Spring 2017

An exploration of the gut bacteria in a mouse model of autism spectrum disorder on a ketogenic diet

Laura M. Nee
Trinity College, Hartford Connecticut, laura.nee@trincoll.edu

TRINITY COLLEGE

AN EXPLORATION OF THE GUT BACTERIA IN A MOUSE MODEL OF AUTISM SPECTRUM DISORDER ON A KETOGENIC DIET

BY

LAURA M. NEE

A THESIS SUBMITTED TO
THE FACULTY OF THE DEPARTMENT OF BIOLOGY
IN CANDIDACY FOR THE BACCALAUREATE DEGREE
WITH HONORS IN BIOLOGY

DEPARTMENT OF BIOLOGY

HARTFORD, CONNECTICUT

5 MAY 2017

AN EXPLORATION OF THE GUT BACTERIA IN A MOUSE MODEL OF AUTISM SPECTRUM DISORDER ON A KETOGENIC DIET

BY

LAURA M. NEE

Honors Thesis Committee

Approved:

_____
Lisa-Anne Foster, Ph.D., Advisor

_____
Kathleen Archer, Ph.D.

_____
Claire Fournier, Ph.D.

Date: _____

ABSTRACT **4**

INTRODUCTION **5**

MATERIALS AND METHODS **12**

Sample selection 12
Polymerase Chain Reaction 13
Cloning of 16s rRNA to plasmids 13
Transformation of TOPO® vector 14
Plasmid minipreps 14
Sanger sequencing of 16s rRNA gene 15
Sequence analysis 15
Species selection for PCR screening 15
PCR primer design 16
Using PCR to screen for specific bacteria 16

RESULTS **17**

Sequence analysis 17
Species selection for PCR screening 17
PCR Screening primer design 19
PCR Primers 20

DISCUSSION **24**

ACKNOWLEDGMENTS **31**

LITERATURE CITED **32**

Abstract

Alterations in the gut bacteria presence and abundance have been noted in certain diseased states like autism spectrum disorders (ASD). The BTBR T+tf/J mouse strain exhibits the three core behavioral symptoms of ASD and is thus used as the model organism to study ASD in this study. Previous studies have shown that a high-fat, ketogenic diet significantly improves all three core symptoms of ASD. The aim of this study was to describe differences between the gut bacteria of healthy mice on standard and ketogenic diets as well as BTBR mice on standard and ketogenic diets. In particular, an unidentified species was found to be significantly present only in the healthