance growth in the culture medium would allow others to potentially replicate experiments. However little experimental details were provided, and attempts to replicate *in vitro* growth were not successful in this project. Host independent viability of *Candidatus* arthromitus was successful using the liquid media recipe recently described (Schnupf et al., 2015).

Our lab is working with avian hosts (turkeys) in order to determine host-independent viability and to develop *in-vitro* culturing methods. The rationale for studying *Candidatus* arthromitus from turkeys as a host stems from the prevalence of LTS in turkeys as previously discussed. If successful cultivation of *Candidatus* arthromitus from turkeys is accomplished, future studies involving poult inoculation will determine whether the presence or absence of CA affects weight of full-grown birds. This would serve to elucidate the observations of the turkey gastrointestinal metagenome made previously (Danzeisen et al., 2013).

3. Comparison of five *Candidatus* arthromitus genomes isolated from mouse, rat, and turkey hosts

3.1 Summary. According to Sczesnak *et al*, SFB are not closely related to any other cluster of bacterial microorganisms at the genomic level and occupy a void space in between minimal relatives such as other *Clostridia* and *Mycoplasma* species (Sczesnak et al.). This provides reason to explore deeper into the genomes of SFB from multiple hosts in order to determine likeness between the genomes as CA is a host-specific and publications based on successes isolating and culturing CA may not be transferrable. The comparison of uniqueness between strains with regard to predicted gene function and metabolic subsystem analysis will be determined in the upcoming section.

3.2 Introduction. RAST is a Rapid Automated Subsystems Technology tool for the annotation of genome contents that can be used to determine similarities and differences at the genomic level. More specifically, the RAST files used and compared for this project were all *Candidatus* arthromitus genomes isolated from three different hosts. With the exception of Mouse and Rat YIT genomes, all other genomes were from different labs. Also discussed in this section are the phylogenetic similarities and differences between CA genomes and between CA genomes, *Bacilli* spp. and *Clostridium botulinum*.

3.3 RAST Metagenomics Functions Results Analysis. Genomes of the 5 CA strains (Sczesnak et al., Pamp et al., 2012, Prakash et al., 2011, Kuwahara et al., 2011) were analyzed using the RAST Analysis Server to generate annotated

genomes to determine similarities and differences at the genetic level. Using Microsoft excel and access, RAST files were used to determine the total number of genes associated with key subsystem functional groups. The key functional groups analyzed were: transporters, permeases, proteases, peptidases, phosphotransferases, catalases, and peroxidases. Appendix A found on page 120 contains the data for all five genomes side by side for comparison. Figure 6 displays a vertical stacked column graph to offer a visual of the information contained in Appendix A. Of these five genomes, only the CA turkey genome contains a gene in the peroxidase category. Overall, the CA Mse_Japan genome contained more genes in each category except for the transporter category. Table 1 shows the relative uniqueness of genomes based on predicted function. The CA turkey genome contained the highest number of unique genes for permeases and transporter genes. All genomes contained no unique catalase genes as they all had the same single gene associated with this function. The CA Mse_NL genome had the most predicted genes encoding proteases at six total genes. Both the CA turkey and CA Rat_YIT genomes contained the highest count of strain-specific unique genes at three genes each. Table 2 shows the overall uniqueness of each genome based on number of unique genes for these combined functional categories. Overall, the CA turkey genome contains 31 unique genes in the selected analyzed functional categories which is the most overall unique gene count out of all five CA genomes. CA Mse_NL and CA Rat_YIT genomes follow with 19 and 14 unique genes within

the analyzed function categories. CA Mse_YIT and CA Mse_Japan genomes contained the fewest unique genes in the analyzed function categories with five and one single unique gene.

Table 1: Comparison of predicted gene functions for 5 genomes of *C.* arthromitus

Genome	No. Transporters	No. Permeases	No. Peptidases	No. Proteases	No. Catalases	No. Phosphotransferases	No. Peroxidases	Total
CA_Turkey	6	7	3	4	0	1	1	22
CA_Rat_YIT	5	3	3	0	0	0	0	11
CA_Mse_YIT	3	0	0	1	0	0	0	4
CA_Mse_NL	0	5	1	6	0	0	0	12
CA_Mse_Japan	1	0	0	0	0	0	0	1

Table 2: Number of Predicted Genes unique to each strain of *C.* arthromitus *classified by* Functional groups

Genome	No. Transporters	No. Permeases	No. Peptidases	No. Proteases	No. Catalases	No. Phosphotransferases	No. Peroxidases
CA_Turkey	68	20	33	24	1	4	1
CA_Rat_YIT	65	24	34	21	1	5	0
CA_Mse_YIT	68	22	32	23	1	5	0
CA_Mse_NL	64	29	36	31	1	5	0
CA_Mse_Japan	61	18	33	21	1	5	0

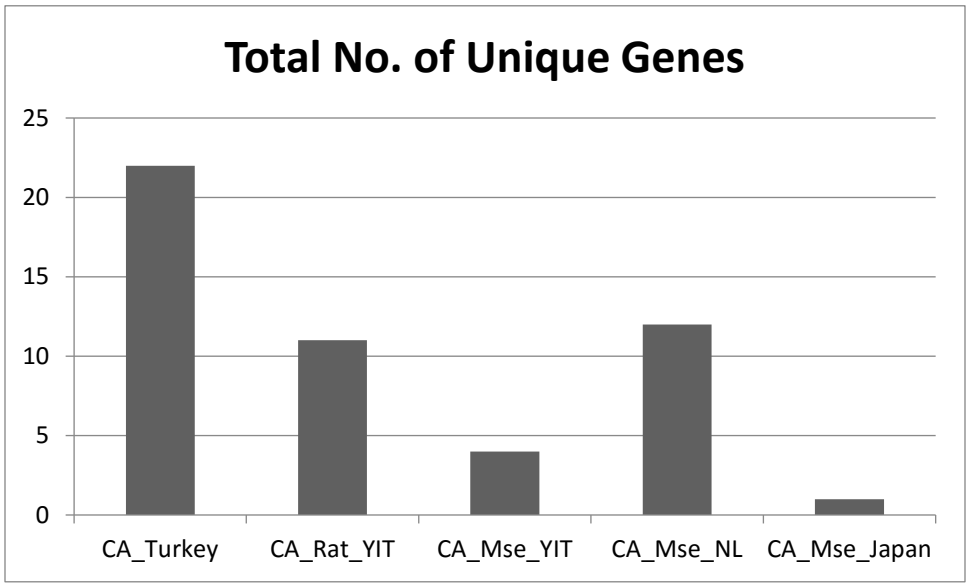

Figure 6: **Total Number of Unique Genes in *C.* arthromitus *g*enomes for Selected Subsystem Catagories**

3.4 RAST Subsystems Analysis*.* The purpose for analyzing the selected gene function and subsystem categories pertains to the analysis of an SFB genome completed by Sczesnak, *et al (Sczesnak et al.)* which compared a genome sequenced from mouse feces to other members of the *Clostridia* family. In comparison to Sczesnak's results, we also found 38 tRNAs in the CA Mse and Rat_YIT genomes, 37 in the CA Mse_Japan genome, and zero in both the CA turkey and CA Mse_NL genomes. Genes encoding CRISPRs were found in all five genomes. The CA turkey genome contained five genes associated with CRISPRs going from genome coordinates 469-2709, 6711-7739, 7754-8044, 2727-3404, 3401-5185 in the genomic sequence. The remaining genomes contained either five or six genes in the CRISPR subsystem located at varying coordinates within the genome. The CA turkey genome did contain a single gene associated with phage tail synthesis;

however, there was no presence of a phage operon in the genome. There was no phage operon found in either the CA Rat_YIT or the CA turkey genomes but all three CA mouse genomes (YIT, NL, and Japan) contained operons associated with genes involved in phage synthesis. It was also found that in the CA turkey genome, no genes are contained in a subsystem found to break down Pantothenate from vitamin B5 but six enzyme CoA biosynthesis genes were found along with a CoA disulfide thiol-disulfide redox system gene at coordinates 76742-79054 that does not appear to be in any of the other analyzed CA genomes. ORFs for V-type ATPase were found in all five analyzed genomes including the turkey genome. Four genomes contained eight ORFs associated with V-type ATPase while the CA Mse_Japan genome contained only seven genes in this subsystem category. All genomes were missing a proton transporter for a gradient to produce ATP discussed *(Sczesnak et al.).* Also mentioned in the work of Sczesnak *et al,* are genes within a flagella synthesis subsystem that may utilize the V-type ATPase gradient. The CA turkey genome contains more genes in a flagella synthesis subsystem than the remaining four analyzed genomes which according to Sczesnak *et al,* dependent on a concentration gradient (V-type ATPase gradient) in order to function. Finally, to compliment the findings by Sczesnak *et al* once more, the CA turkey genome contained a comparable number of transporter genes which demonstrates that it also has many

transporter pathways in comparison to other auxotrophic organisms. Subsystem results can be found in Appendix B on page 100.

3.5. Construction of a Phylogenetic Tree Using Sequences from C. arthromitus and a Clostridia Out-Group Aligned in Mauve. Sequenced genomes aligned in using the comparative analysis tool Mauve were loaded into the project space and SNPs were exported using the SNP export function under tools in the Mauve alignment project space (Darling et al., 2004). SNP files were saved and opened as a .txt file for further modification using the general command line. To open the PC's general command line, the start tab was selected and in the search bar "cmd" was entered. A black icon with cmd appeared and it was this icon that was selected for use in modifying the .txt file containing the SNP information. The following functions were performed to modify the file: cut –f1 (File.txt)>(Filenew.txt)**,** grep –v "N" (Filenew.txt)>(Filenewer.txt)**,** grep –v "-" (Filenewer.txt)>(Filenewest.txt)**.** This will delete all but the first column of the SNPs file along with any "#NAME?" and "---*--" rows. The modified file was saved and imported into excel with cells set to fixed width under the data to columns tab. The excel file contains a single column containing many rows of letters. The number of letters in each row is the same and corresponds to the number of genomes in the alignment the information was exported from. To modify the file in MS excel, each letter was placed in its own column with the genome name placed in the first cell of each column. A ">" sign was directly placed before the name and no spaces were in each genome name.

The excel file was saved in order to be converted into a .fasta file using Sublime text editor. Sublime text was opened and a new file was created. Each genome was copied and pasted into the new sublime text file, one on top of the other so that a singular vertical column was created. This file was saved to the computer's desktop as a .fasta file by saving as file.fasta in the file's name. The .fasta file was then opened in Mega software by selecting to open a new file/session and selecting "analyze" (Kumar et al., 2008). The phylogeny tab was selected after analyzing and the "Maximum Likelihood" option was chosen to construct the phylogenetic tree. Bootstrap values were displayed.

The phylogenetic tree constructed from the SNPs extracted from a six-genome alignment in Mauve showed CA genomes Mouse_YIT, Mouse_NL, and Mouse_Japan are most related (Figure 7). The phylogenetic distance between each of these genomes on the Maximum Likelihood rooted tree is 0.0 meaning that these three genomes are closest in predicted relation. The CA Rat_YIT genome was a phylogenetic distance of 0.2 away from CA Mouse_YIT, Mouse_NL, and Mouse_Japan making the relationship between the three mouse genomes and the one rat genome close but not as close as the relationship between the three mouse genomes themselves. The CA Turkey genome was a phylogenetic distance of 0.2 away from the CA Rat_YIT genome and 0.3 in phylogenetic distance away from the three CA mouse genomes. The out group, *C. Botulinum*, was a phylogenetic distance of 0.6 away from the three CA mouse

genomes, 0.3 away from the CA Turkey genome, and 0.4 away from the CA Rat_YIT genome.

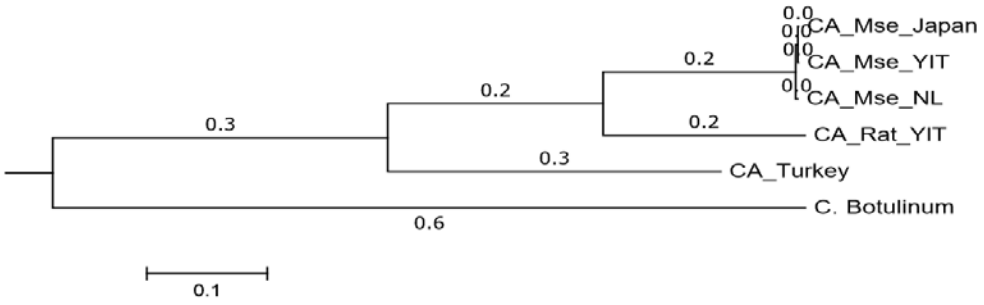

Figure 7: **Phylogenetic Tree Relating CA genomes to *C. botulinum***

3.6 Phylogenetics Using Muscle: An Alternative Multiple Alignment Source for 16S Region Alignment. 5 CA 16S rRNA gene regions were aligned along with a *Bacillus cereus* and a *Bacillus anthracis* 16S sequence. The purpose for aligning the five studied CA sequences with these Bacilli sequences was to determine the reasoning behind obtaining *Bacillus* matching hits while using NCBI's Nucleotide Basic Local Alignment Tool (NBLAST) (Dumontier and Hogue, 2002). The amplicon from our CA turkey genome was in the 16S rRNA region too, explaining the overall comparison of 16S rRNA regions for all organisms. When performing standard PCR, two sets of primers were used. One set termed the Universal primers (F: CCGCCTGAGGAGTACGA, R: GGTAATTCAAGGTATGTCAAGTTTAGG) was designed to amplify the entire 16S rRNA region instead of a CA-specific region16S. The purpose for doing so was to determine the purity of the culture, potentially identify other contaminating

organisms within the culture, and for molecular determination and verification of CA presence. There was no success in identifying CA using universal primers in this project which originally led us to believe that the culture had several organisms growing even though the morphology observed while Gram staining did not agree. In order to understand the reasoning behind these sequences being returned in NBLAST, we constructed a multiple 16S rRNA region alignment in Muscle (Edgar, 2004) to determine the phylogenetic distance between our CA strains and the two most commonly seen organisms returned using NBLAST with our configured sequences from PCR conducted with our universal primers.

Margulis *et al (Margulis et al., 1998)* explains the likeness of soil *Bacillus* species and CA. Although arthropods all harbor one type of arthromatid (*Lachnospiraceae*) and vertebrates harbor another, avian hosts harbor a third type of arthromatid that is less relatable to mammalian SFB and more relatable to *Bacillus* species such as *B. cereus* and *B. anthracis* (Margulis et al., 1998). Figure 8 contains the phylogenetic tree with branch lengths generated from the multiple alignment completed in Muscle (Edgar, 2004) which serves as supportive evidence that the CA turkey strain we study is more closely related to the soiled *Bacilli* species studied by Margulis *et al* (Margulis et al., 1998). An alternative conclusion could be that the amplified organism was not truly CA but instead was a contaminant within the co-culture.

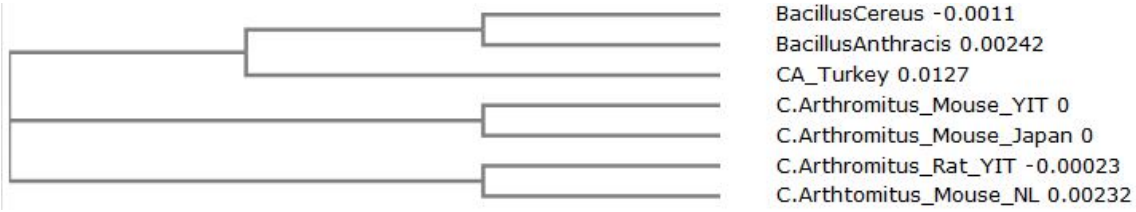

Figure 8: Phylogenetic Tree Relating CA 16s rRNA genes to *B. cereus* and *B. anthracis*

3.7 Conclusion. The RAST selected functions results demonstrated areas of uniqueness for each genome as well as overall uniqueness. Based on these results, it is clear that the Turkey genome has the most unique genes within the selected subsections. The turkey genome also has a gene associated with peroxidase while none of the other CA genomes do. The RAST subsystems results followed the works of Sczesnak *et al* supporting the idea of CA from turkeys being of auxotrophic nature (*Sczesnak et al.)*. Phylogenetic results comparing CA genomes with *C. botulinum* as an out-group show that *C. botulinum* is less related to CA genomes and that CA genomes are more related to one another. Finally, phylogenetic results comparing CA genomes with *B. cereus* and *B. anthracis* shows that CA from an avian host (turkey) are more related to these soil-dwelling bacilli than to the other genomes of CA from mammals (mouse and rat). The importance of these findings further support Sczesnak *et al.* in determining the relatedness of CA from avian hosts to soil bacilli. These findings also support the flow of the phylogenetic tree in Figure 7 as the avian host derived CA from both chickens and turkeys would be expected to reside on close branches.

4. New Culture Methods Optimizing Growth of *Candidatus* arthromitus.

4.1 Summary. The contents of this chapter describe the materials used and methods followed to conduct the described research on CA and provide the contents of the results section. This process began with *in silico* simulations and genomic comparisons for determination of genetic uniqueness and potential nutrient utilization. Built on these methods and results were the wet lab experiments of isolation and extraction of CA from turkey ileums, molecular verification for identity and growth, and visual determination. Unsuccessful methods are also documented in this section to serve as documentation for what has been attempted experimentally.

4.2 Introduction. Initial meta-genome information was provided by Dr. Timothy Johnson's lab. This information along with the prevalence of LTS served as the basis to begin research on CA from turkeys. The genome from avian host-residing (turkey) CA obtained from Dr. Johnson's lab enabled the genomic comparisons and the GEM creation for metabolic predictions and *in silico* analysis. Turkey Ileum samples were provided by Dr. Sally Noll from poults. Bacterial isolation is described below, along with purification, culturing techniques, and experiments serving as supporting evidence of viability and growth.

4.3 Materials and Methods.

4.3.1 Genome Shotgun, Extraction, and Comparison. Bacterial genome extraction and shotgun on CA from turkey was completed by Dr. Timothy Johnson and colleagues (Danzeisen et al., 2013). Extraction and ordering of the genome was achieved by using a metagenome shotgun technique of the entire intestinal microbiome. The results were then utilized for further research in Dr. David Baumler's laboratory where the draft turkey *Candidatus* athromitus genome was compared using the *Mauve* multiple genome alignment tool to align this new draft genome with four other published *Candidatus* arthromitus genomes (Darling et al., 2004). Aligned genomes came from four locations. Mouse-YIT and rat-YIT strains were sequenced from the same lab at the University of California at Berkley (Prakash et al., 2011). The Mouse-Japan strain was sequenced as a joint project between several educational institutions in Japan (Kuwahara et al., 2011). The SFB-Mouse-NL strain was sequenced and originated from the Netherlands, supported by the French National Research Agency (FNRA) (Bolotin et al., 2014a). The last strain was derived from turkeys, and was sequenced at the University of Minnesota in Dr. Timothy Johnson's lab (Danzeisen et al., 2013). The first alignment (Figure 10) also contained a published *Clostridium botulinum* genome, which was included to show gene comparisons between CA and a phylogenetically close relative. The alignment found in Figure 11 displays the five CA genomes aligned without the *Clostridium botulinum* genome. Figure 12 displays the backbone view of the 5 CA strains.

4.3.2 Genome-Scale Metabolic Modeling. Genomes were uploaded to KBase, an online database sponsored by the U.S. Department of Energy (Ranjan, 2015), and draft genome-scale metabolic models (GEMs) were created using the semi-automated tools provided by KBase. Using KBase, gap-filling was then performed on the draft models to identify missing metabolic reactions for D-glucose catabolism under both aerobic and anaerobic conditions. Then the draft models were downloaded in systems biology markup language (SBML) file format, and these SBML files were converted to General Algebraic Modeling System (GAMS) file format to enable their utilization by GAMS software. Five GEMs were created using KBase; one for each CA genome: Mouse_NL, Mouse_YIT, Mouse_Japan, Rat_YIT, and Turkey CA strains.

4.3.3 Using Gems for Metabolic Predictions. Constructed GEMs were used by General Algebraic Modeling System (GAMS) software. Flux balance analysis was conducted to optimize the biomass produced and to screen for nutrient utilization that may enhance growth. Flux balance analysis was completed under both aerobic and anaerobic environments to determine the sole carbon sources predicted to be utilized by each strain, the compounds secreted by strain when grown on solely on D-glucose, and the nutrients that may increase biomass production of each strain.

4.3.4 Gut Extraction I. Five week-old poults were humanely gassed with CO_2 gas in a secure chamber for 10 min. Subjects were determined dead by looking for any optical activity or respiratory action. Subjects were laid on a sterile

surface and systematically cut open underneath the rib cage. The dermal layers were pulled down and away from the gut and the thin membrane layer covering the gastrointestinal system was removed carefully by making a small incision in the center near the rib cage and pulling down and apart. The intestines were removed from the body cavity by making incisions near the stomach and the rectum. The ileal portion of the intestine was removed by cutting the intestinal tract at the ileocecal junction and the jejunum-ileum junction. Ileum samples were placed into sterile whirl-pak bags, labeled and placed into 4°C storage until gut contents were harvested from ileum samples. Gut extraction for gut contents was performed as in Gut Extraction II.

4.3.5 Gut Extraction II. Ileum samples of young turkeys were obtained from Willmar Farm in Willmar, MN. Samples were obtained on five separate collection dates associated with five different ages of birds ranging from 0-14 days old. Ileum samples were extracted for intestinal contents after each collection date. In order to extract gut contents, 1.2 mL centrifuge tubes, 6-well plates, Dulbecco's Modified Eagle Medium (DMEM), cell scrapers, 25 mL pipettes, an autopipetter, aluminum foil, 70% concentrated ethanol, paper towels, and a working HEPA filtered exhaust hood were used. Aluminum foil was placed onto the bottom of the experimental hood as a working surface. Ileum samples contained in Whirl-Pak bags were marked with collection number, date, bird number, and cage number. All information was transferred onto 1.2 mL centrifuge tubes and 6-well plates. Gut extractions were done individually, one ileum sample at a time. Each

sample was taken from its bag, placed onto the foil surface and extracted using gloved hands. Using a pulling motion towards either opening of the sample aided in extracting contents without tearing the intestinal tract. The contents were placed into the corresponding 1.2 mL centrifuge tube by using a cell scraper to lift contents off of foil and transfer them into the tube. Foil and cell scraper were wiped clean with 70% ethanol in between samples. After samples had all been extracted and placed into labeled 1.2 mL tubes, each well of a 6-well plate was filled with ~5 mL of DMEM using a 25 mL pipette and autopipetter. The contents of each 1.2 mL tube were placed into the corresponding well in a 6-well plate filled with DMEM and mixed gently. Samples were incubated at 42°C in an anaerobic tent until further use. Following incubation, samples were removed from anaerobic conditions and placed back in to the hood. Samples were transferred from 6-well plates into glass test tubes with added DMEM and put back into 42°C incubators in anaerobic tent until further use.

4.3.6 SFB Sample Storage. Samples in test tubes were removed from 42°C anaerobic conditions and placed into the anaerobic hood for separation. Each test tube contained 2 samples from a 6-well plate from the same collection date. Test tube contents were filtered through 100 µm mesh into 15 mL conical tubes. 15 mL tubes were then incubated at 42°C in the anaerobic tent until further separation was done.

4.3.7 SFB culture media and technique. The culture media for *Candidatus* arthromitus is comparative to other types of media used to culture fastidious

organisms. CA media consists of: Caco-2 cells (TC7), grown in DMEM/F-12 advanced media with 10% fetal calf serum (FCS). DMEM/F-12 advanced medium with 2% FCS added, 1:100 dilution BHI prepared [BD Difco 237500] with10% peptone/yeast and 5% casein amino acids; 200 mM ribose/cellobiose/mannose , 10 mM 1:1,000 dilution ferrous sulfate, 12.5 mM ferric ammonium citrate, 30mg/mL 1:10,000 dilution retinoic acid in DMSO, 10mg/mL 1:500 dilution sperm DNA in diethylpyrocarbonate (DEPC) digested for 1 h with 10 µm DNase I at 37°C and heat inactivated at 75°C for 30 min, 10 mg/mL RNA undigested, and 0.2% yeast extract (1:100 dilution from 20%). Media was pre-equilibrated in an anaerobic tent prior to use. Frozen Caco-2 cells at passage 3 stage were obtained from Dr. Roy's laboratory (Moos Tower, Minneapolis, MN) and grown aseptically in separate laboratory facility. Caco-2 cells were grown in tissue-culture grade 10 cm petri dishes, then passaged and transferred into trans-well plates with a 0.4 µm membrane. Pure CA cells in CA media were to be layered onto Caco-2 cells in trans-well plates and placed into 42°C incubators in an anaerobic tent set to 1-4% oxygen content however additional steps to combine culture with tissue cells were not taken do to the findings of additional working media not requiring the merriment of live cells with media.

4.3.8 CA Cell Extraction from Samples and Purification. CA cells contained in DMEM/F12 media were taken from 42°C incubators and aseptically transferred to 50 mL sterile Falcon tubes and centrifuged at 8,000 g for 5 min. Pellets were

re-suspended in phosphate buffered saline (PBS) buffer in sterile test tubes and layered onto 3 mL 50% and 2 mL 30 % Nycodenz solution made with PBS in 15 mL conical tubes. Layered solution was centrifuged at 4,000 g for 10 min. The 30% fraction was collected, which is a milky white color and contains bacteria, and diluted in PBS in 15 mL falcon tubes. Bacteria were centrifuged at 8,500 g for 10 min and then re-suspended again in 15 mL PBS that was pre-equilibrated in anaerobic conditions. Pellets were then broken up once more, homogenized by vortexing and strained through 70 µM cell strainers. Cell strainers were scraped using sterile cell scrapers and large cells were then placed into pre-equilibrated SFB media. Samples were placed into 42°C incubators in anaerobic conditions to proliferate. 100 µL of large cells in SFB media were placed into columns A-G on a 96 well-plate and 10^{-1} to 10^{-13} dilutions were performed for each row of the plate. The well-plates were incubated at 42°C in anaerobic conditions for 7 days to achieve the first round of dilution to extinction. Optical Density (OD) readings were taken with a Biotek® plate reading device every 24 h at 630 nm and recorded for comparison. The most diluted sample with the highest OD reading was then used for the next dilution to extinction. This method for dilution to extinction was carried out sequentially three times to obtain a CA cell culture. Upon completion of the third dilution to extinction experiment, the most diluted sample with the highest optical density was spread plated out on Biolog Universal Growth + Blood (BUG +B) media and placed into 42°C incubators in anaerobic conditions for 24-72 h to obtain colony growth. Further

dilutions using the serial dilution method were plated and incubated using the same parameters described above.

4.3.9 BUG + B Liquid Media. BUG + B liquid media was prepared similar to original BUG + B agar media however the first step for preparation of BUG + B media is to fill to volume (950 mL) with cold DDW and to place into the cooler for ~1h. The agar in the media settles out on the bottom and the liquid can then be decanted off into another flask. This liquid is then heated and adjusted for pH. The adjusted media is then autoclaved at 121°C for 20 min. The media is placed into a water bath until it reaches 45°C. 50 mL of defibrillated sheep's blood is aseptically added. This media remains in the flask and is stored at 4°C.

4.3.10 Preparing BUG Media Agar Plates. Preparing BUG agar plates is similar to preparing BUG + B media. Instead of using 950 mL of DDW, 1000 mL of DDW is used. The media is heated to dissolve media contents and the pH is adjusted. Autoclaving is done at 121°C for 20 min. The media is placed in a water bath to cool down to 50-55°C and then plates are aseptically poured in a hood with lights and vent turned on. Plates sit in the hood with the vent on for ~24h. After 24h, plates are placed back into bags and stored at 4°C.

4.3.11 Preparing BUG Liquid Media. Preparing BUG liquid media is similar to preparing BUG + B liquid media. BUG powder media is mixed into 1000 mL of DDW and then allowed to sit at 4°C for ~1h. The liquid is then decanted off into a new flask which then is heated to dissolve media and the pH is adjusted. Media

is then autoclaved at 121°C for 20 min. No defibrillated sheep's blood is added to this media. Media is stored at 4°C.

4.3.12 Growth on Solid Media. BUG +B media is a universal growth source known to grow a wide variety of organisms. To prepare media, 37 g of BUG media powder is mixed into 950 mL of distilled water while stirring on a heated stir plate until all materials were sufficiently dissolved. The solution was then autoclaved at 121°C for 30 minutes to sterilize and cooled to 50-55°C in a water bath. 50 mL of defibrillated sheep's blood was added to the cooled solution and then poured onto sterile petri plates. Plates were left in laminar flow hood with blower on overnight.

Two methods were utilized when preparing for growth on solid media: the three-phase streaking method and the spread plate method. The spread plate method was also used for spreading the diluent onto plates when serial dilutions were carried out in tubes prior to plating. Serial dilutions were performed to 10^{-13} in order to ensure a pure culture. This process was carried out three times in a row with the most dilute plate being the plate harvested from and subsequently plated. Completing this serial dilution to extinction method three-fold is necessary for trying to create a pure bacterial culture. The three-phase streak method was also used to isolate single colonies for visual identification (i.e. Gram stain) followed by re-streaking.

4.3.13 DNA isolation. Genomic DNA was extracted using the MoBio Power Soil DNA Extraction Kit and used to perform PCR using SFB specific forward primer

(CCGCCTGAGGAGTACGA) and universal reverse primer (GGTAATTCAAGGTATGTCAAGTTTAGG). Gel Electrophoresis was done to confirm the presence of PCR products. PCR products were then sent to the Integrated DNA Technologies (IDT) center for Sanger Sequencing. Sequencing was analyzed by using a Minnesota Supercomputing Institute (MSI) account and the Seqman Pro Database to configure contigs from the raw sequence files. Contig files were saved in FASTA file format and compared to all available genomes using the National Center for Biotechnology Information's (NCBI) nucleotide Basic Local Alignment Search Tool (BLAST) (Dumontier and Hogue, 2002).

4.3.14 Determining Optimal Bands on Agarose Gel from PCR Products. In order to determine optimal bands on an agarose gel from PCR products, two DNA ladders must be run on either side of the gel. For this project, a 1 Kb and a 100 base pair ladder were use on either side of the gel to determine the relative size of the PCR produced DNA fragments in the gel.

4.3.15 Determining SFB Visual Characteristics.

Visual SFB characteristics were determined via Gram stain. Figure 9 shows images of both differentiated and undifferentiated organisms.

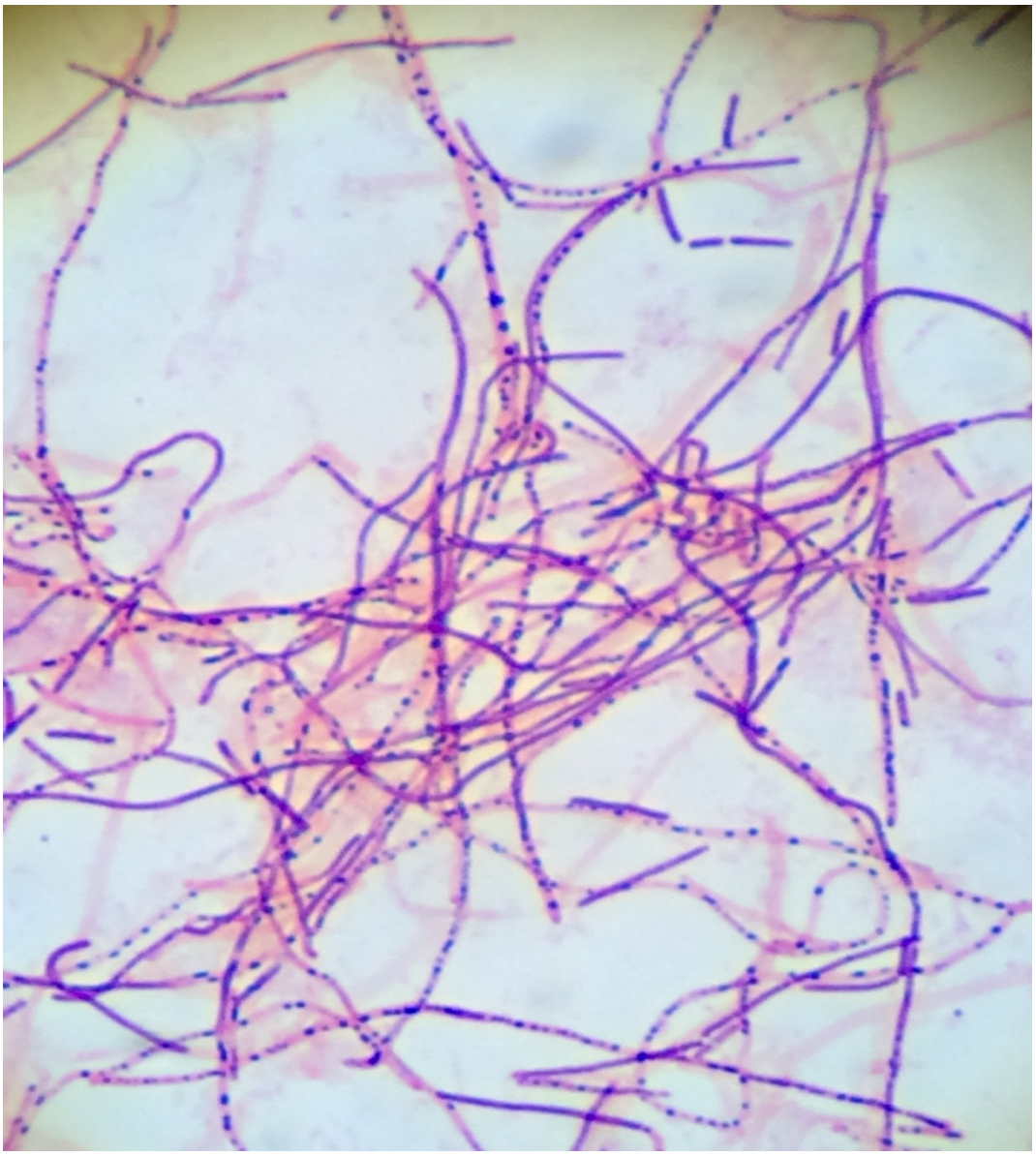

Figure 9: **Differentiated Cells of *Candidatus* arthromitus**

4.3.16 Determining Nutrient Use of the Consortia Using Biolog Plates. To determine nutrient use of the microbial consortia contained in the gastrointestinal tract of turkeys, the bacterial layer was isolated from the rest of the intestinal debris of three replicates. Gut contents suspended in SFB liquid media were strained through 100 µM mesh. The effluent was separated from the solids

trapped on the mesh and pipetted into micro centrifuge tubes. Tubes containing effluent were then centrifuged at 8,000 g for 5 min. Supernatant was discarded and pellets were resuspended in 5 mL PBS buffer. Diluted samples were layered onto 2 mL 30% and 3 mL 50% Nycodenz solutions in 15 mL conical tubes and centrifuged at 4,000 g for 10 min. The tubes were carefully removed from the centrifuge and the middle layer containing bacteria was isolated and placed into micro centrifuge tubes. Bacteria were centrifuged at 8,500 g for 10 min to separate bacteria from Nycodenz solution. Pellets were resuspended in 1 mL PBS. 300 µL of each sample was spread plated onto BUG +B agar plates in triplicate. Plates were incubated at 42°C in an anaerobic tent (1.4% O_2) for 24 h. Preparation of Biolog IF-0 and IF-0+ solutions were completed prior to harvesting BUG +B plates. IF-0 solution was prepared by obtaining a 1:5 ratio of sterile water/IF-0 stock solution. 1.8 mL of this solution was added to 15 mL conical tubes. IF-0+ solution was prepared by adding 23.2 mL sterile water and 1.8 mL IF-0 dye to 125 mL stock solution. 9 mL of this solution was added to 15 mL tornado tubes to create IF-0+ tubes. BUG +B plates with growth were harvested and the cells were placed into corresponding IF-0 tubes (without dye). 150 µL of this cell suspension was transferred into a new 15 mL tornado tube and diluted until the optical density at 590 nm (OD_{590}) reached 0.171+/- 0. 001. This procedure was carried out for all samples individually. Once the desired OD_{590} was reached for each replicate, 1.8 mL of this diluted cell suspension was transferred to tubes containing 9 mL IF-0+. For Phenotypic Microarray's (PM) 3-

4, 108 µL of 200M sodium succinate/10 M ferric citrate solution was added to IF-0+ tubes and mixed gently by inversion. Individually, samples were transferred to a sterile reservoir and pipetted into Biolog PM 1-4 plates. PM 1-4 plates were prepared in duplicate for each consortia replicate to yield 24 total PM plates. 100 µL were pipetted into each well using a multichannel pipette. OD_{600} readings were taken for time 0, 12, 24, 48, and 72 h using a Biotek® microplate reader.

4.3.17 qPCR for quantification.

4.3.17.1 Standard Curve Preparation and Construction. A standard curve was constructed using an artificial DNA fragment resembling the PCR amplicon known as a G-block (sequence: F:5'CCCCGCCTGAGGAGTACGATCGCAAGATTAAAACTCAAAGGAATTGACGGGGGACCCGCACAAGCAGCGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCTAAACTTGACATACCTT3'). This product was ordered from Integrated DNA Technologies (IDT) (Owczarzy et al., 2008). In order to prepare the standard curve, the lyophilized product had to be dissolved in TE buffer to yield a final concentration of 50 ng/µL. In this case, 25 µL of TE were added to the original vial. The vial was then incubated at 50°C for 20 min as per IDT's specifications for the product. 5 µL aliquots of DNA fragment were then stored at -20°C until standard curve construction was performed using the Roche® Lightcycler 96 device. When constructing a standard curve, samples were thawed and diluted using Roche® qPCR external DNA diluent for standards (product #). The stock solution was first diluted 1:100 and then serially diluted to

10^{-9} to obtain 0.5, 0.05, 0.005, 0.0005, 0.00005, 0.000005, 0.0000005, 0.00000005, 0.000000005 ng/20uL reaction volume. Thermocycling conditions for constructing a standard curve can be found in the Lightcycler 96 Reaction Protocol below. Upon completion of thermocycling, amplification results were used to find a linear dynamic range in order to measure subsequent sample runs for absolute quantification.

4.3.17.2 Preparation of Samples and Reagents. DNA samples were collected and isolated from colonies grown on BUG + B plates, in BUG broth, in Brain Heart Infusion (BHI) broth, and in SFB media (Schnupf et al., 2015) for 1, 2, 3, 4, and 5 days. The protocol for DNA isolation was conducted as described previously. After collection, samples were stored at -20°C. 10 µL of Roche® Sybr Green Master Mix [04707516001] with intercalating dyes was used for each reaction. Also present in each reaction was 3 µL of Millipore water and 0.75 µL each of forward and reverse primers. The final concentration of DNA template for each reaction was 10-50 ng/20 µL. The final concentration for each primer was 275 nM for each reaction.

4.3.17.3 Lightcycler 96 Reaction Protocol. Reactions were carried out in a Roche Lightcycler 96 machine. A 600 s pre-incubation step was initiated first, followed by a 3-step amplification portion that lasted for 45 cycles. The steps in the amplification cycles were to first denature at 95°C for 10 s, anneal at 54-55°C for 15 s, which was the optimum annealing temperature for the forward primer, and measure at 72°C for 30 s, for a total run time of 1 h and 40 min. The final

step after all cycles were completed was a high resolution melting step carried out at 95°C for 10 s, 40°C for 60 s, 65°C for 1 s, and 97°C for 1 s in order to determine the specificity of the amplified product. Products were stored at 4°C for later use.

4.3.17.4 Primer Concentration Optimization for PCR. Different concentrations of primers were used to create a gradient. A 1:10, 1:5,2.5:10, 7.5:10, 5:1, and 10:1 ratio was used in separate round-bottom, flat-top PCR strips for a final total of six strips. All strips were placed into a Bio-Rad thermocycler PCR amplification machine. Parameters were set to 95°C denaturation, 72°C annealing temperature, 54°C melting temperature, and a 95°C extension phase. An infinite hold was set to 4°C to avoid degradation of PCR products. A heated lid was used to mitigate temperature differences between the sample and the PCR strip. Gel electrophoresis was conducted using a 1.7% agarose gel solution run at 90V. ~5µL of PCR product was mixed with ~1µL gel loading dye by pipetting back and forth several times. Each sample was loaded into a separate well in the 1.7% agarose gel. The gel was run for ~1h. The gel was then stained in a 10µL:100mL heated (45°C) solution of 1X TAE buffer and Ethidium Bromide for ~30 min. A gel doc Bio Rad imaging machine was used to identify bands of DNA in the gel using UV transmission light and a Chemidoc XRS scanner.

4.3.17.5 Primer Design and Reconstruction for qPCR. Original thermocycle PCR primers (Schnupf et al., 2015) were unsuccessful when utilized to conduct

qPCR. A new set of PCR primers was designed first by using Muscle (Edgar, 2004) to construct a multiple alignment of successfully sequenced amplicons using the primers above, then selecting the area of sequence within the alignment that all amplicons shared as the sequence input in IDT's Primer Design tool for qPCR (Owczarzy et al., 2008). Primer Design generates five sets of forward and reverse primers as well as a probe in between the two. The probe was not ordered because Roche Sybr Green Master Mix with intercalating dyes was being used for fluorescence. The decision to order the generated set 3 of primers from IDT was determined by using NCBI's NBLAST tool (Dumontier and Hogue, 2002) to determine what each primer sequence would align with. Set 3 was the most precisely alignment with detecting *Candidatus* arthromitus.

4.3.18 Preparation for Sanger Sequencing. PCR purification was performed on successful PCR results determined by gel electrophoresis. The purification process was carried out using the QIAGEN® Mini Kit PCR purification system (Cat No./ID: 28106). 15 µL of purified samples were dispensed into separate 0.5 mL sterile micro centrifuge tubes. An order was placed with the University of Minnesota Genomic Sequencing Center (UMGC) for "Individual Tubes" and "Sanger QC." Each tube had the initials of the account holder along with a number. The numbers for each sample in a single order were selected consecutively. Standard desalting was selected on the order sheet along with a paper specification sheet and the pick-up location as St. Paul, Snyder Hall. Once the order sheet was filled out and the samples aliquoted and labeled, samples

were taken to 20 Snyder Hall (UMGC) for quality control (QC) and Sanger sequencing. Sequence results were reported as exportable files on the UMGC portal from which they were downloaded and analyzed further.

4.3.19 Sequence Analysis. Sequences from PCR products were downloaded from The University of Minnesota's website. Ab1 files were then analyzed further using the Minnesota Supercomputing Institute (MSI) and the Citrix receiver (Design). Once logged onto the Citrix receiver, Seqman Pro (Kim and Mundt, 2011) was utilized to configure contigs in the ab1 files. To upload ab1 files into Seqman Pro, files must be emailed and downloaded onto the Citrix receiver default desktop then added into Seqman Pro. Contigs were assembled for each sequence and single files were exported onto the Citrix desktop and transferred back to the original desktop via email. Configured sequences (FASTA files) were compared to all other sequenced genomes using NCBI's nucleotide BLAST tool with parameters selected to exclude models and non-cultured bacteria and to only include genome sequences with high similarity.

DNA concentration and purity were determined prior to sample submission for Sanger Classic QC sequencing. The University of Minnesota Genomics Facility performed quality control tests to determine these parameters. 384 well plates were used to hold 20 µL of each sample. A purity level of ~1.8 or higher is defined as "pure." The optimal concentration for the DNA template generated by SFB specific primers was 25-50 ng/µL due to the ~200 bp length of the fragment. Once purity and concentration were determined, the correct amount of DNA was

added to 500 µL submission tubes along with 6.4 pmoles of primer. The 6.4 pmoles of primer were prepared by creating a 1:100 dilution of 1 µL of 100 uM primer “stock” solution into 99 µL of nuclease-free water. The concentration of the dilute solution was 1 pmol/µL. 6.4 µL of dilute primer solution were added to 500 µL submission tubes. 4.6 µL of nuclease-free water were added to each tube to bring the final volume up to 12 µL. Samples were then ready to be submitted for sequencing.

4.3.20 Freezing Down Cells. Colonies growing on BUG +B plates determined to be pure were harvested by placing 250 µL of PBS buffer solution onto agar media containing colonies. Using a sterile cell scraper, colonies were scraped off of agar and placed in 1.2 mL micro centrifuge tubes along with 500 µL of additional PBS buffer. Tubes containing colonies and PBS were centrifuged at 8,000 g for 4 min. The supernatant was drained off and cells were resuspended in 500 µL of PBS buffer. Each tube was adjusted to an OD_{600} of 0.8 -1.2 using a spectrophotometer. Cell suspensions were divided into 250 µL aliquots and placed into sterile cryo tubes. 250 µL of sterile glycerol was added to cryo tubes and mixed so that cells were coated evenly for freezing at a final 50% glycerol concentration. Cryo tubes containing cell suspensions and glycerol were sealed and placed into a -80°C freezer.

4.13.21 Growth Curve Analysis. A control growth curve was completed using 200 µL of sterile BUG broth in all inner wells of a 96-microwell, lidded plate. 200 µL of sterile water was placed into the outer wells of the plate to avoid drying out

of the media. The plate was run in a Biotek® Microwell Plate automated spectrophotometer inside of an anaerobic chamber containing 1-4% oxygen. The Biotek machine was set to 42°C incubation for the entirety of the run. The growth curve protocol was created on the Biotek® machine and consisted of a 168 h (7 d) run at 42°C with constant shaking and a singular OD_{600} reading every hour. After the 168 h run was completed for the control, the degree of drying out in wells was assessed.

To run a live culture instead of a sterile media control, the same protocol was followed except 5 mL aliquots of BUG broth media were inoculated with a single colony of SFB grown on BUG + B agar plates in the anaerobic chamber at 42°C. These tubes were incubated for 48 h or until turbid in the anaerobic chamber at 42°C prior to use in the Biotek® machine. Turbid cultures were adjusted to an OD_{600} of 0.04 prior to plating in a 96-well plate for growth curve construction. After inoculation of 96 microwell plates, the protocol laid out above was used as described. Results were obtained after 168 h and analyzed.

4.3.22 Growth curve construction on BHI. BHI broth containing SFB was incubated for ~17days in order to reach an OD_{600} of 0.1±0.01. A 1:10 dilution of the culture was then made using fresh, sterile BHI broth. Diluted culture was transferred to columns 3-10 and rows C-F of a 96-well plate in 200 µL aliquots. Empty perimeter wells were filled with 600 µL of sterile water to avoid evaporation and concentration of the media that would affect OD readings. When using a 384-well plate, sterile water was pipetted into rows A, B, C, N, O, P

and columns 1-3, and 21-24 to avoid evaporation of the culture/media. Each triplicate sample was distributed into two rows instead of one single row as in the 96-well plates. A protocol was constructed in a Biotek® microwell plate reader for a 168 h (7 day) run with continuous double orbital shaking, an incubation temperature of 42°C and OD_{600} readings taken every 1 h for 168 h.

4.3.23 Growth Curves on BHI Supplemented with Individual Nutrients. BHI containing 37 g/L was supplemented with ten individual nutrients. The nutrients in order from one to ten were: sucrose, ribose, cellobiose, D-mannose, D-glucosamine, adenosine, L-aspartic acid, L-asparagine, maltose, and melibiose. Each compound was prepared as an 85% solution, then added to 9 mL of sterile BHI to yield 0.1-1% and 1-10% supplemented solutions of BHI broth. 90 µL of each solution was distributed into a single well of a 384-well plate going in order from 0.1-10% in a single row. Rows C-L contained compounds 1-10 in columns 3-21. Column 22 contained 90 µL of blank BHI solution in each well. Columns 1, 2, 23, and 24 contained 100 µL of sterile water. Rows A, B, M, N, O, and P also contained 100 µL of sterile water. All wells containing 90 µL of media also had 10 µL of SFB culture added to them in order to create a 1:10 dilution of the stock culture. The original reading of the stock culture prior to dilution was 0.1 at OD_{600}.

Once prepared, the 384-well plate was placed into a Biotek microwell automated spectrophotometer inside of an anaerobic chamber containing 1-4% O_2 and the experiment was conducted at 42°C. The protocol set up on the

machine took an OD reading at 600nm each hour for 72 hours while continually shaking. The lid was left on for the duration of this experiment. After the 72-hour experiment was completed, the plate was removed from the plate reader, the data taken for analysis, and the water wells of the plate were refilled. The plate was then placed back into the machine to repeat the experiment for another 3 days in order to collect more data points for the growth curve.

Once optimal amounts for each nutrient that increased growth were determined, a new 384-well plate was pipetted out containing the amount of nutrient that increased growth for each compound, plated out in triplicate 100 µL amounts. A BHI blank well consisting of 100 µL of control media was pipetted for each set of triplicates along with a blank BHI media well that was not to be inoculated by culture, but to be used for subtracting the blank BHI optical density from the inoculated wells to determine the changes in optical density from bacterial growth. Also included in this plate was a row containing BHI media supplemented with all ten compounds at their optimal amounts in order to determine if combining them all would yield an even larger increase in growth. This inoculated microwell plate was run using the same parameters as described above.

4.3.24 Plate-Count Method on BHI. Two mL BHI broth tubes were prepared for two separate isolated SFB colonies growing on BHI agar plates. Each of these tubes became the stock solution once inoculated aseptically with a single colony from BHI agar plates. Tubes were sealed and vortexed to break up the colonies

to ensure homogenous suspension of bacterial cells. Four additional tubes of two mL BHI broth were prepared for each initial tube. 200 µL of stock bacterial cell culture was transferred into a sterile 1.8 mL BHI tube to create a 1:10 dilution. Serial dilutions were completed for each of the two stock solutions out to 10^{-4}. Each serial dilution tube was vortexed prior to transferring to make sure the cells were suspended and had not settled at the bottom. 100 µL of each stock solution and serial dilution were plated onto individual BHI agar plates using the spread plating method. Plates were sealed with parafilm and placed upright into a 42°C incubator in microaerophilic conditions for 24 h. Stock solutions and serial dilutions were also incubated using identical conditions. Plating of stock solutions and serial dilution was repeated at 24, 48, 72, 96, 120, 144, and 168 h in order to generate Day 0-7 bacterial enumeration values for solid media growth curve construction.

4.3.25 Dilution to Extinction on BHI. The BHI plate count method above was utilized to create a third 2 mL stock solution and set of serial dilutions for plating on BHI plates using the same amounts in the above protocol. Instead of bacterial enumeration, the goal of this experiment was to determine how far SFB can be diluted on BHI before it stops growing and to determine the purity of colonies growing on further diluted plates.

4.3.26 Dry Cell Weight. 30 mL of CA in BHI was grown in triplicate to yield three biological replicates for 14 days to an OD ranging from 0.076-0.090 at 600 nm. 25 of the 30 mL for each sample were used to inoculate 225 mL of BHI

supplemented with 10% w/v D-Glucosamine and 8% w/v Sucrose. Samples were incubated until the log phase of growth was attained. Once log phase was attained, 50 mL samples at 0, 4, 8, and 16 hours were harvested and filtered using 0.22 µm glass microfiber filters and a Buchner funnel connected to a sidearm flask and a vacuum system. Samples were dried at 80°C for 24 h and weighed on an analytical scale. Original filter weight was subtracted from final filtered sample weight + filter weight to yield the filtered dry cell weight.

4.3.27 Illumina Next-Generation Sequencing. Illumina Next-Generation Sequencing was performed by CLC bio, a Qiagen ® company on two CA samples provided by the Baumler lab. Samples were prepared using the DNA isolation protocol and kit mentioned above. One sample (1) was isolated from cells taken from a previously frozen stock while the other sample was isolated from fresh cells in colonies on a BUG+B media plate (2). Nucleotide distribution, contig measurements, accumulated contig lengths, summary statistics, distribution of read length, matched read length, non-matched read length, and paired reads distance distribution were determined.

4.3.28 Fluorescent In-Situ Hybridization. Fluorescent In-Situ Hybridization (FISH) was performed using three probes containing different fluorophores in order to determine presence of SFB, other Bacillus cells, and potential contamination by *Achromobacter spp.* Probes for *Bacillus* spp. (CTT CAG CAC TCA GGT TCG), *Achromobacter* (CGC TCY AAT AGT GCA AGG TC), and SFB (GGG TAC TTA TTG CGT TTG CGA CGG CAC) targeting the 16S rRNA region

of each organism. (Snel et al., 1995) Snel *et al*. provided the SFB 16S rRNA probe while both the *Bacillus* spp. and *Achromobacter* probes were found on the PROBER (Navin et al., 2006) online database. Probes were ordered through IDT (Owczarzy et al., 2008) and rehydrated with 1X TE buffer. 5 ng amounts of each probe were utilized in 20 µL aliquots of hybridization solution for this experiment. Cells were fixed using 4% paraformaldehyde in PBS buffer in a 1:3 cell culture/fixative ratio, incubated for 16h at 4°C and pelleted using a table top micro centrifuge at 4°C and 12, 000 rpm. Cells were washed with PBS and pelleted again, then resuspended in PBS to yield a ~110 cells/mL concentration. For storage, one volume of fixed cell solution was mixed with one volume of cold absolute 100% ethanol and stored at -20°C for two weeks (Amann et al., 1990). Fixed cells were removed from storage and applied to glass slides and allowed to air dry for 60 min. A 20 µL aliquot of hybridization solution (0.9 M NaCl, 20 mM Tris/HCl, 0.01% SDS, pH 7.2) containing 5 ng of each probe per slide were spotted onto fixed cells and incubated in a dry heating oven at 50°C overnight. Cells were submerged in washing solution (0.9 M NaCl, 20 mM Tris/HCl, 0.01% SDS, pH 7.2) at 48°C for 15 min. Slides were rinsed with sterile water and allowed to air dry and then counter-stained with (2-(4-amidinophenyl)-6-indolecarbamidine dihydrochloride) DAPI for 15 min away from light. Slides were rinsed with 80% ethanol and sterile water and allowed to air dry. One drop VectaShield (Catalog #H-1200) mounting media containing 1.5 µg/mL DAPI was placed onto each slide along with a clean coverslip prior to storage and imaging.

Slides were stored in a slide case wrapped in foil at 4°C. Slides were imaged at the University of Minnesota Imaging Center using a Nikon Eclipse 90i light microscopy scope and camera with RGB filters to capture fluorescence (Snel et al., 1995). When imaging, the exposure time was set separately for each probe in order to eliminate background signal. The *Bacillus* spp. probe had an exposure time of 400 ms, the *Achromobacter* probe had an exposure time of 700 ms, and the SFB probe had an exposure time of 2s.

JPG images of FISH results were opened with Fiji (Fiji is just Image J) program in order to adjust the background threshold and measure only fluorescence beyond a signal strength of 10. To do so, images were individually opened in Fiji, the “Image” tab was selected and the image “type” was changed from RGB color to grayscale 8-bit. In the “Image” tab, the image was then duplicated and the threshold was adjusted on the new duplicate. The red background was set to 255 and the black background was set to 0. The black background was then adjusted to a value of 10 for signal intensity in order to eliminate anything below this threshold. To save the threshold setting, the “Edit” tab was opened and the “create selection” option was chosen. The “Edit” tab was opened again and the “add to selection manager” option was chosen which caused a Region of Interest (ROI) manager tab to open. In order to adjust the threshold on the original image, the selection number in the ROI manager tab was highlighted and the measurement was taken by using “ctrl + M.” Mean,

mode, median, standard deviation, maximum, minimum, and total measured area were recorded.

4.4 Results & Discussion

4.4.2 Mauve multiple genome alignments. The Mauve multiple genome alignments can be found in figures 10, 11, and 12 with results summarized in figure 13. Overall, the CA genomes used for generation of the five GEMs were found to be highly similar however, differences among *Candidatus* arthromitus strains were found and genes encoding unique reactions were identified for the strain isolated from turkeys. A genome for a *Clostridium botulinum* strain was included in the first comparative genome alignment (Figure 10) to demonstrate the level of genomic conservation between a free living organism and an auxotrophic one within the same family of organisms. Figure 13 (graph) summarizes the metabolic differences of gene content between strains. As can be seen, the *Clostridium botulinum* strain contains more genes associated with the metabolism of carbohydrates, amino acids, co-factors and vitamins (Figure 13). The *Candidatus* arthromitus strains contain more genes associated with the metabolism of proteins to break them down to their amino acid components. *Candidatus* arthromitus may rely on the host to break down amino acids into their subsequent carbohydrate and nitrogen sources. Figures 11 and 12 contain an alignment with the *Clostridium botulinum* strain removed in both regular and backbone view. When aligned in Mauve, the five CA genomes reveal a high level of genome conservation as indicated by the similarity of the Local Colinear

Blocks (LCBs) displayed in the Mauve viewer, although regions in LCBs with white indicate some sequence variation and some genomic islands unique to some of the CA strains. The genome alignment in backbone view demonstrates regions of all five CA genomes that are conserved for all strains (colored in Mauve), and regions shared by subsets of strains (For example regions in yellow are shared in the four CA genomes from mice and rat hosts, yet absent in the CA isolate from turkeys), or genes that are completely unique to a single strain. Blank areas within the alignment are associated with genomic islands unique to a single CA genome and are of interest since they may contain genes that may be involved in host specificity.

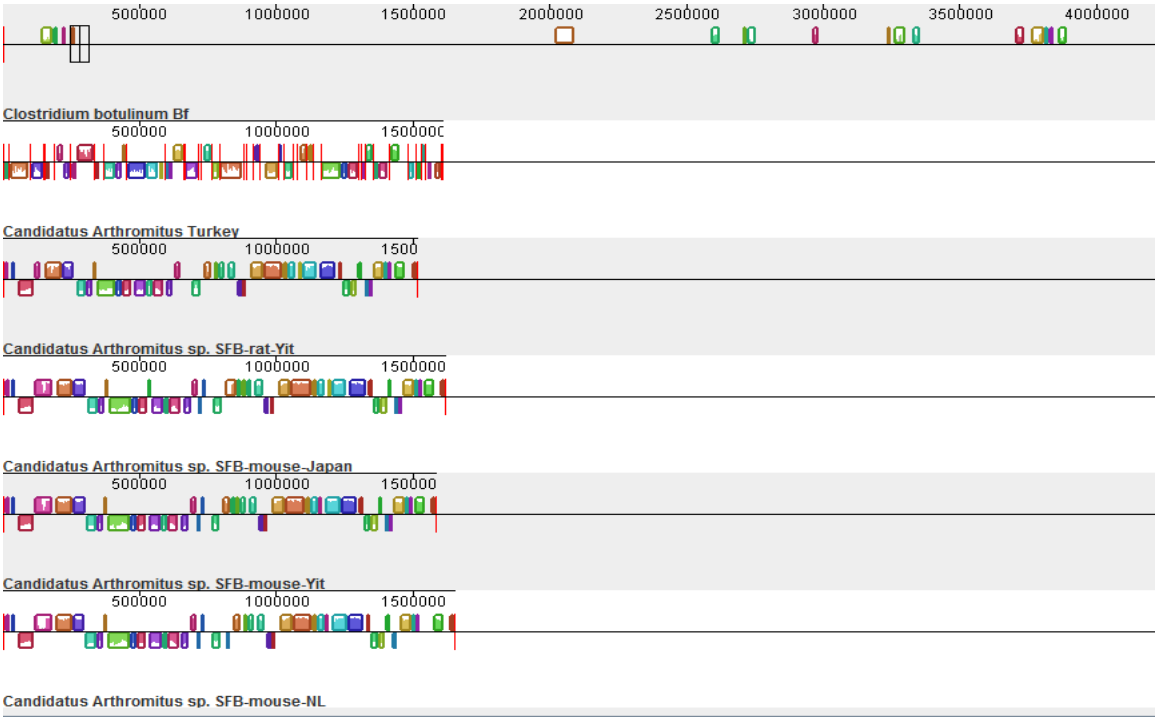

Figure 10: Mauve Genome Alignment Including *C. botulinum*

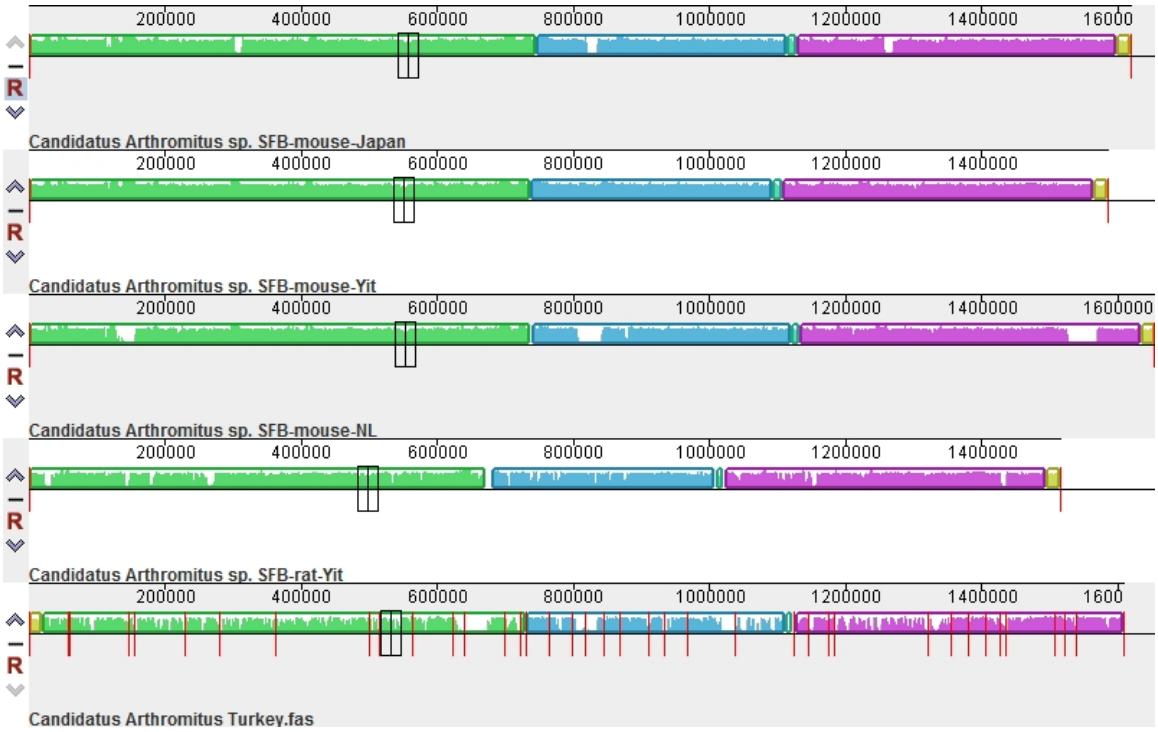

Figure 11: Mauve Genome Alignment Containing only the 5 *Candidatus* arthromitus Strains

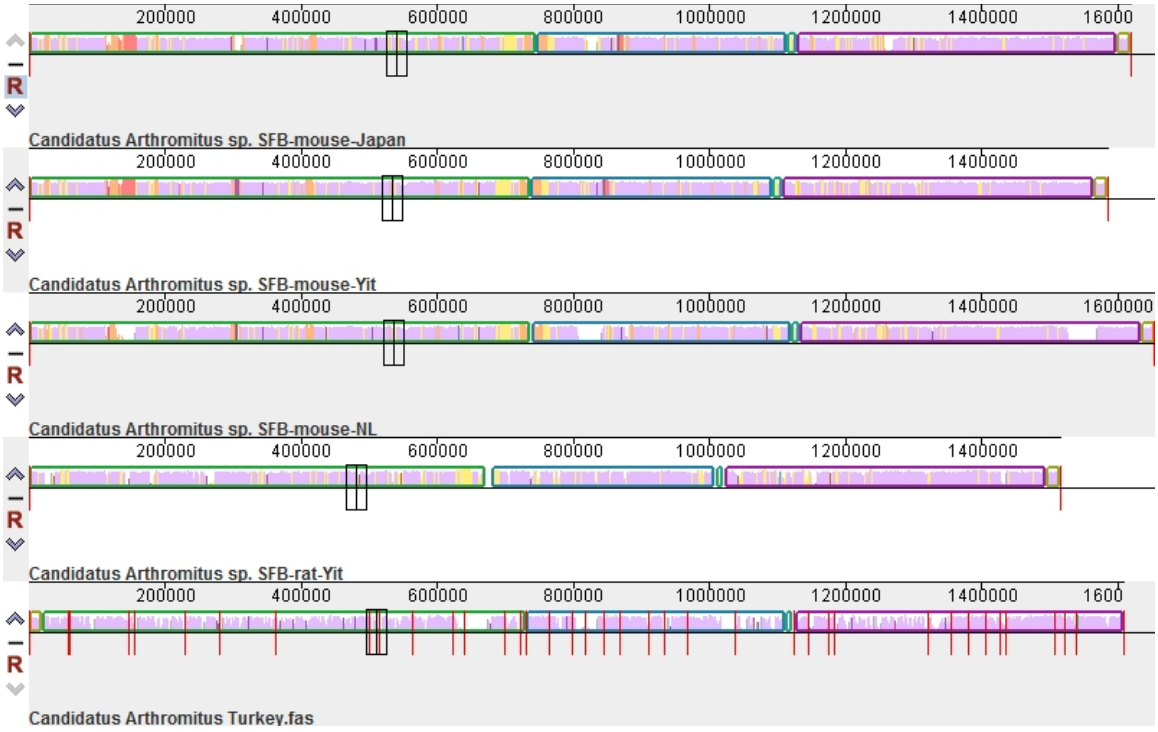

Figure 12: Backbone View of 5 *Candidatus* arthromitus Strains Mauve Genome Alignment

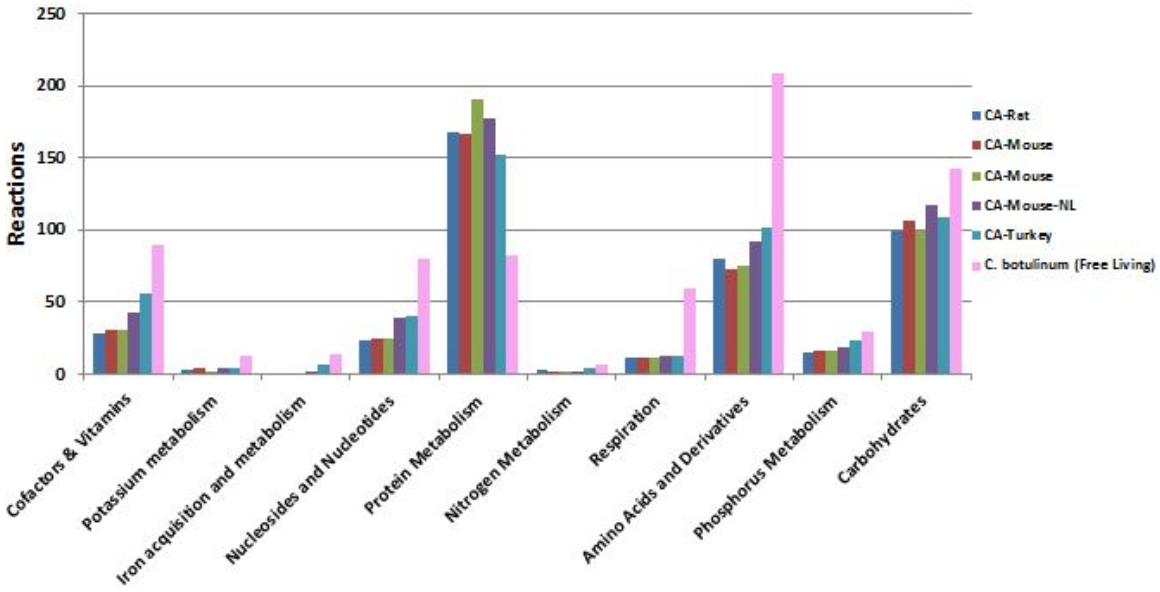

Figure 13: Predicted metabolic reactions contained in the genomes of 5 CA strains to *Clostridium botulinum*

4.4.3 Genome-Scale Metabolic Models. GEMs were constructed using the Department of Energy's new genome-scale metabolic network reconstruction pipeline KBase. FASTA genome sequences were uploaded into KBase and then annotated. Once annotated, the genomes were analyzed and all genes predicted to encode proteins involved in metabolic reactions were identified and reconstructed to create metabolic network containing all metabolites and metabolic reactions to generate GEMs to be used as prediction tools. Five GEMs were created, one for each CA strain. The five CA GEMs were converted into SBML files and downloaded to be used with the optimization software package GAMS. When the reactions comprising GEMs were compared for all five CA strains, several reactions unique to individual strains were identified. Table 4 summarizes the total number of reactions found to be common in all five CA

strains as well as reactions found in some and not others. KBase also provided an actual gene count from FASTA genome sequences, the relative number of genes contained in each model, and the number of reactions in each GEM. Table 3 provides a comparison of GEM statistics for the composition of all five models.

Table 3: Overall summary of gene count, genes in GEMs, and metabolic reactions for 5 CA strains

Genome	Genes	Abbreviation	Genes in Model	Metabolic Reactions
Mouse Japan	1,546	Mse_Japan	348	860
Mouse YIT	1,473	Mse_YIT	351	861
Mouse NL	1,589	Mse_NL	351	859
Rat YIT	1,401	Rat_YIT	336	863
Turkey	1,554	Turkey	365	863

Table 4: Summary of common reactions for all 5 CA GEMs and reactions missing or unique to the turkey CA GEM

Missing in Turkey	Found in All 5 Models	Unique to Turkey
58	801	47

Though unique reactions were found for all models, the main focus of this project is *Candidatus* athromitus isolated from turkeys. The genome-scale metabolic model for this strain contained 47 unique reactions not present in any of the other

four CA GEMs. Of these unique reactions, two major reactions were most interesting. raffinose fructohydrolase (Figure 14) and biotin synthase (Figure 15) were both found to be reactions specifically unique to the metabolic model generated for *Candidatus* arthromitus isolated from turkeys. Raffinose fructohydrolase is part of the galactose pathway (KEGG database #R02410) and reacts with water to produce melibiose and D-fructose. According to the metabolic model analysis made for this strains GEM, raffinose is not predicted as a carbon source that can be utilized by this strain of *Candidatus* arthromitus*.* However, the D-fructose formed as a product once the reaction completes may be utilized by turkey-isolated *Candidatus* arthromitus. Biotin synthase is part of the biotin metabolism pathway (#R01078, KEGG) and functions as a catalyst for this reaction. It belongs to the sulfurtransferase family, transferring sulfur groups. One co-factor for this reaction is Iron-Sulfur that is found within the reaction pathway. Biotin, also known as vitamin H, is a water-soluble B vitamin and also a coenzyme for carboxylase enzymes involved with the synthesis of fatty acids, valine, and isoleucine as well as playing a role in gluconeogenesis, which allows an organism to produce sugars for catabolic reactions from non-carbohydrate precursors. Intestinal bacteria such as SFB (CA) are known to produce Biotin that the host scavenges from waste products before they are excreted. The requirements for Biotin are small however, if SFB (CA) are not present in the intestinal tract, this may be causative of a lighter weight and thus poor performing host as well as other issues with immunodeficiency (Danzeisen et al., 2013).

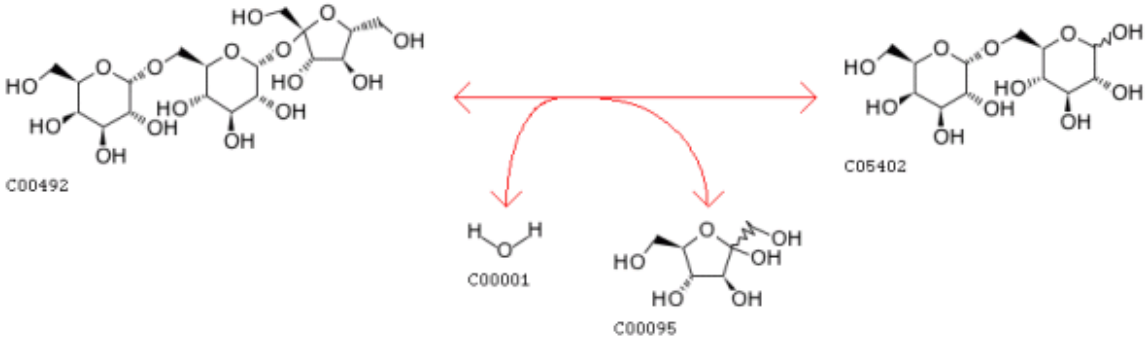

Figure 14:Raffinose Fructohydrolase Reaction
Source: http://www.genome.jp/dbgetbin/www_bget?R02410+RP02188+RC00077

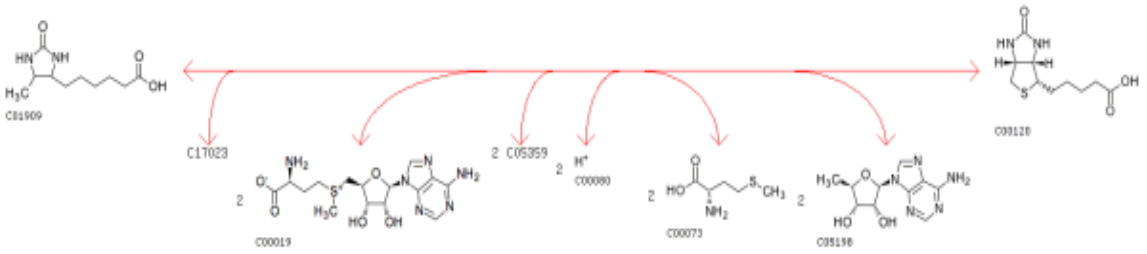

Figure 15: Biotin Synthase Pathway
Source: http://www.genome.jp/dbget-bin/www_bget?rn:R01078
Kuwahara, T, *et al* (2011) DNA research
Carey, C, *et al* (1976) The Journal of Nutrition
Hernández-Vázquez, A, *et al* (2012) Molecular Genetics and Metabolism

One study that focused on Biotin deficiency indicated slower growth in experimental mice during the first 6 weeks of Biotin deficient diet consumption, followed by progressive weight-loss in weeks 7-20. Also indicated upon completion of this experiment was the inability of subjects to mount an adequate immune response (Baez-Saldana et al., 1998).

4.4.4 Use of GEMs for Metabolic Inquiry. GAMS, the General Algebraic Modeling System was used conduct flux balance analysis (FBA) on models while

optimizing the biomass of cells. GEMs were used to screen a cascade of nutrients for sole carbon source usage, identify secreted metabolites, and nutrient growth sources predicted to increase biomass. Table 5 contains the results for *in silico* sole carbon source utilization for all GEMs. The turkey GEM was predicted to utilize sucrose as a carbon source while no other GEM was predicted for this metabolic capability. D-ribose was predicted for utilization as a sole carbon source by all three mouse isolated CA GEMs. Glucosamine was predicted for utilization only by the rat-YIT isolated CA GEM. Also, cellobiose was predicted for utilization by all five CA GEMs and is found primarily in plants. The diets of the animal hosts that all strains correspond to have plant-based diets. All models generated the same four secreted compounds when grown on D-glucose: H_2O, O_2, H+, and CO_2.

Table 5: Sole carbon source predicted for utilization for all 5 CA models

Carbon Source	Mouse_NL	Mouse_YIT	Mouse_Japan	Rat_YIT	Turkey
Sucrose	-	-	-	-	+
Cellobiose	+	+	+	+	+
Maltose	+	+	+	+	+
D-Fructose	+	+	+	+	+
D-Glucose	+	+	+	+	+
D-Mannose	+	+	+	+	+
D-Ribose	+	+	+	-	-
Glucosamine	-	-	-	+	-

Table 6: Nutrient growth sources that are predicted to increase biomass in decreasing order for all 5 CA strains

Mouse_NL	Mouse_YIT	Mouse_Japan	Rat_YIT	Turkey
Cellobiose	Cellobiose	Cellobiose	Cellobiose	Sucrose
D-Maltose	D-Maltose	D-Maltose	D-Maltose	Cellobiose
D-Glucose	Mannitol	D-fructose	Mannitol	D-Maltose
D-Fructose	D-Fructose	D-Mannose	D-Fructose	D-Fructose
D-Mannose	D-Mannose	Glucosamine	D-Mannose	D-Mannose
D-Ribose	Glucosamine	D-Ribose	Glucosamine	D-Mannitol
N/A	D-Ribose	Mannitol	D-Glucose	D-Glucose
N/A	D-Glucose	D-Glucose	D-Glucosamine	D-glucosamine
N/A	Deoxyribose	Deoxyribose	Deoxyribose	Glucosamine

Table 6 summarizes the *in silico* predictions for all GEMs when screened for nutrient growth sources increasing biomass in decreasing order. A 1.0 value for *in silico* predicted biomass threshold was set as a baseline. Several amino acids increased biomass to some extent but did not reach this threshold. As can be seen in Table 7, cellobiose is the nutrient that is predicted to increase biomass the most in all but the turkey CA model. Sucrose replaces cellobiose for *in silico*

predictions to increase biomass for the turkey GEM. This is supported by the work of Thaxton and Parkhurst (Thaxton and Parkhurst, 1976) who state the feeding poults a 10% sucrose solution in water prior to feeding exposure enhances their numerical weight.

The turkey GEM was expanded with additional nutrient exchange reactions to include 190 carbon, 95 nitrogen, and 95 phosphorous and sulfur sources present on Biolog ™ PM 1-4 sources for comparison with Biolog ™ plate results. The results can be found in table 7 below for nutrient sources found to be utilized experimentally.

Table 7: Nutrients found to increase growth from PM 1-4 Biolog ™ plates

Mouse_NL	Mouse_YIT	Mouse_Japan	Rat_YIT	Turkey
Cellobiose	D-Glucose	Cellobiose	Maltose	Sucrose
D-Fructose	D-Mannose	D-Maltose	D-Maltose	D-Maltose
D-Glucose	D-Ribose	D-Fructose	D-Mannose	Cellobiose
D-Mannose	N/A	D-Mannose	D-Fructose	D-Fructose
D-Ribose	N/A	D-Ribose	D-Glucose	D-Mannose
N/A	N/A	D-Glucosamine	N/A	Melibiose
N/A	N/A	N/A	N/A	Adenosine
N/A	N/A	N/A	N/A	L-Asparagine
N/A	N/A	N/A	N/A	L-Aspartic Acid

4.4.5 Gut Extractions I and II. Gut extraction I resulted in extracted ileums from poults that would then be dissected for the ileum contents for isolation of CA. The ileum sections were successfully dissected in gut extraction II to yield intestinal contents that was processed further by filtering, centrifuging, and using Nycodenz solutions in order to isolate the bacterial layer and finally CA. This isolation method has been successful in isolating CA from most other bacteria in the gastrointestinal tract.

4.4.6 SFB Media and Technique

4.4.6.1 Growth on solid BUG +B Media. Isolated CA from consortia samples plated onto BUG +B plates yielded growth within 24 h. Growth and development of bacteria continued for 72 h when bacteria would begin to produce intrasegmental bodies. In the first 24-48 h bacteria stain primarily Gram negative. From 48-72 h and thereafter, bacteria stain primarily Gram positive. Spores within bacterial filaments form and are likely the reason behind Gram positive staining in later stages of development. Figure 17 shows isolated CA staining as primarily Gram negative with Gram positive spores inside the bacterial filament. Growth on BUG +B plates has shown to be the most promising media method to grow CA. Additional parameters for CA bacterial growth are incubation temperature of 42°C and microaerophilic conditions between 1-4% O_2. CA growth was determined on BUG +B plates in aerobic conditions as well as after prolonged exposure to 4°C temperatures for storage.

4.4.6.2 Other Working Growth Mediums. Media described in Schnupf *et al*, 2015 did not yield significant growth in our lab**.** DMEM media used as the base of media described in Schnupf *et al*, 2015 also did not yield significant growth. DMEM media was tested along-side of BHI with supplemented nutrients in a 96-well plate put into a microwell plate reader to construct a growth curve in order to come to this conclusion.

Working media that optimized growth conditions consisted of Biolog™ universal growth medium both with and without 50mL of defribilated sheep's blood per liter (BUG +B/BUG media). BUG + B and BUG prepared liquid media also yielded growth. Use of BUG media resulted in contamination by mold, however in order to limit the growth of fungi while still optimizing the growth of CA, a topical antifungal agent was added to the media recipe. 3.35 µL/L of 100X cycloheximide effectively limited the growth of fungi while still allowing for growth of CA. NBLAST was used for sequence analysis and results did not reveal any residual contamination by mold spores.

Other working growth mediums tested and verified visually via Gram staining but not on a molecular basis (qPCR, PCR/sequencing) include Brain Heart Infusion (BHI) Agar plates and "OK" agar plates. "OK" agar plates contained minimal growth at the inoculation site that appeared to be prone to contamination with mold spores. BHI plates had ~10 CFU's on each plate that formed after 96 h of incubation at 42°C. When Gram stained, no contamination

was apparent. SFB cells were slow growing and still producing spores and elongating after 96 h.

4.4.7 Isolation of CA. Liquid media recipe outlined in Schnupf, 2015 (Schnupf et al., 2015), preserved SFB bacteria effectively however, the growth by bacteria in liquid media was minimal. Optical density readings were taken with a Chromate Reader microplate reader set to 630 nm wavelength starting from day zero and finishing after seven days. DMEM media was also used to generate a growth curve however as can be seen in Figure 16 below, no statistically significant growth occurred. Trends in optical density increases over time were plotted however trends in optical density for each sample do not follow a normal growth curve or a general increase over time. Samples of liquid media containing culture was plated onto BUG +B plates after seven days to allow for visual colonies to grow. These colonies were harvested, Gram stained and visualized at 100X magnification to determine bacterial morphology. Bacterial colonies growing on BUG +B plates from liquid media did not represent the bacterium of interest. Colonies contained Gram negative amorphous bacteria instead of segmented filamentous bacteria that stain both Gram positive and Gram negative. A visual representation of these bacteria are shown in Figure 17.

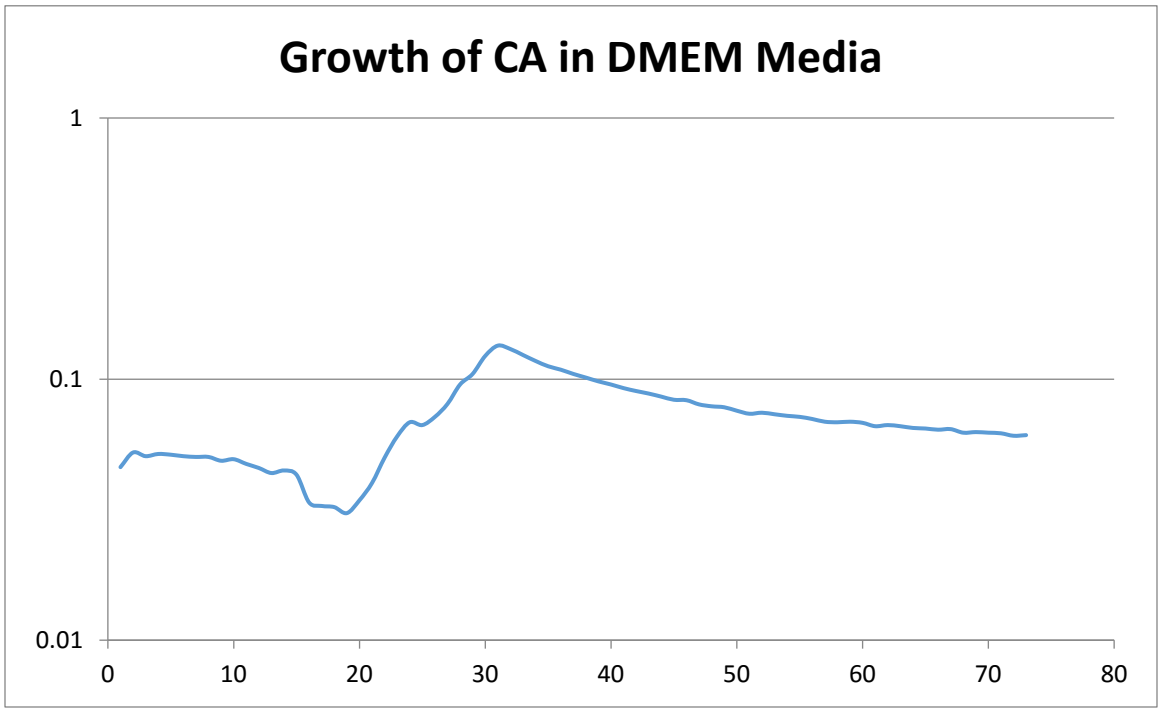

Figure 16: CA growth in DMEM base media based on optical density measurements over time

4.4.8 Determining SFB Visual Characteristics. As stated in the literature review, CA cells contain intrasegmental bodies as they mature. The intrasegmental bodies can be seen in Figure 16 below as small dots inside of Gram-negative stained cells. CA cells can stain either Gram negative or Gram positive, depending on their level of maturity which likely explains the pink and purple filamentous cells imaged below.

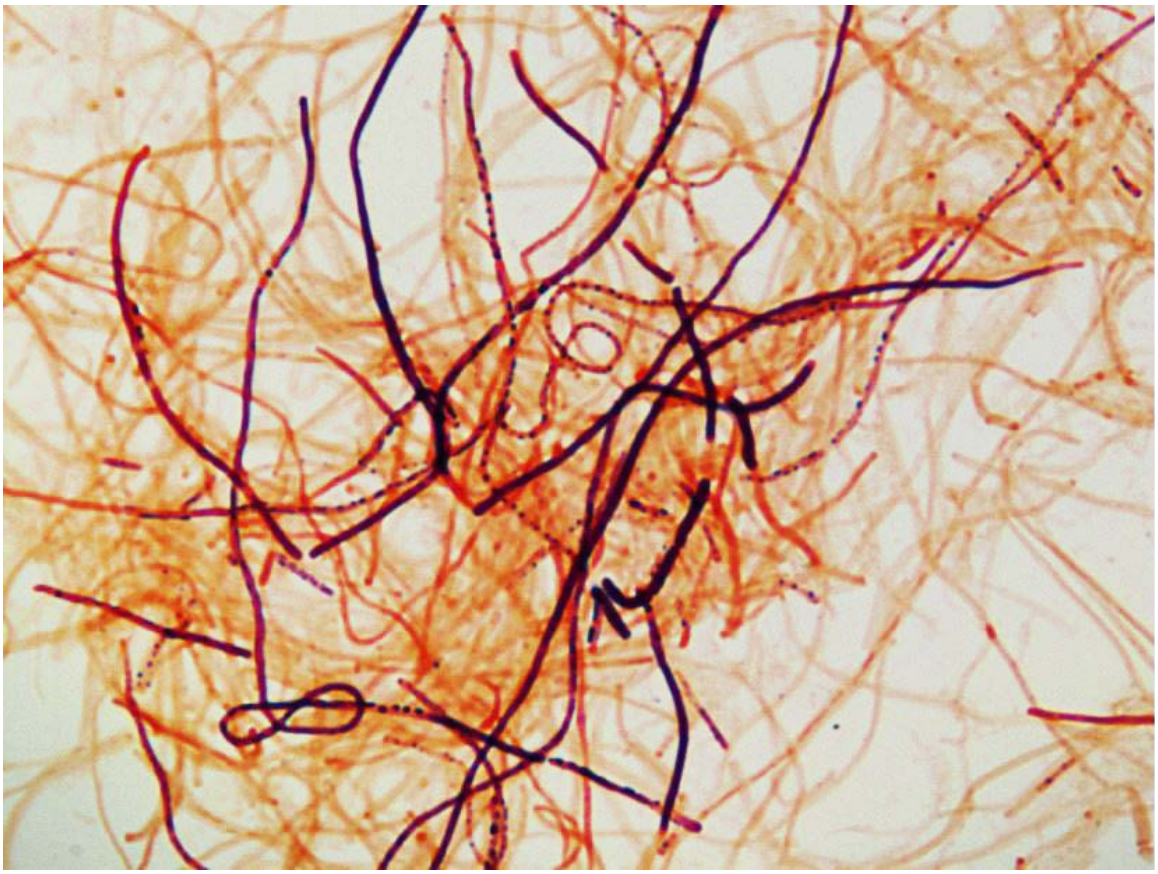

Figure 17: Gram stain of *Candidatus* arthromitus using Nikon Eclipse 90i camera from University of Minnesota Imaging Center (UIC)

4.4.9 Biolog™ phenotypic assays.

4.4.9.1 Microbial Consortia from turkey ileum. Figure 18 represents the results for Biolog ™ Phenotypic Microarray plates 1-2 completed for three samples with duplicates for each plate. As shown in Figure 19, the gastrointestinal consortia can utilize a wide variety of carbon, nitrogen, phosphorous, and sulfur nutrients. Utilized nutrients for PM1 plates ranged from 64-96 nutrients generating growth between the three samples. Nutrients contained on PM2 plates that were utilized ranged from 67-95 utilized nutrients. PM plates 1-2 contained single carbon sources. PM3 plates contained individual

nitrogen sources within each well and ranged from 85-95 utilized nutrients between samples. PM4 plates contained individual sulfur and phosphorous source and ranged from 85-96 utilizable ingredients between samples. Figure 18 provides a comparison of all nutrients that do not support growth for any or all of the three consortia samples.

Carbon Sources					
PM1 Compounds	PM1 Consortia	PM1 CA Co-Culture	PM2 Compounds	PM2 Consortia	PM2 CA Co-Culture
Negative Control			Negative		
L-Arabinose			Chondroitin Sulfate C		
N-Acetyl-D-Glucosamine			alpha-Cyclodextrin		
D-Saccharic Acid			beta-Cyclodextrin		
Succinate			gamma-Cyclodextrin		
Galactose			Melitose		
L-Aspartate			Gelatin		
L-Proline			Glycogen		
D-Alanine			Inulin		
TRHL			Melitose		
D-Mannose			Maltotetraose		
Dulcose			Pectin		
L-Serine			UDP-N-Acetyl-D-Galactosamine		
Sorbitol			N-Acetyl Neuraminic Acid		
Glycerol			beta-D-Allose		
L-Fucose			Amygdalin		
D-Glucuronic Acid			L-Lyxitol		
GLCN			L-Lyxitol		
D,L-α-Glycerol-Phosphate			L-Lyxitol		
L-Arabinose			Ursin		
D-Lactate			2-Deoxy-D-Ribose		
Formate			i-Erythritol		
D-Mannitol			D-Fucose		
L-Glutamate			3-O-beta-D-Galactopyranosyl-D-Arabinose		
D-Glucose-6-Phosphate			Gentiobiose		
D-Galactonic acid-γ-lactone			beta-D-Glucose		
D,L-Malic Acid			Lactitol		
D-Ribose			Melitose		
Tween 20			Maltitol		
L-Rhamnose			alpha-Methyl-D-glucoside		
D-Fructose			beta-Methyl-D-Galactoside		
Acetate			3-Methyl Glucose		
beta-D-Glucose			beta-Methyl-D-Glucuronic Acid		
Maltose			alpha-Methyl-D-Mannoside		
Melibiose			beta-Methyl-D-Xyloside		
Thymidine			Palatinose		
L-Asparagine			D-Raffinose		
D-Aspartic Acid			Salicin		
D-Glucosaminic Acid			Sedoheptulosan		
1,2-Propanediol			L-Sorbose		
Tween 40			Stachyose		
α-Keto-Glutaric Acid			D-Tagatose		
2-Oxobutyrate			Turanose		
α-Methyl-DGalactoside			L-Lyxitol		
α-D-Lactose			N-Acetyl-D-Glucosaminitol		
LACT			GABA		
Sucrose			delta-Amino Valeric Acid		
Uridine			Butyrate		
L-Glutamine			Capric Acid		
M-Tartaric Acid			Caproic Acid		
Glucose-1-Phosphate			Citraconic Acid		
D-Fructose-6-Phosphate			Citramalic Acid		
Tween 80			D-Glucosamine		
α-Hydroxy Glutaric Acid-γ-Lactone			2-Hydroxy Benzoic Acid		
α-Hydroxy Butyric Acid			4-Hydroxybenzoate		
beta-Methylglucoside			beta-Hydroxy Butyric Acid		
L-Lyxitol			gamma-Hydroxy Butyric Acid		
Melitose			a-Keto-Valeric Acid		
DeoxyAdenosine			Itaconic Acid		
Adenosine			5-Keto-D-Gluconic Acid		
Glycyl-L-Aspartic Acid			D-Lactic Acid Methyl Ester		
Citrate			Malonic Acid		
M-Inositol			Melibionic Acid		
L-Threonine			Oxalic Acid		
Fumarate			Oxalomalic Acid		
Bromo Succinic			Quinic Acid		
Propionate			D-Ribono-1,4-Lactone		
Mucic Acid			Sebacic Acid		
Glycolate			Sorbic Acid		
Glyoxalate			L-2-Amino-acetoacetate		
CELB			D-Tartaric Acid		
Inosine			L-Tartaric Acid		
Glycyl-Lglutamic acid			Glycine		
Tricarballylic Acid			L-Alaninamide		
L-Serine			N-Acetyl-L-glutamate		
L-Threonine			L-Arginine		
L-Alanine			Glycine		
L-alanylglycine			L-Histidine		
Acetoacetate			L-Homoserine		
N-Acetyl-D-mannosamine			5-Aminolevulinate		
Mono Methyl Succinate			L-Isoleucine		
Methyl Pyruvate			L-Leucine		
D-Malic Acid			L-Lysine		
L-Malate			L-Methionine		
gly-pro-L			Ornithine		
p-Hydroxy Phenyl Acetic Acid			L-Phenylalanine		
m-Hydroxy Phenyl Acetic Acid			L-pyroglutamic Acid		
Tyramine			L-Valine		
D-Psicose			Carnitine		
L-Arabinose			Sec-Butylamine		
Glucuronamide			D,L-Octopamine		
Pyruvate			Putrescine		
L-Galactonic Acid-γ-Lactone			D-Glyceraldehyde		
D-Galacturonic Acid			2,3-Butanediol		
Phenylethylamine			2,3-Butanone		
Aminoethanol			3-Hydroxy 2-Butanone		

Figure 18: Positive and negative growth of consortia and CA Co-Culture on Biolog ™ PM 1-2 plates

4.4.9.2 Isolated CA. Biolog^{TM} assays were conducted on the biological replicates of isolated CA in BHI broth. The results of the three biological replicates were compared and combined into a single reported outcome based on OD_{600}. PM plates 1-4 were tested. On PM1 plates, CA did not grow on any compounds. On PM2 plates CA grew L-phenylalanine. On PM3 plates that screen for nitrogen utilization, CA grew on several amino acids: L-glutamine, L-histidine, L-serine, L-tryptophan, L-valine, L-tyrosine, D-glutamate, and D-lysine as well as other compounds histamine, inosine, guanosine, and a couple of complex amino acid complexes listed in Table 8. On PM4 plates CA grew on lipoamide, N-Acetyl-L-cysteine, thiosulfuric acid, phosphocreatine, phosphoserine, 6-phospho-D-gluconate, 2',3' cyclic guanosinemonophosphate, and pyrophosphate. Figures 18 & 19 provides a comparison of all nutrients that do not support growth for any or all of the three samples isolated from consortia.

Nitrogen Sources			Phosphorous/Sulfur Sources		
PM3 Compounds	PM3 Consortia	PM3 CA Co-Culture	PM4 Compounds	PM4 Consortia	PM4 CA Co-Culture
Negative			Negative Control		
NH3			Phosphate		
Nitrite			PPi		
Nitrate			Trimetaphosphate		
Urea			Triphosphate		
Biuret			Triethyl Phosphate		
L-Alanine			Hypophosphite		
L-Arginine			Adenosine-2'-monophosphate		
L-Asparagine			Adenosine-3'-monophosphate		
L-Aspartate			AMP		
L-Cysteine			Adenosine-2',3'-cyclic monophosphate		
L-Glutamate			Adenosine-3',5'-cyclic monophosphate		
L-Glutamine			Thiophosphate		
Glycine			Dithiophosphate		
L-Histidine			Glycerol-3-phosphate		
L-Isoleucine			beta-Glycerol Phosphate		
L-Leucine			Carbamyl Phosphate		
L-Lysine			2-Phospho-D-glycerate		
L-Methionine			3-Phosphoglycerate		
L-Phenylalanine			Guanosine-2'-monophosphate		
L-Proline			Guanosine-3'-monophosphate		
L-Serine			GMP		
L-Threonine			2',3'-Cyclic GMP		
L-Tryptophan			2',3'-Cyclic GMP		
L-Tyrosine			Phosphoenolpyruvate		
L-Valine			Phospho-Glycolic Acid		
D-Alanine			Glucose-1-phosphate		
D-Asparagine			D-Glucose-6-Phosphate		
D-Aspartic Acid			D-Glucosamine phosphate		
D-Glutamate			D-Glucosamine phosphate		
D-Lysine			6-Phospho-D-gluconate		
D-Serine			CMP		
D-Valine			3'-CMP		
Citrulline			CMP		
L-Homoserine			2',3'-Cyclic CMP		
Ornithine			2',3'-Cyclic CMP		
N-Acetyl-L-glutamate			D-Mannose1-phosphate		
N-Phthaloyl-L-Glutamic Acid			D-Mannose-6-Phosphate		
L-Pyroglutamic Acid			Cysteamine-S-Phosphate		
Hydroxylamine			Phospho-L-Arginine		
Methylamine			Phosphoserine		
N-Amylamine			Phosphoserine		
N-Butylamine			L-Threonine phosphate		
Ethylamine			Uracil		
Aminoethanol			3'-UMP		
Ethylenediamine			UMP		
Putrescine			2',3'-Cyclic UMP		
Agmatine			2',3'-Cyclic UMP		
Histamine			O-Phospho-D-Tyrosine		
beta-Phenylethylamine			O-Phospho-L-Tyrosine		
Tyramine			Phosphocreatine		
Acetamide			Phosphoryl Choline		
Formamide			O-Phosphoryl-Ethanolamine		
Glucuronamide			Acetylphosphate		
D-Alanine			2-Aminoethyl Phosphonic Acid		
D-Glucosamine			Methylene Diphosphonic Acid		
D-Galactosamine			dTMP		
N-Acetyl-D-Mannosamine			dTMP		
N-Acetyl-D-Glucosamine			Inositol Hexaphosphate		
N-Acetyl-D-Galactosamine			Thymidine-3',5'-cyclic monophosphate		
N-Acetyl-D-Mannosamine			Negative Control		
Adenine			Sulfate		
Adenosine			H2S2O3		
Cytidine			Tetrathionate		
Cytosine			Thiophosphate		
Guanine			Dithiophosphate		
Guanosine			L-Cysteine		
Thymine			L-Cysteine		
Thymidine			Cys-Gly		
Uracil			L-Cysteic Acid		
Uridine			Cysteamine		
Inosine			L-Cysteine Sulfinic Acid		
Guanosine			N-Acetyl-L-Cysteine		
Xanthosine			homocysteine		
Thiamin			Cystathionine		
Alloxan			Lanthionine		
Allantoin			GSH		
Parabanic Acid			D,L-Ethionine		
GABA			L-Methionine		
L-Leucine			D-Methionine		
L-Isoleucine			Gly-Met		
D,L-alpha-Amino-Caprylic Acid			N-Acetyl-D,L-Methionine		
delta-Amino-N-Valeric Acid			L-Methionine S-oxide		
alpha-Amino-N-Valeric Acid			L-Methionine Sulfone		
ala-L-asp-L			L-Djenkolic Acid		
Ala-Gln			Thiourea		
ala-L-glu-L			1-Thio-beta-D-Glucose		
L-alanylglycine			Lipoamide		
Ala-His			Taurocholic Acid		
Ala-Leu			Taurine		
ala-L-Thr-L			Hypotaurine		
gly-asn-L			p-Amino Benzene Sulfonic Acid		
Gly-Gln			Butane Sulfonic Acid		
gly-glu-L			2-Hydroxyethane Sulfonic Acid		
Gly-Met			Methane Sulfonic Acid		
met-L-ala-L			Tetramethylene Sulfone		

Figure 19: Positive and negative growth of consortia CA Co-Culture on Biolog™ PM 3-4 plates

Table 8: List of compounds in PM plates 1-4 utilized for growth by CA

Use of Compounds by Isolated CA			
PM1 **(Carbon)**	PM2 (Carbon)	PM3 (Nitrogen)	PM4 (Phosphorous or Sulfur)
None	Sedoheptulosan	L-Glutamine	Ppi
	Butyrate	L-Histidine	2',3'-Cyclic Guanosinemonophosphate
	Ornithine	L-Serine	Phosphoserine
		L-Tryptophan	Phosphocreatine
		L-Tryosine	Thiosulfuric Acid
		L-Valine	N-Acetyl-L-Cysteine
		D-Glutamate	Lipoamide
		D-Lysine	
		Histidine	
		Inosine	
		Guanosine	
		ala-L-asp-L	
		Ala-Gln	

4.4.9.2.1 Comparison to Growth Curve Results. Growth curve experiments and analysis revealed increased growth when supplemented with D-mannose or D-sucrose with statistical significance compared to the growth control ($P<0.005$) while carbon sources contained in Biolog ™ PM1 plates did not yield growth on either of these carbon sources. This may have been due to the short experiment time, or that Biolog ™ assays test for growth on a nutrient as a sole source of carbon, whereas growth curve assays had nutrients supplemented to an existing growth media.

4.4.9.2.2 Comparison to *in silico* Predictions. *In silico* predictions and Biolog ™ experimental results demonstrated an 80.00% agreement for PM1 plates with nine false positives predicted by the *in silico* model. PM2 plates had a 96.00% agreement compared to the Biolog ™ results with only one false negative compound reported as no growth based on *in silico* predictions by the model. PM3 plates had an 80.00% agreement rate compared to the *in silico* predictions with three false positives and nine false negatives. PM4 plates had a 90.70% agreement with *in silico* predictions with only four false negatives. All false positives/negatives can be found listed in Appendix C.

4.4.10 Dilution to Extinction Well-Plate Method. Dilution to extinction was performed on consortia samples having gone through the SFB specific isolation process outlined in Schnupf et al., 2015 (Schnupf et al., 2015). Although the microplate reader determined growth occurred over the course of each week-long period taken to grow and dilute each plate, the growth was not significant for

this project. The SFB media used with a minute number of cells did not allow for the survival of SFB bacteria. Tan mucoid colonies formed on new samples plated out after three weeks (three rounds) of dilution to extinction had been completed. When Gram stained, the bacteria were found to be gram positive coccoid bacteria forming clusters.

4.4.11 Dilution to Extinction on BHI Agar Medium.

Dilution to extinction on BHI agar plates was not successful. Dilutions did not yield consistent results.

4.4.12 Plate Count Method Using BHI Broth and Agar Medium.

Plate count method using BHI broth and agar medium was not successful, as bacterial cultures had a tendency to form too few colonies early on in all dilutions and then form bacterial lawns as the culture reaches 48 h onward.

4.4.13 qPCR Absolute Quantification. Many attempts were made at carrying out qPCR reactions using the Roche Lightcycler 96 apparatus. Initially, parameters were set to conduct a standard curve using artificial DNA emulating the DNA amplicon fragment cleaved during PCR. The first run with standard curve dilutions, non-template controls (NTCs), negative controls, positive control plasmid DNA, and several samples was not successful as no amplification occurred. In order to obtain a successful run, the Tm was adjusted down from 60°C to a 52-54°C gradient, corresponding to the annealing temperature of the forward primer. Original runs were completed with the use of a qPCR probe with FAM dye at the start, and a double quencher of ZEN/IBFQ. Use of the probe

however was unsuccessful. Instead, Roche Sybr Green Master Mix with intercalating dyes was used. Although not as specific as the use of a specially designed probe, use of Sybr green master mix allowed for amplification to occur and fluorescence to be recorded. Specificity of the reaction was determined using a melting curve at the end of the run. The melting curve determines the number of different fragment sizes in the PCR product as different-sized amplicons will all have different melting temperature. Several adjusted reactions using Sybr green and a melting curve were also unsuccessful. Many peaks upon completion of melting curve analysis displayed that the reaction was overall non-specific warranting the design of a new set of primers.

New primers were designed using a sequence of DNA contained in three amplicons successfully sequenced from PCR products. These three amplicon sequences were compared using the Muscle (Edgar, 2004) multiple sequence alignments system. The matching sequence portion revealed by the alignment was copied and pasted into IDT's primer design tool (Owczarzy et al., 2008). The primer design tool generates five sets of primers and probes that can then be used with NCBI's NBLAST database (Dumontier and Hogue, 2002). After using NBLAST for all five sets of primers and probes, it was determined that set three had the most optimal set of primers and thus that set of primers was ordered via the University's IDT portal (Owczarzy et al., 2008). The probe designed in between the forward and reverse primer was not ordered as the Sybr green dye was still in use.

Once available for use, primers were rehydrated in PCR-grade sterile water and optimized using SFB-positive samples in a standard Bio-Rad thermocycler. The melting temperature for this primer set was 56.3°C and was optimized at 55°C by using two samples of isolated SFB DNA, one initially producing two melting curve peaks, and one initially producing one melting curve peak. A temperature gradient was set from 54-55°C across a 96-microwell plate with each sample spanning the length of the plate. Samples were placed in center rows D and E of the plate with a row of non-template controls below in row F and sterile water in rows G and C to balance the plate out. Thermocycling conditions included an initial 600s 95°C incubation period, and 3-step amplification as specified above with the temperature gradient used instead of a single temperature across the entire plate. Once the reaction was complete, qPCR reaction file was used with the Roche Lightcycler 96 analysis software. Two analysis methods were applied to determine the results of primer optimization: Tm calling and Absolute Quantification. Tm calling addresses the melting curve step conducted at the end of the qPCR run. Each reactive well's peak is overlaid over the others in order to determine how closely aligned the melting points for each well are. During this optimization run, several melting peaks appeared. To address the possibility of impurities within samples, digested qPCR reactions were sent in for Sanger sequencing. Absolute quantification was used to determine an arbitrary "call" for each well. This means that the software gives a positive or negative symbol for each well to display whether or not

amplification occurred for each well. Also determined by absolute quantification was the relative Ct values of NTC's and sample wells. Ct values correspond to the degree of amplification/fluorescence detected for each well. Figure 20 displays samples run with an acceptable standard curve and Figures 21 and 22 display the values associated with the acceptable standard curve and the linearized trend resulting from standard curve construction. Absolute quantification was achieved however demonstrating growth of CA in this way was not successful. DNA template samples were obtained from plates containing growth from 1, 2, 3, and 4 days. The concentrations of DNA for these samples did not reflect an increase in growth. Figure 23 shows the amplification curves of both the standard curve samples and the samples run for absolute quantification. This may have to do with the DNA isolation technique. Running qPCR did support our belief that we have isolated CA and maintained its viability. Digested qPCR products were purified and sent in for sequencing to reveal their direct match with CA using NCBI's NBLAST (Dumontier and Hogue, 2002) network. Reconfiguring the contigs of these samples was not necessary as the base pair length for each sample was too short.

qPCR samples and reagents were sent out to the University of Minnesota's Expression Analysis Center to conduct an assay in order to determine if growth of CA was occurring and was not successfully determined in the Baumler lab. Table 9 displays the Ct values for each replicate along with averages and Table 10 shows the growth results in the form of Ct value

difference between days 1-2, days 2-4, and days 1-4. Figure 24 contains the amplification plot along with standard curves visible in addition to sample amplification curves. Figure 25 shows a bar graph associated with the growth founded in the decrease in Ct values from days 1-4. Figure 26 shows the Ct values for days 1, 2, and 4 plotted to show the decrease in value. Day three values did not fit the trend and thus were not used in the analysis process. The decrease in Ct value is associated with an increase in DNA copy number. Because the Ct values decreased from days 1-4, the DNA copy number for samples increased and thus growth of CA is supported based on these results. A Single-Tail Paired Student T-test was run on Ct values between triplicate samples from day 1 and day 4; T-test results can be found in Table 9. T-test results ($P < 0.01$) show that the differences in Ct values between days 1 and 4 are statistically significant thus the growth of CA in isolated DNA samples from days 1-4 is supported.

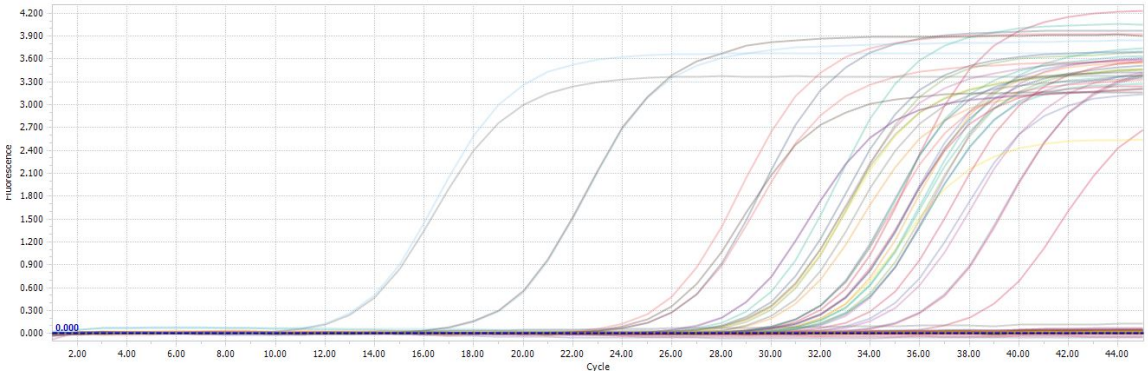

Figure 20: Amplification curves of CA samples grown in BHI broth
*** Standard curve points are included**

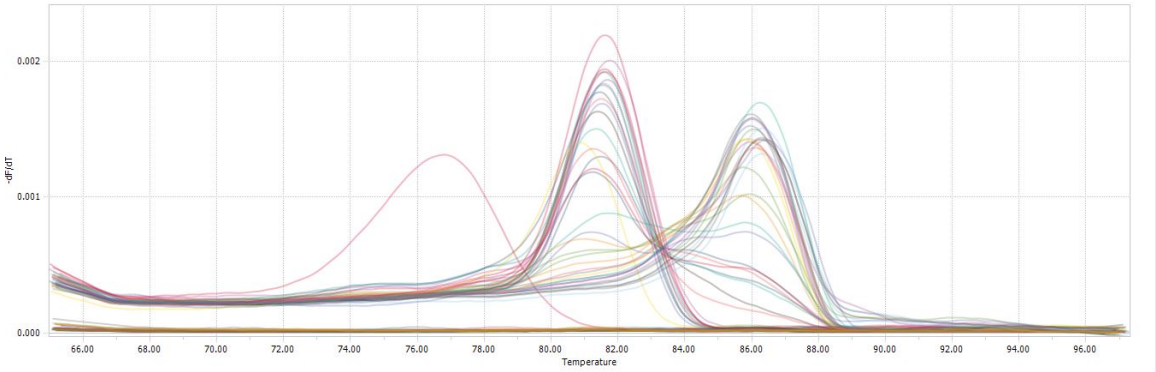

Figure 21: Melting peaks of CA samples grown in BHI
*** Standard curve points are included**

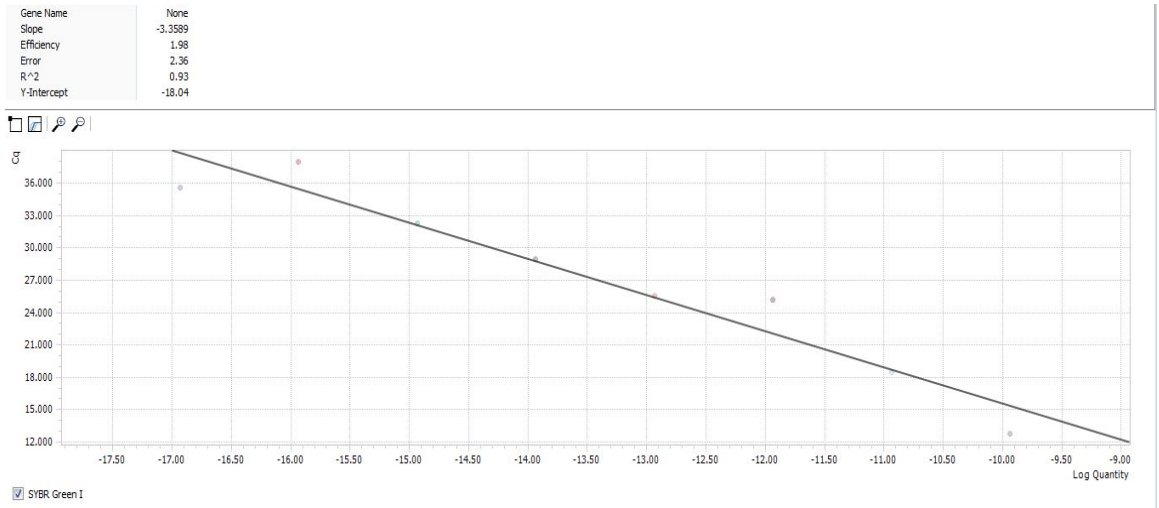

Figure 22: Standard Curve (SC) results for CA samples grown in BHI

Gene Name	None
Slope	-3.3589
Efficiency	1.98
Error	2.36
R^2	0.93
Y-Intercept	-18.04

Figure 23: Standard Curve values for CA samples grown in BHI

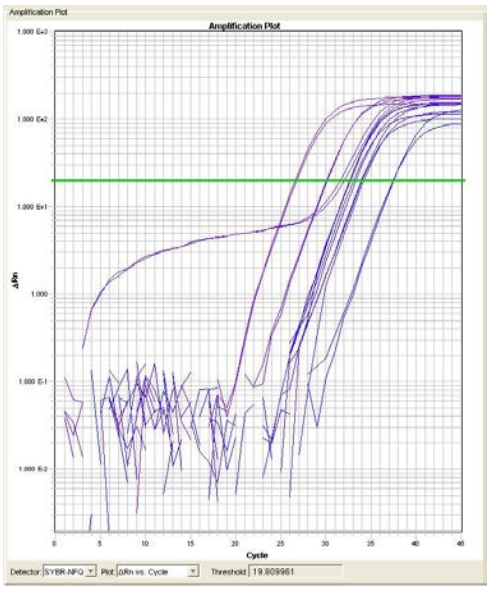

Figure 24: Amplification plot containing CA days 1-4 from triplicate samples and triplicate G-Block standard curves

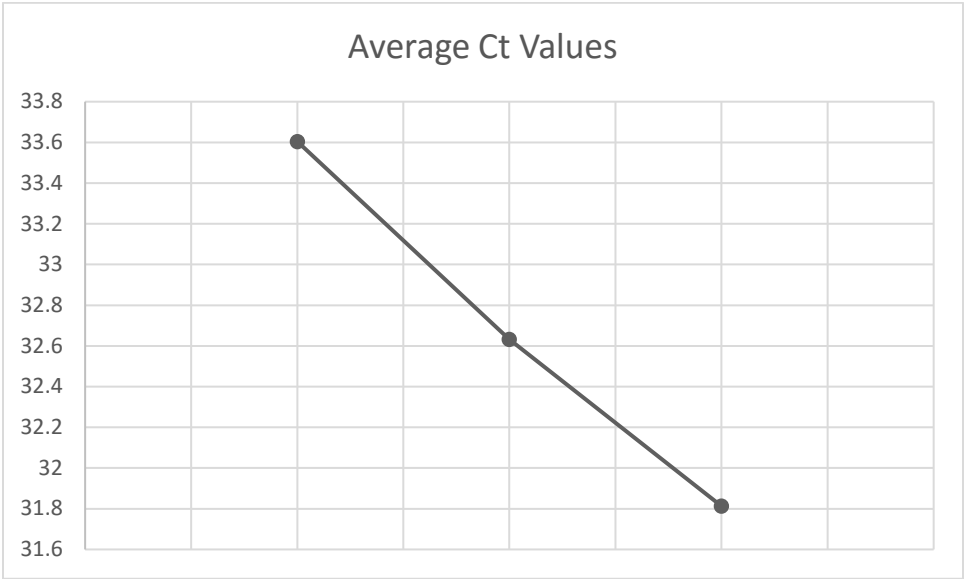

Figure 25: Average Ct values for days 1, 2, and 4 CA culture determined by qPCR

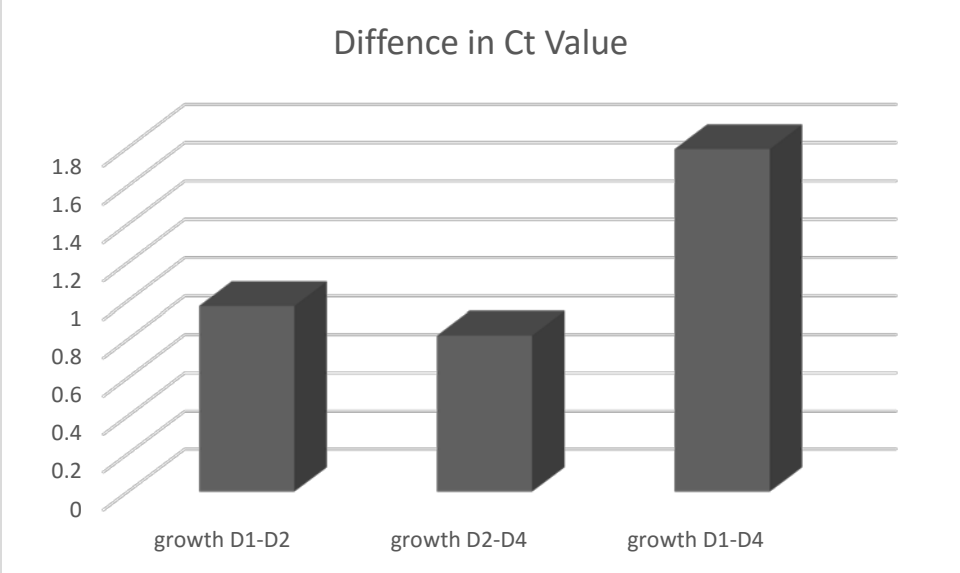

Figure 26: Ct value differences between days 1, 2, 4 CA culture run on ABI 7900HT

Table 9: Ct values for replicates 1-3, days 1-4 CA culture including Ct value mean and Single Tail Paired Student T-Test.

Samples	Ct Values	Ct Values	Ct Values	Mean
SFB D1	33.96	33.25	33.60	33.60
SFB D2	32.60	32.66	32.63	32.63
SFB D4	31.59	32.03	31.81	31.81
Day 1: Day 4 T-Test	0.01632537			

Table 10: Differences in Ct values between days 1-2, 2-4, and 1-4

Samples	Difference in Ct Value
growth D1-D2	0.972124
growth D2-D4	0.8197525
growth D1-D4	1.7918765

4.4.14 Growth Curve Construction Using a Biotek Microwell Plate Reader.

Growth curves were set up both in a 96 well plate and a 384 well-plate in order to determine consistency of results. The results of both 168 h runs yielded a positive growth curve constructed by averaging the triplicate samples from growth in microwell plates. At OD_{600}, a lag phase was observed in both plates from approximately zero to fifteen hours (Figure 27). A log or exponential phase of growth occurred from 15 to 50 hours. After 50 hours, the growth curve leveled out and a minor death phase set in around 75-100 hours. A diauxic growth phase occurred in the 96-well plate and less so in the 384-well plate after 100 hours but before 150 hours (Figures 27 & 28). This may have happened because living cells can scavenge nutrients from degradation of dead cells.

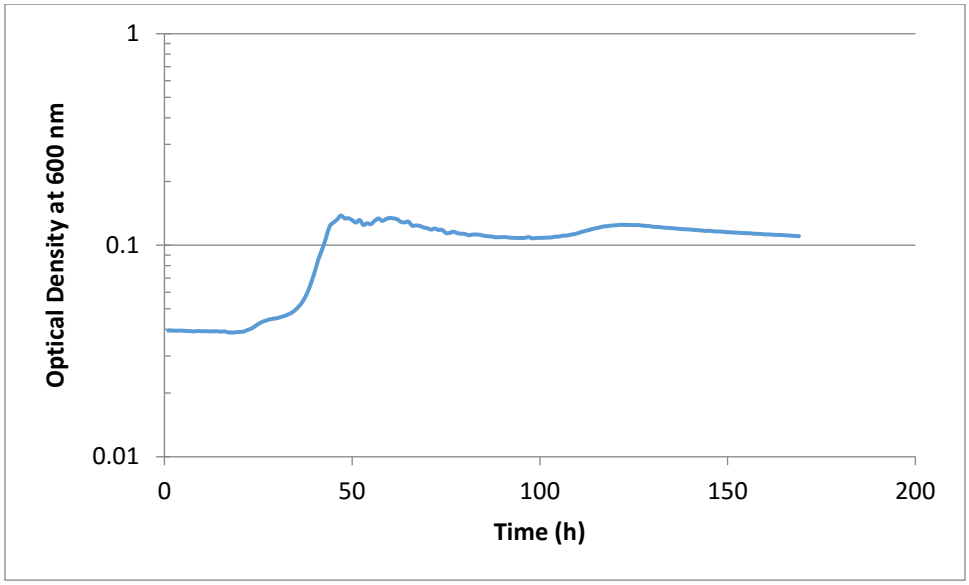

Figure 27: Growth curve of CA in BHI broth in 96-well plate

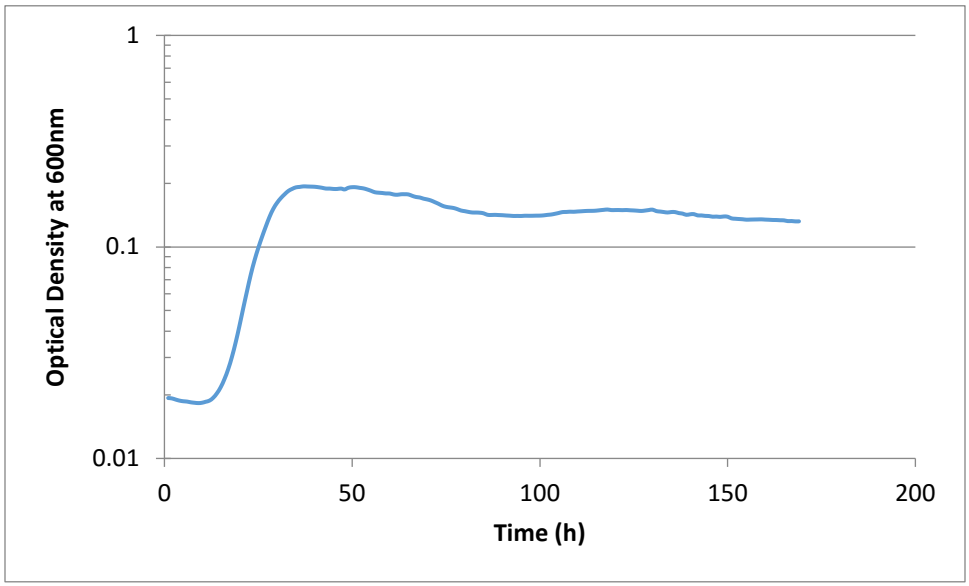

Figure 28: Growth curve of CA in BHI broth using 384-well plate

4.4.15 Growth Curve using BHI Supplemented with Nutrients. Experiments completed for 72h were successful. All compounds except for D-Maltose (Compound 9) at varying concentrations enhanced the growth of CA to some extent. Figure 29 displays growth curves to show increases in OD_{600} by adding compounds to BHI to screen for additional growth beyond the non-supplemented BHI control. Further experiments conducted on increasing amounts s of nutrients that increased growth past the baseline growth curve using only BHI. These values (w/v) were 8% sucrose, 1% D-ribose, 8% cellobiose, 10% D-mannose, 10% D-glucosamine, 2% adenosine, 9% of L-aspartic acid, 9% of L-asparagine, or 10% of melibiose in BHI. These percentages were tested in triplicated and a student t-test was conducted to determine statistical significance of results. Furthermore, a single row of wells was analyzed for growth containing BHI supplemented with the optimal percentage of all compounds in order to

determine if growth can be increased when all compound amounts were supplemented together in combination or used in tandem.

Nutrient supplements that supported increases in growth of each individual compound tested in triplicate are discussed below. Sucrose at 8%, D-glucosamine at 10%, and D-mannose at 10% in BHI were the most effective at increasing growth with statistical significance ($P<0.005$). Ribose at 1%, cellobiose at 8%, and adenosine at 2% all in BHI were effective at increasing growth in the diauxic phase of growth that occurs later in the curve with statistical significance ($P<0.005$). D-fructose at 0.7%, D-glucose at 8%, and melibiose at 10% in BHI were found to increase growth beyond the control but were not found to be statistically significant ($P = 0.0053$, .067, or .66, respectively). L-asparagine at 9%, and L-aspartic acid at 9% in BHI were not found to increase growth and were statistically insignificant. A combination containing the optimum percentages of all compounds had a potentially toxic effect as the control growth was greater than that observed using the nutrient mixture with statistical significance ($P<0.005$).

Combinations of nutrients demonstrating statistically significant positive growth during the initial logarithmic growth phase compared to the control were tested in combinations in order to determine the optimal set of compounds to be added to base BHI media for optimized growth of CA. D-glucosamine at 10%: D-mannose at 10%, D-mannose at 10%: sucrose at 8%, D-glucosamine at 10% : sucrose at 8%, and a combination of all three compounds were tested using the

same growth curve protocol described above. All combinations of two compounds resulted in statistically significant positive growth compared to the control (P<0.005). Figures 29, 30, 31, and 32 show the growth curves plotted over time. A slurry of all three compounds tested against the control did not have statistically significant growth (P= 0.065). Overall, the combination that had the lowest P value (P<0.0006) contained 8% sucrose and 10% D-glucosamine in BHI. These compounds in BHI were used in future dry cell weight experiments below. DMEM was screened for growth using this method as well. Growth did occur but results were statistically insignificant. Growth curve analysis and raw data results can be found in Appendix B.

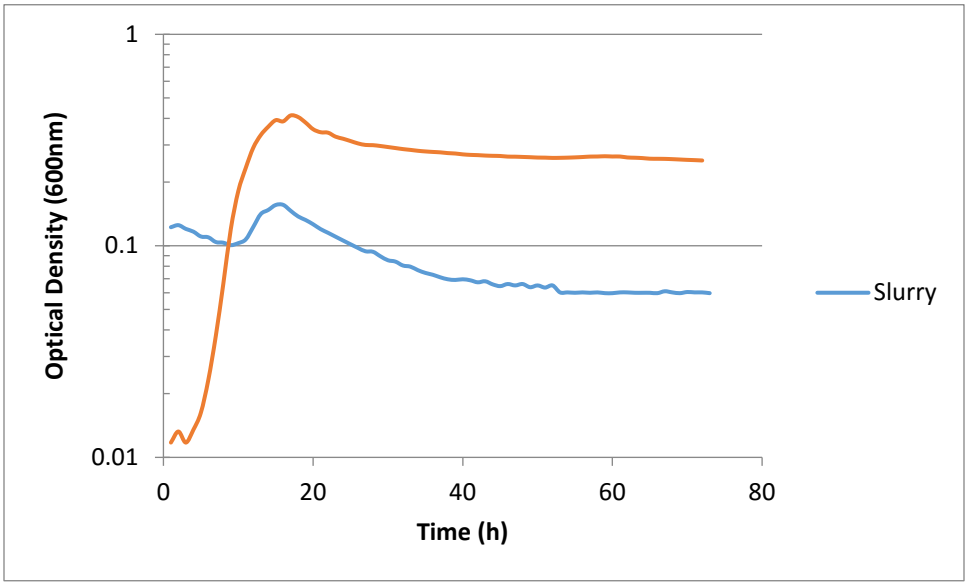

Figure 29: Growth curve of CA in BHI supplemented with optimal percentage of all 10 initially screened nutrients

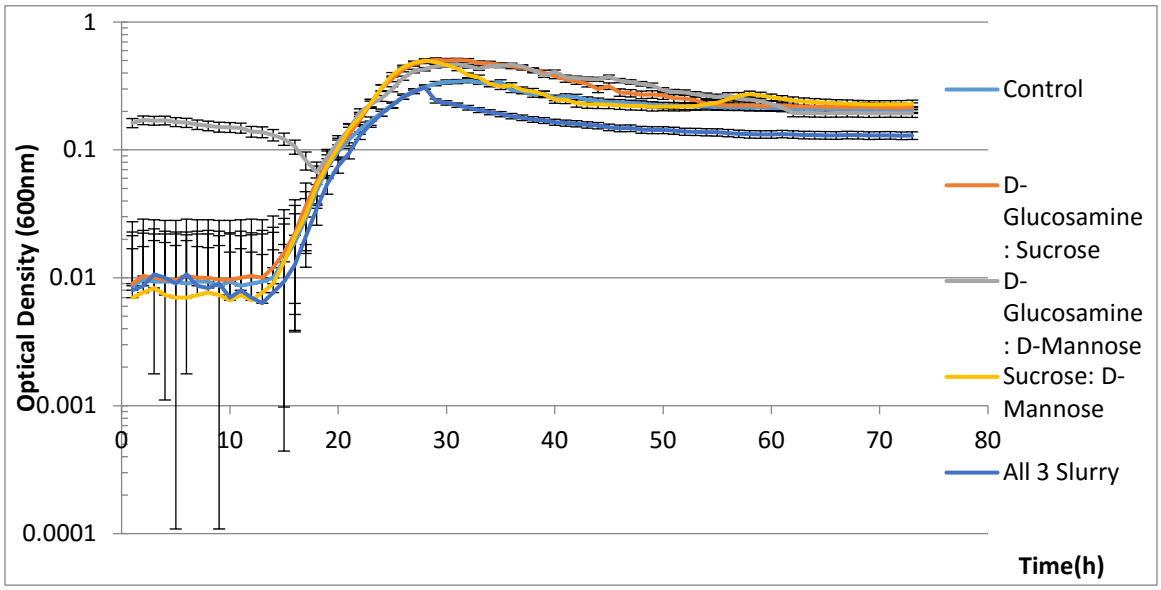

Figure 30: Growth curve of CA in BHI, all three statistically significant supplemented nutrient compound combinations vs control

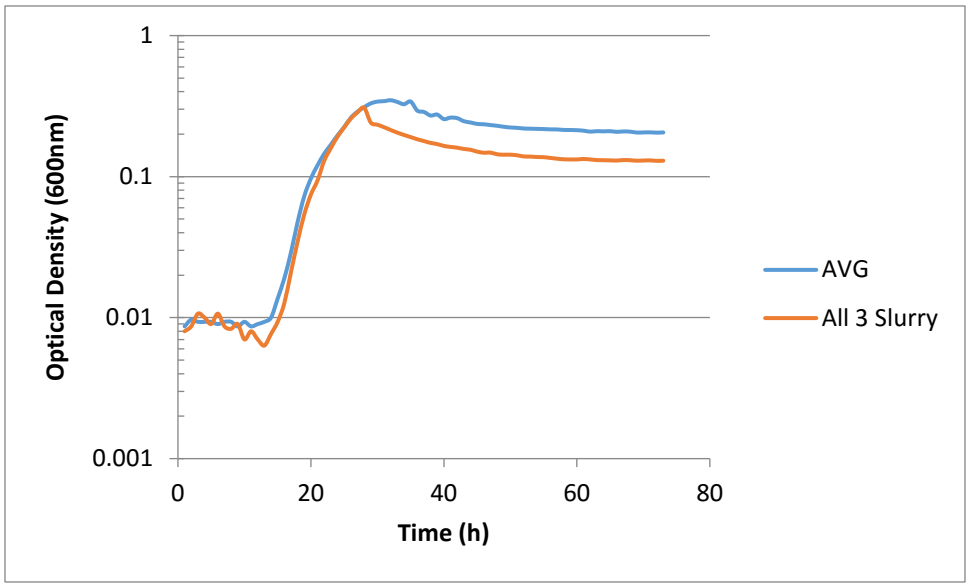

Figure 31: Growth curve of CA in BHI supplemented with 8% sucrose, 10% D-glucosamine, and 10% D-mannose

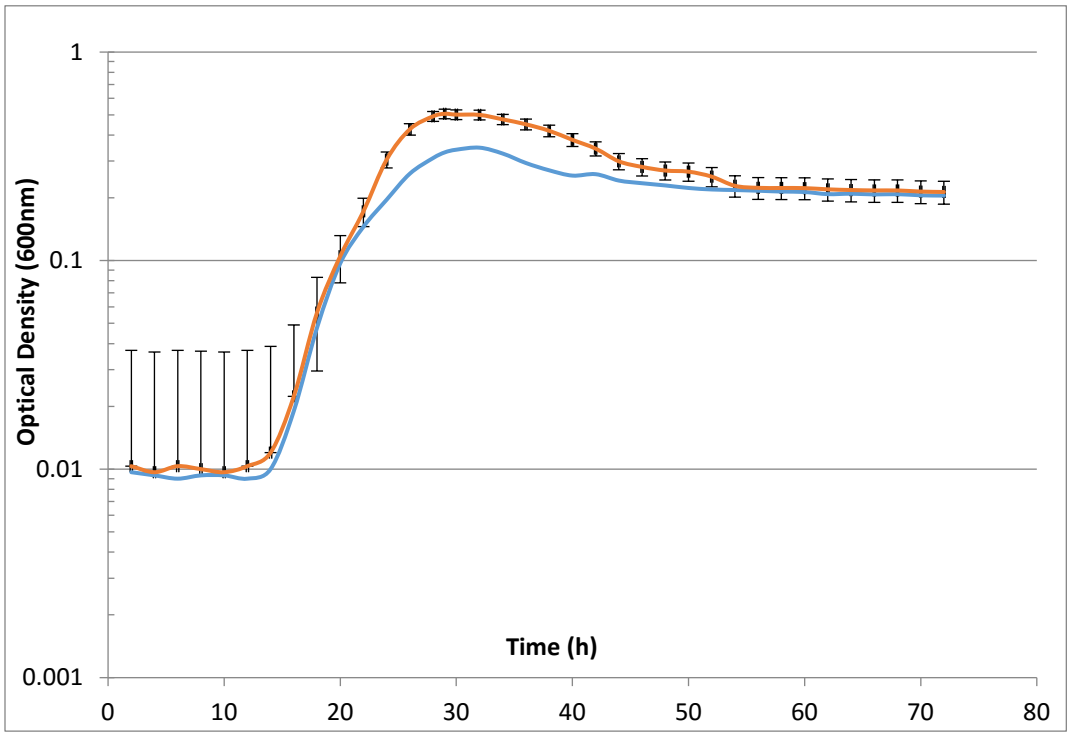

Figure 32: Growth of CA in BHI supplemented with 8% sucrose and 10% D-glucosamine

4.4.16 Illumina Next-Generation Sequencing. Nucleotide distribution between Sample 1 and Sample 2 differed significantly. Sample 1 had a GC content of 70.4% while Sample 2 had a GC content of 48.9%. The average contig measurement including scaffolding regions contains 713 base pairs in Sample 1 while Sample 2 has a slightly longer average of 862 base pairs. The paired reads distance distribution for each sample can be found in Appendix D along with the contents of the official report for each sample from Qiagen ®. Histograms are provided as a visual. The paired reads distance distribution for Sample 1 is slightly less defined than Sample 2 meaning that in Sample 2, there are more paired reads occurring at a certain base pair length.

Unfortunately, the results from samples sent in for Illumina next generation sequencing did not pertain to CA. Samples sent in for Illumina next generation sequencing contained an abundance of contigs that did not match up to anything indicating that media material may have gotten into the spin filter during DNA isolation from cells grown on agar plates. Samples sent in for Illumina did not detect CA, but did detect *Bacillus* spp. and *Achromobacter*. Additional samples will be sent in to re-run and obtain new results and fluorescent *in situ* hybridization experiments were completed to determine presence of CA, *Bacillus* spp., and *Achromobacter* spp. in samples.

4.4.17 Fluorescent In-Situ Hybridization. Figures 33-37 provide visual images for FISH experimental results. Although images are dim, differences in signal intensity were observed. In Figure 33, image A is a negative control using an active current culture with no hybridization probes in the hybridization solution then counterstained with DAPI. Images A-E remain the same for Figures 33-35. Image B contains current culture hybridized with the SFB 16S rRNA probe, counterstained with DAPI. Image C contains frozen culture stock cells hybridized with the SFB 16S rRNA probe, counterstained with DAPI. Image D contains current culture hybridized with the SFB, *Bacillus* sp., and *Achromobacter* probes, counterstained with DAPI. Image E contains frozen culture stock cells hybridized with the SFB, *Bacillus* spp., and *Achromobacter* probes, counterstained with DAPI. Images A-E utilized the DAPI filter of the Nikon Eclipse 90i. Image F

contains the negative control using the same parameters but captured using the GFP filter. This image serves as a control for images G & H of Figure 33. Image G contains actively growing current culture hybridized with all three probes and imaged using the GFP filter to capture fluorescence of the Alexa488 fluorophore attached to the *Bacillus* spp. 16S rRNA probe. Image H contains frozen culture hybridized with all 3 probes also imaged using the GFP filter. No significant results were obtained from Figure 33 other than a slight increase in signal intensity between Image F and G. A 1.18:1 signal ratio was observed between these two images demonstrating that the culture may contain Bacillus sp. but results are not conclusive enough to tell. However, when the mean signal ratios are used, a 6.5:1 ratio is observed. The numeric results for analysis of FISH images using Fiji (Schindelin et al., 2012) can be found below in Table 11.

Figure 34 contains the same Images A-E in its first row. In the second row, image I contains a negative control fixed to the same parameters outlined above but imaged using the DS-RED filter associated with the Alexa546 fluorophore attached to the *Achromobacter* 16S probe. Image J contains current culture hybridized with all 3 probes and imaged using the DS-RED filter. Image K contains frozen culture hybridized with all 3 probes also imaged using the DS-RED filter. The current culture in Image K had the same signal intensity as in control in Image J yielding a 1:1 signal ratio. The mean signal ratio was even lower at 0.98:1. Numeric results for FISH image analysis using Fiji (Schindelin et al., 2012) can be found below in Table 11. This warrants further investigation of

the potential contamination of the SFB culture with *Achromobacter* however it does not seem probable.

Figure 35 contains Images A-E from Figure 31 in the first row. In the second row, Image L represents the negative control that was captured using the CY5 filter in order to create a control for the CY5-labled SFB 16S probe. Image M contains current culture hybridized with the SFB 16S probe, imaged using the CY5 filter. Image N contains frozen culture hybridized with the SFB 16S probe also imaged using the CY5 filter. Image O contains current culture hybridized with all 3 probes imaged with the CY5 filter. Image P contains frozen culture hybridized with all 3 probes. The most significant findings when conducting FISH for this project were found between Image L and Image M. The negative control for CY5 and the current culture in Image M had an 11.16:1 signal ratio between the two. The current culture captured in the CY5 filter had a signal intensity of 67 and the negative control filtered in CY5 had a maximum signal intensity of 6. The mean signal ratio was slightly lower, at 9.13:1. Numeric results for FISH analysis using Fiji (Schindelin et al., 2012) can be found below in Table 11. Figure 36 shows a comparison of Images L and M on a larger scale. Images N, O, and P also contained positive results. Image N containing frozen culture and labeled only by the SFB 16S probe had a maximum signal ratio of 1.83:1 and a mean signal ratio of 6.21:1. Image O containing current culture and labeled by all 3 probes had a maximum signal ratio of 4.83:1 and a mean signal ratio of 7.38:1. Image P containing frozen culture and labeled with all 3 probes had a maximum

signal ratio of 1.33:1 and a mean signal ratio of 1.25:1. These results are listed in Table 11 below.

In order to relate the signal ratios for each probe, a pie chart was constructed for the maximum and mean signal ratios for both the frozen culture and the current culture samples labeled with all 3 probes. Figure 37 displays the 4 charts below. As can be seen in Chart A containing the maximum signals for the current culture, 69% of signal pertains to signal emitting from SFB 16S probe, 14% from the *Achromobacter* 16S probe, and 17% from the *Bacillus* spp. 16S probe. Chart B containing the maximum signal ratios for frozen culture labeled with all 3 probes shows that 46% of signal is coming from SFB 16S probe, 29% is coming from *Achromobacter* 16S probe, and 25% is coming from *Bacillus* spp. 16 s probe. In Chart C, the mean signal ratios for the current culture labeled with all probes shows that 50% of mean signal comes from the SFB 16S probe, 44% comes from the *Bacillus* sp. 16S probe, and 6% comes from the *Achromobacter* 16S probe. Chart D displays the mean signal ratios for the frozen culture labeled with all 3 probes demonstrating that 48% of signal is coming from the SFB 16S probe, 40% of signal is coming from the *Bacillus* spp. 16S probe, and 12% of signal is coming from the *Achromobacter* 16S probe.

FISH experiments were carried in order to determine SFB presence and potential contamination presence by *Achromobacter* and *Bacillus* spp. The high intensity signal found in a current culture sample demonstrates the presence of SFB in current cultures while the weak signals for *Bacillus* spp. and

Achromobacter require further investigation to prove their presence. The stronger maximum and mean signals of *Bacillus* spp. 16S probe vs *Achromobacter* 16S probe leads to the potential conclusion of SFB-*Bacilli* co-culturing in all samples since isolation from the turkey host, and also indicate that *Achromobacter,* a bacterium known for contamination of tissue culture medium may have been introduced into the SFB-*Bacilli* co-culture as determined through FISH analysis (Figure 37).

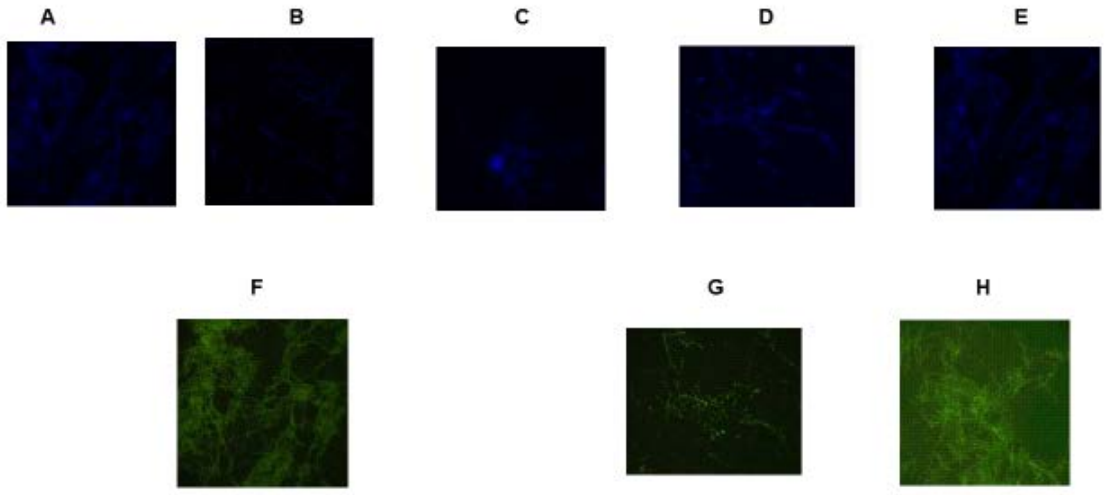

Figure 33: FISH imaging of control DAPI vs. GFP filters using SFB, *Achromobacter*, and *Bacillus* spp. 16S probes

1. Images A-E (left to right): negative control stained with DAPI, current culture*, frozen culture*, current culture**, frozen culture**
2. Images F-G second row (left-right): negative control using GFP filter, current culture**, frozen culture**
*hybridized with SFB 16S probe
** hybridized with SFB, *Achromobacter*, Bacillus sp. 16S probes

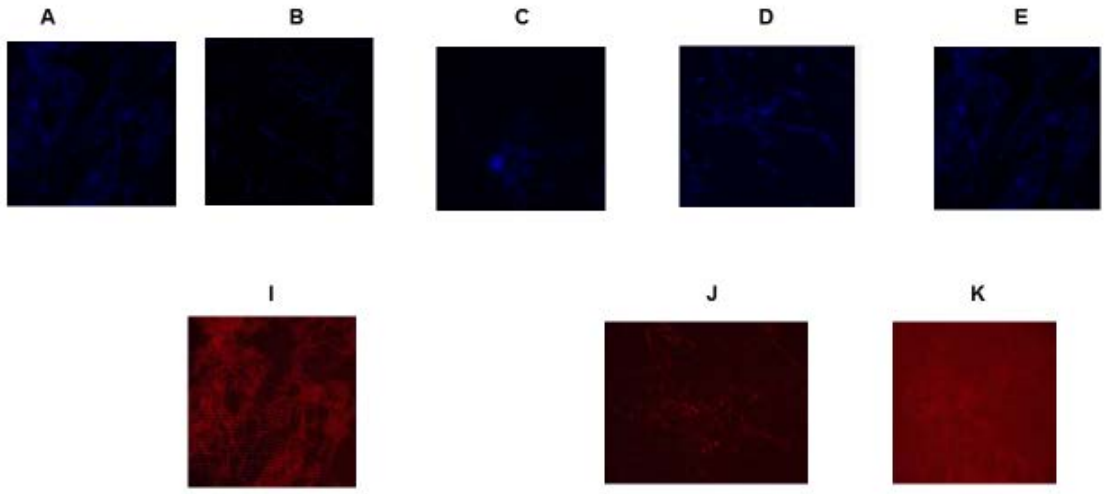

Figure 34: FISH imaging of control DAPI vs DS-RED filters using SFB, *Achromobacter*, and *Bacillus* spp. 16S probes

1. Images A-E (left to right): negative control stained with DAPI, current culture*, frozen culture*, current culture**, frozen culture**
2. Images I-K second row (left-right): negative control using DS-RED filter, current culture**, frozen culture**

*hybridized with SFB 16S probe

** hybridized with SFB, *Achromobacter*, Bacillus sp. 16S probes

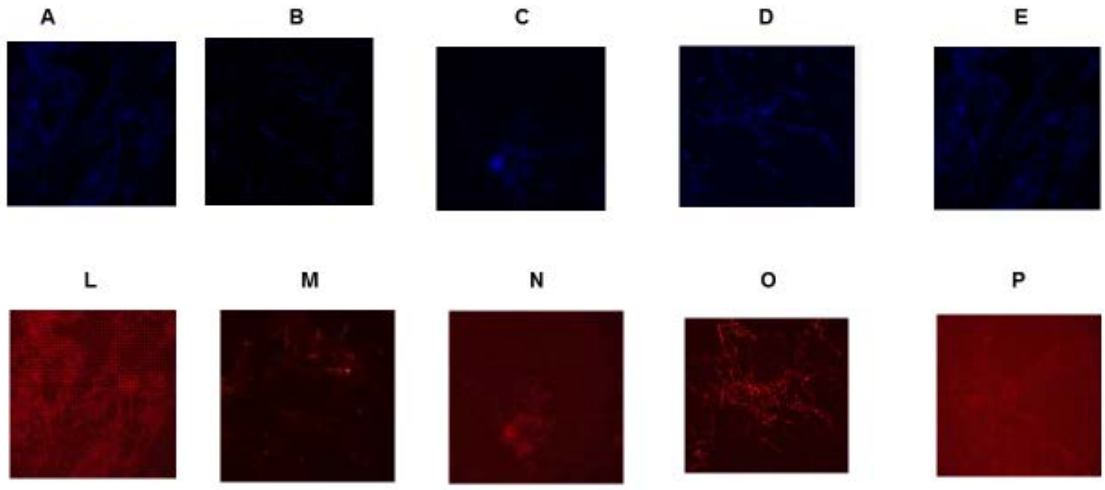

Figure 35: FISH imaging of control DAPI vs CY5 filters using SFB, *Achromobacter*, and *Bacillus* spp. 16S probes

1. Images A-E (left to right): negative control stained with DAPI, current culture*, frozen culture*, current culture**, frozen culture**
2. Images L-P second row (left-right): negative control using CY5 filter, current culture*, frozen culture*, current culture**, frozen culture**

*hybridized with SFB 16S probe

** hybridized with SFB, Achromobacter, Bacillus sp. 16S probes

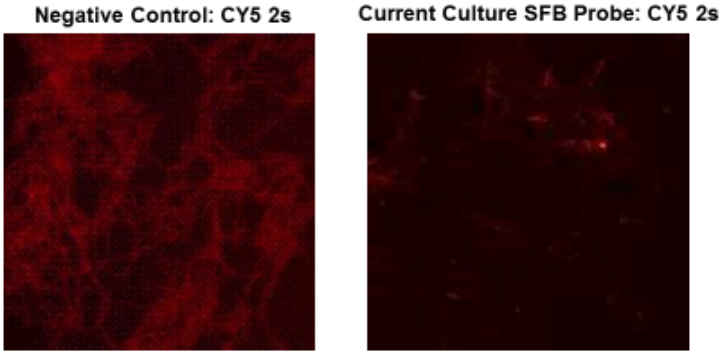

Figure 36: FISH imaging of negative control and current culture hybridized with SFB 16S probe using CY5 filter

Table 11: FISH analysis results using Fiji application

Image Name	Pixel Area	Mean Signal	Standard Deviation	Mode	Min Signal	Max Signal	Max signal:blank ratio	Mean Signal: blank ratio
All_CC_CY5_2s	8397	12.043	2.496	10	10	29	4.833333333	7.383813611
All_CC_DAPI_1s	16529	10	0	10	10	10	NA	NA
All_CC_DSRED_700ms	6	10.167	0.408	10	10	11	1	0.983934966
All_CC_GFP_400ms	155	9.716	0.931	9	9	13	1.181818182	6.490313961
All_ff_CY5_2s	4194304	2.037	0.688	2	0	8	1.333333333	1.248927039
All_ff_DAPI_1s	1833433	12.823	3.431	11	10	86	NA	NA
All_ff_DSRED_700ms	4194304	3.089	0.625	3	0	9	0.818181818	0.298945127
All_ff_GFP_400ms	4194304	1.53	0.697	1	0	8	0.727272727	1.022044088
Blank_CC_DAPI_1s	484326	18.23	10.108	11	10	87	NA	NA
BLANK_ff_CY5_2s	4194304	1.631	0.762	1	0	6	NA	NA
BLANK_ff_DAPI_1s	1006483	15.874	5.143	11	11	86	NA	NA
BLANK_ff_DSRED_700ms	6	10.333	0.516	10	10	11	NA	NA
BLANK_ff_GFP_400ms	4194304	1.497	0.818	1	0	11	NA	NA
SFB_CC_CY5_2s	4179	14.893	10.164	10	10	67	11.16666667	9.131207848
SFB_CC_DAPI_1s	192488	18.894	11.544	11	10	87	NA	NA
SFB_ff_CY5_2s	8	10.125	0.354	10	10	11	1.833333333	6.207847946
SFB_ff_DAPI_1s	157377	16.332	8.333	11	10	73	NA	NA

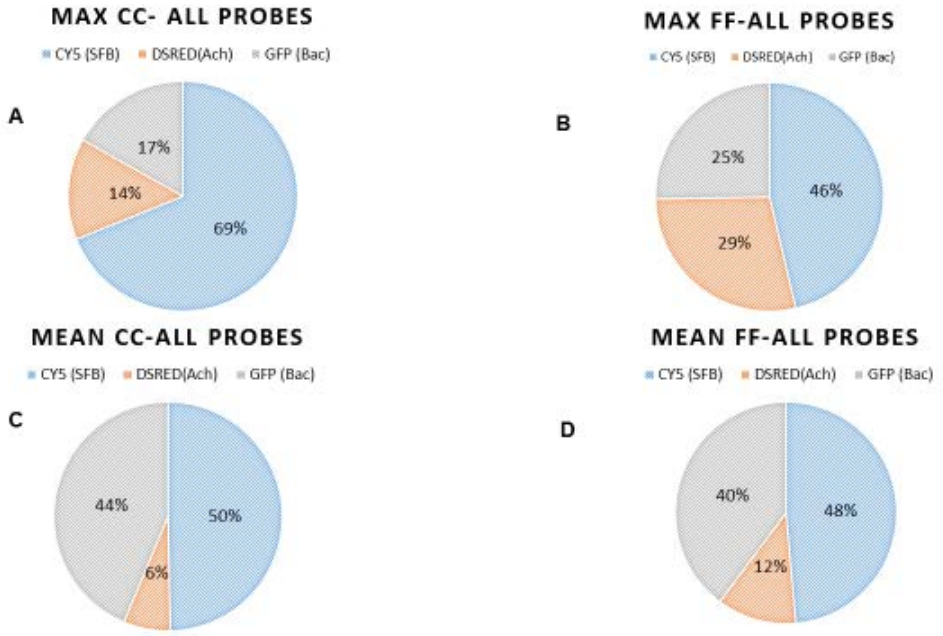

Figure 37: Maximum and mean signal ratio comparisons for *Bacillus* spp., *Achromobacter*, and SFB 16S probes

1. Charts A&B display maximum signal ratios
2. Charts C&D display mean signal ratios
3. All charts contain samples fluoresced with all 3 probes
4. "CC" signifies current culture
5. "FF" signifies frozen culture

***4.5* Conclusion.** Using 16S rRNA and Sanger gene sequencing, results support the isolation and growth of viable cells of *C.* arthromitus in BUG, BUG+B, and BHI media. 16S rRNA metagenome sequencing is the next step in order to determine the proportions of these three organisms in of CA samples. Results do not support the isolation of a pure culture of CA. All results of isolated CA refer to CA isolated along-side a *Bacillus* spp. While CA was detected through PCR, and increases in CA in growth medium is supported by the qPCR results, FISH experiments were completed as supporting evidence that CA exists in samples. FISH experiments successfully supported the presence of CA in samples as the signal intensity for the SFB 16S probe was much stronger than that of the

Achromobacter or *Bacillus* spp. 16S probes. Background noise probe signal ratios were much larger as well. qPCR was completed using the University of Minnesota's Expression Analysis Center in order to obtain accurate results. The qPCR assay run on the ABI 7900 HT supports the growth of CA in days 1-4 DNA isolated samples from 10% D-glucosamine/8% sucrose supplemented in BHI broth.

Future studies will work to separate these two cultures and determine if CA can grow in a lab medium as a monoculture. Additional studies will also work on further supporting evidence of CA's out of host viability, and work to further optimize growth yields and new culturing methods. Once the separation of the co-culturing organisms has been explored, future studies will be conducted in order to determine an effective way to optimize growth of isolated CA cells (or co-culturing system if they cannot survive on their own) in large quantities for further investigation of CA effects in young poults. In addition, experiments to determine the best mode of deliverance of this CA "probiotic" to poults is warranted. Some ideas that will be explored are: pelletizing, encapsulation, and freeze drying of cells for addition to turkey feed.

Overall, the content of this thesis embodies the trial and error involved while working on a very novel topic that has stymied scientists for a long time. Although many experiments were unsuccessful, the ones that were successful provide many new exciting research directions.

***4.6* Future Work.** The difficulty of culturing *C.* arthromitus has been emphasized by the extensive work of this project. With a lack of adequate sources and protocols to start from, a larger majority of materials and methods accessed for completion of this project resulted in failure or troubleshooting. Through troubleshooting, however, a new isolation and culturing methods have been discovered that is more simplistic and repeatable allowing for additional studies focusing on CA to ensue in the future.

Also, attention should be paid to potential symbiotic relationships between organisms of the gastrointestinal tract. Single species isolation of CA was not achievable using the newly developed culturing method. This warrants further studies focusing on the relationship of *Bacillus* spp. in the turkey gut microbiome and CA.

The major findings from this study are the multiple types of supporting evidence collected pertaining to CA's out of host viability, use of GEMs for new inquiry, and promising leads to optimize growth in a new culture medium. The diversity of data generated supports the detection, isolation, and growth of CA on both a molecular and a visual basis. However, because a co-culturing system has been created experiments generating growth data are limited in their usefulness and future studies will ascertain the metabolic capabilities of the *Bacillus* spp. member if monoculture is feasible. Additional work should be done to determine the reason behind the symbiotic relationship i.e. what might the *Bacillus* spp. be supplying CA with and how can this be supplied *in vitro.*

Determining the potential reason for the reliability on *Bacillus* spp. may help to further isolate CA into a pure monoculture and thus carry out growth phenotypic experiments (Biolog ™ and batch growth assays) without concern for what is contributing to the indicator of growth.

One major concern is that CA may not be capable of robust *in vitro* growth as an isolated organism. The co-culturing system with *Bacillus* spp. may be contributing to the ease of growth in new media. A reliable culturing method is important in order for future steps to take place. Understanding the importance of the relationship between CA and *Bacillus* spp. is invaluable to the future studies involving CA as more work is done on this exciting microorganism and more isolates of CA from a wider variety of vertebrate hosts including humans.

Sources Cited.

AMANN, R. I., BINDER, B. J., OLSON, R. J., CHISHOLM, S. W., DEVEREUX, R. & STAHL, D. A. 1990. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl Environ Microbiol,* 56**,** 1919-25.

BAEZ-SALDANA, A., DIAZ, G., ESPINOZA, B. & ORTEGA, E. 1998. Biotin deficiency induces changes in subpopulations of spleen lymphocytes in mice. *Am J Clin Nutr,* 67**,** 431-7.

BAUMLER, D. J., PEPLINSKI, R. G., REED, J. L., GLASNER, J. D. & PERNA, N. T. 2011. The evolution of metabolic networks of E. coli. *BMC Syst Biol,* 5**,** 182.

BERTELSEN, H., ANDERSEN, H. & TVEDE, M. 2011. Fermentation of D-Tagatose by Human Intestinal Bacteria and Dairy Lactic Acid Bacteria. *Microbial Ecology in Health and Disease,* 13.

BOLOTIN, A., DE WOUTERS, T., SCHNUPF, P., BOUCHIER, C., LOUX, V., RHIMI, M., JAMET, A., DERVYN, R., BOUDEBBOUZE, S., BLOTTIERE, H. M., SOROKIN, A., SNEL, J., CERF-BENSUSSAN, N., GABORIAU-ROUTHIAU, V., VAN DE GUCHTE, M. & MAGUIN, E. 2014a. Genome Sequence of "*Candidatus* Arthromitus" sp. Strain SFB-Mouse-NL, a Commensal Bacterium with a Key Role in Postnatal Maturation of Gut Immune Functions. *Genome Announc,* 2.

BOLOTIN, A., DE WOUTERS, T., SCHNUPF, P., BOUCHIER, C., LOUX, V., RHIMI, M., JAMET, A., DERVYN, R., BOUDEBBOUZE, S., BLOTTIÈRE, H. M., SOROKIN, A., SNEL, J., CERF-BENSUSSAN, N., GABORIAU-ROUTHIAU, V., VAN DE GUCHTE, M. & MAGUIN, E. 2014b. Genome Sequence of "*Candidatus* Arthromitus" sp. Strain SFB-Mouse-NL, a Commensal Bacterium with a Key Role in Postnatal Maturation of Gut Immune Functions. *Genome Announcements,* 2**,** e00705-14.

BORENSTEIN, E. 2012. Computational systems biology and in silico modeling of the human microbiome. *Brief Bioinform,* 13**,** 769-80.

CASELLI, M., HOLTON, J., BOLDRINI, P., VAIRA, D. & CALÒ, G. 2010. Morphology of segmented filamentous bacteria and their patterns of contact with the follicle-associated epithelium of the mouse terminal ileum. *Gut Microbes,* 1**,** 367-372.

CERF-BENSUSSAN, N. & GABORIAU-ROUTHIAU, V. 2010. The immune system and the gut microbiota: friends or foes? *Nat Rev Immunol,* 10**,** 735-744.

DANZEISEN, J. L., CALVERT, A. J., NOLL, S. L., MCCOMB, B., SHERWOOD, J. S., LOGUE, C. M. & JOHNSON, T. J. 2013. Succession of the turkey gastrointestinal bacterial microbiome related to weight gain. *PeerJ,* 1**,** e237.

DARLING, A. C., MAU, B., BLATTNER, F. R. & PERNA, N. T. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res,* 14**,** 1394-403.

DEL-POZO, J., CRUMLISH, M., TURNBULL, J. F. & FERGUSON, H. W. 2010. Histopathology and Ultrastructure of Segmented Filamentous Bacteria–Associated Rainbow Trout Gastroenteritis. *Veterinary Pathology Online,* 47**,** 220-230.

DESIGN, C. X. Designing XenServer Network Configurations. Citrix.

DEVOID, S., OVERBEEK, R., DEJONGH, M., VONSTEIN, V., BEST, A. A. & HENRY, C. 2013. Automated genome annotation and metabolic model reconstruction in the SEED and Model SEED. *Systems Metabolic Engineering: Methods and Protocols***,** 17-45.

DING, T., CASE, K. A., OMOLO, M. A., REILAND, H. A., METZ, Z. P., DIAO, X. & BAUMLER, D. J. 2016. Predicting Essential Metabolic Genome Content of Niche-Specific Enterobacterial Human Pathogens during Simulation of Host Environments. *PloS one,* 11**,** e0149423.

DONIA, M. S., CIMERMANCIC, P., SCHULZE, C. J., WIELAND BROWN, L. C., MARTIN, J., MITREVA, M., CLARDY, J., LININGTON, R. G. & FISCHBACH, M. A. 2014. A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. *Cell,* 158**,** 1402-14.

DUMONTIER, M. & HOGUE, C. W. 2002. NBLAST: a cluster variant of BLAST for NxN comparisons. *BMC Bioinformatics,* 3**,** 1-4.

EDGAR, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res,* 32**,** 1792-7.

ERICSSON, A. C., HAGAN, C. E., DAVIS, D. J. & FRANKLIN, C. L. 2014a. Segmented filamentous bacteria: commensal microbes with potential effects on research. *Comp Med,* 64**,** 90-8.

ERICSSON, A. C., TURNER, G., MONTOYA, L., WOLFE, A., MEEKER, S., HSU, C., MAGGIO-PRICE, L. & FRANKLIN, C. L. 2014b. Isolation of segmented filamentous bacteria from complex gut microbiota. *Biotechniques,* 59**,** 94-8.

FERGUSON, D. J. & BIRCH-ANDERSEN, A. 1979. Electron microscopy of a filamentous, segmented bacterium attached to the small intestine of mice from a laboratory animal colony in Denmark. *Acta Pathol Microbiol Scand B,* 87**,** 247-52.

FREILICH, S., ZARECKI, R., EILAM, O., SEGAL, E. S., HENRY, C. S., KUPIEC, M., GOPHNA, U., SHARAN, R. & RUPPIN, E. 2011. Competitive and cooperative metabolic interactions in bacterial communities. *Nat Commun,* 2**,** 589.

GLASS, E. M., WILKENING, J., WILKE, A., ANTONOPOULOS, D. & MEYER, F. 2010. Using the metagenomics RAST server (MG-RAST) for analyzing shotgun metagenomes. *Cold Spring Harbor Protocols,* 2010**,** pdb. prot5368.

GONG, J., SI, W., FORSTER, R. J., HUANG, R., YU, H., YIN, Y., YANG, C. & HAN, Y. 2007. 16S rRNA gene-based analysis of mucosa-associated bacterial community and phylogeny in the chicken gastrointestinal tracts: from crops to ceca. *FEMS Microbiology Ecology,* 59**,** 147-157.

GREENBLUM, S., CHIU, H. C., LEVY, R., CARR, R. & BORENSTEIN, E. 2013. Towards a predictive systems-level model of the human microbiome: progress, challenges, and opportunities. *Curr Opin Biotechnol,* 24**,** 810-20.

GREENBLUM, S., TURNBAUGH, P. J. & BORENSTEIN, E. 2012. Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. *Proceedings of the National Academy of Sciences,* 109**,** 594-599.

HE, G., SHANKAR, R. A., CHZHAN, M., SAMOUILOV, A., KUPPUSAMY, P. & ZWEIER, J. L. 1999. Noninvasive measurement of anatomic structure and intraluminal oxygenation in the gastrointestinal tract of living mice with spatial and spectral EPR imaging. *Proceedings of the National Academy of Sciences of the United States of America,* 96**,** 4586-4591.

HEINKEN, A., SAHOO, S., FLEMING, R. M. & THIELE, I. 2013. Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. *Gut Microbes,* 4**,** 28-40.

HENRY, C. S., DEJONGH, M., BEST, A. A., FRYBARGER, P. M., LINSAY, B. & STEVENS, R. L. 2010. High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nature biotechnology,* 28**,** 977-982.

IBRAHIM, M. & ANISHETTY, S. 2012. A meta-metabolome network of carbohydrate metabolism: interactions between gut microbiota and host. *Biochem Biophys Res Commun,* 428**,** 278-84.

JEPSON, M. A., CLARK, M. A., SIMMONS, N. L. & HIRST, B. H. 1993. Actin accumulation at sites of attachment of indigenous apathogenic segmented filamentous bacteria to mouse ileal epithelial cells. *Infect Immun,* 61**,** 4001-4.

KARLSSON, F. H., NOOKAEW, I., PETRANOVIC, D. & NIELSEN, J. 2011. Prospects for systems biology and modeling of the gut microbiome. *Trends in Biotechnology,* 29**,** 251-258.

KIM, T. & MUNDT, E. 2011. Metagenomic analysis of intestinal microbiomes in chickens. *High-throughput next generation sequencing: methods and applications***,** 185-194.

KLAASEN, H. L., KOOPMAN, J. P., POELMA, F. G. & BEYNEN, A. C. 1992. Intestinal, segmented, filamentous bacteria. *FEMS Microbiol Rev,* 8**,** 165-80.

KLITGORD, N. & SEGRE, D. 2010. Environments that induce synthetic microbial ecosystems. *PLoS Comput Biol,* 6**,** e1001002.

KLITGORD, N. & SEGRÈ, D. 2011. Ecosystems biology of microbial metabolism. *Current Opinion in Biotechnology,* 22**,** 541-546.

KOOPMAN, J. P., STADHOUDERS, A. M., KENNIS, H. M. & DE BOER, H. 1987. The attachment of filamentous segmented micro-organisms to the distal ileum wall of the mouse: a scanning and transmission electron microscopy study. *Lab Anim,* 21**,** 48-52.

KRIEGEL, M. A., SEFIK, E., HILL, J. A., WU, H. J., BENOIST, C. & MATHIS, D. 2011. Naturally transmitted segmented filamentous bacteria segregate with diabetes protection in nonobese diabetic mice. *Proc Natl Acad Sci U S A,* 108**,** 11548-53.

KRISHNAN, S., ALDEN, N. & LEE, K. 2015. Pathways and functions of gut microbiota metabolism impacting host physiology. *Curr Opin Biotechnol,* 36**,** 137-145.

KUMAR, S., NEI, M., DUDLEY, J. & TAMURA, K. 2008. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in bioinformatics,* 9**,** 299-306.

KUWAHARA, T., OGURA, Y., OSHIMA, K., KUROKAWA, K., OOKA, T., HIRAKAWA, H., ITOH, T., NAKAYAMA-IMAOHJI, H., ICHIMURA, M., ITOH, K., ISHIFUNE, C., MAEKAWA, Y., YASUTOMO, K., HATTORI, M. & HAYASHI, T. 2011. The lifestyle of the segmented filamentous bacterium: a non-culturable gut-associated immunostimulating microbe inferred by whole-genome sequencing. *DNA Res,* 18**,** 291-303.

LU, Y., LEVIN, G. V. & DONNER, T. W. 2008. Tagatose, a new antidiabetic and obesity control drug. *Diabetes Obes Metab,* 10**,** 109-34.

MARGULIS, L., JORGENSEN, J. Z., DOLAN, S., KOLCHINSKY, R., RAINEY, F. A. & LO, S. C. 1998. The Arthromitus stage of Bacillus cereus: intestinal symbionts of animals. *Proc Natl Acad Sci U S A,* 95**,** 1236-41.

MARTIN, F.-P. J., SPRENGER, N., YAP, I. K. S., WANG, Y., BIBILONI, R., ROCHAT, F., REZZI, S., CHERBUT, C., KOCHHAR, S., LINDON, J. C., HOLMES, E. & NICHOLSON, J. K. 2009. Panorganismal Gut Microbiome−Host Metabolic Crosstalk. *Journal of Proteome Research,* 8**,** 2090-2105.

MORISHITA, T. Y., LAM, K. M. & MCCAPES, R. H. 1992. Research note: isolation of two filamentous bacteria associated with enteritis in turkey poults. *Poult Sci,* 71**,** 203-7.

MUSSO, G., GAMBINO, R. & CASSADER, M. 2011. Interactions between gut microbiota and host metabolism predisposing to obesity and diabetes. *Annu Rev Med,* 62**,** 361-80.

NAVIN, N., GRUBOR, V., HICKS, J., LEIBU, E., THOMAS, E., TROGE, J., RIGGS, M., LUNDIN, P., MANER, S., SEBAT, J., ZETTERBERG, A. & WIGLER, M. 2006. PROBER: oligonucleotide FISH probe design software. *Bioinformatics,* 22**,** 2437-8.

NOGUCHI, H., PARK, J. & TAKAGI, T. 2006. MetaGene: prokaryotic gene finding from environmental genome shotgun sequences. *Nucleic Acids Res,* 34**,** 5623-30.

OWCZARZY, R., TATAUROV, A. V., WU, Y., MANTHEY, J. A., MCQUISTEN, K. A., ALMABRAZI, H. G., PEDERSEN, K. F., LIN, Y., GARRETSON, J., MCENTAGGART, N. O., SAILOR, C. A., DAWSON, R. B. & PEEK, A. S. 2008. IDT SciTools: a suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Res,* 36**,** W163-9.

PAMP, S. J., HARRINGTON, E. D., QUAKE, S. R., RELMAN, D. A. & BLAINEY, P. C. 2012. Single-cell sequencing provides clues about the host interactions of segmented filamentous bacteria (SFB). *Genome Research,* 22**,** 1107-1119.

PRAKASH, T., OSHIMA, K., MORITA, H., FUKUDA, S., IMAOKA, A., KUMAR, N., SHARMA, V. K., KIM, S. W., TAKAHASHI, M., SAITOU, N., TAYLOR, T. D., OHNO, H., UMESAKI, Y. & HATTORI, M. 2011. Complete genome sequences of rat and mouse segmented filamentous bacteria, a potent inducer of th17 cell differentiation. *Cell Host Microbe,* 10**,** 273-84.

RANJAN, P. The DOE Systems Biology Knowledgebase (KBase): Progress Toward a System for Inference and Modeling of Biological Function in Plants. Plant and Animal Genome XXIII Conference, 2015. Plant and Animal Genome.

SCHELLENBERGER, J., QUE, R., FLEMING, R. M. T., THIELE, I., ORTH, J. D., FEIST, A. M., ZIELINSKI, D. C., BORDBAR, A., LEWIS, N. E., RAHMANIAN, S., KANG, J., HYDUKE, D. R. & PALSSON, B. O. 2011. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat. Protocols,* 6**,** 1290-1307.

SCHINDELIN, J., ARGANDA-CARRERAS, I., FRISE, E., KAYNIG, V., LONGAIR, M., PIETZSCH, T., PREIBISCH, S., RUEDEN, C., SAALFELD, S. & SCHMID, B. 2012. Fiji: an open-source platform for biological-image analysis. *Nature methods,* 9**,** 676-682.

SCHNUPF, P., GABORIAU-ROUTHIAU, V. & CERF-BENSUSSAN, N. 2013. Host interactions with Segmented Filamentous Bacteria: An unusual trade-off that drives the post-natal maturation of the gut immune system. *Seminars in Immunology,* 25**,** 342-351.

SCHNUPF, P., GABORIAU-ROUTHIAU, V., GROS, M., FRIEDMAN, R., MOYA-NILGES, M., NIGRO, G., CERF-BENSUSSAN, N. & SANSONETTI, P. J. 2015. Growth and host interaction of mouse segmented filamentous bacteria in vitro. *Nature,* 520**,** 99-103.

SCZESNAK, A., SEGATA, N., QIN, X., GEVERS, D., PETROSINO, JOSEPH F., HUTTENHOWER, C., LITTMAN, DAN R. & IVANOV, IVAYLO I. The Genome of Th17 Cell-Inducing Segmented Filamentous Bacteria Reveals Extensive Auxotrophy and Adaptations to the Intestinal Environment. *Cell Host & Microbe,* 10**,** 260-272.

SHOAIE, S., KARLSSON, F., MARDINOGLU, A., NOOKAEW, I., BORDEL, S. & NIELSEN, J. 2013. Understanding the interactions between bacteria in the human gut through metabolic modeling. *Sci. Rep.,* 3.

SNEL, J., HEINEN, P. P., BLOK, H. J., CARMAN, R. J., DUNCAN, A. J., ALLEN, P. C. & COLLINS, M. D. 1995. Comparison of 16S rRNA sequences of segmented filamentous bacteria isolated from mice, rats, and chickens and proposal of "*Candidatus* Arthromitus". *Int J Syst Bacteriol,* 45**,** 780-2.

SRIDHARAN, G. V., CHOI, K., KLEMASHEVICH, C., WU, C., PRABAKARAN, D., PAN, L. B., STEINMEYER, S., MUELLER, C., YOUSOFSHAHI, M., ALANIZ, R. C., LEE, K. & JAYARAMAN, A. 2014. Prediction and quantification of bioactive microbiota metabolites in the mouse gut. *Nat Commun,* 5**,** 5492.

TANNOCK, G. W., MILLER, J. R. & SAVAGE, D. C. 1984. Host specificity of filamentous, segmented microorganisms adherent to the small bowel epithelium in mice and rats. *Appl Environ Microbiol,* 47**,** 441-2.

THAXTON, J. P. & PARKHURST, C. R. 1976. Growth, Efficiency, and Livability of Newly Hatched Broilers as Influenced by Hydration and Intake of Sucrose. *Poultry Science,* 55**,** 2275-2279.

THIELE, I., HEINKEN, A. & FLEMING, R. M. 2013. A systems biology approach to studying the role of microbes in human health. *Curr Opin Biotechnol,* 24**,** 4-12.

TURNBAUGH, P. J., LEY, R. E., MAHOWALD, M. A., MAGRINI, V., MARDIS, E. R. & GORDON, J. I. 2006. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature,* 444**,** 1027-31.

YIN, Y., WANG, Y., ZHU, L., LIU, W., LIAO, N., JIANG, M., ZHU, B., YU, H. D., XIANG, C. & WANG, X. 2013. Comparative analysis of the distribution of segmented filamentous bacteria in humans, mice and chickens. *ISME J,* 7**,** 615-621.

Appendix A: RAST Metagenomics Functions Results Analysis

*Please double-click on link below to view data

RAST Compiled File_HR.xlsx

Appendix B: RAST Metagenomics Subsystems Analysis

*Please double-click on link below to view data

RAST Subsystems Turkey Excel.xlsx

Appendix C: Biolog ™ Comparison to *In-Silico* Predictions

*Please double-click on link below to view data


Biolog_Compiled_Analysis_CAPM1.pdf


Biolog_Compiled_Analysis_CAPM2.pdf


Biolog_Compiled_Analysis_CAPM3.pdf


Biolog_Compiled_Analysis_CAPM4.pdf


PM1-PM10_biolog.pdf

Appendix D: Illumina Next-Generation Sequencing

** Please double-click on links in order to view data


SFB_Old.pdf


SFB_New.pdf

Appendix E: PLOSone article

PLOSone article to include as appendix.p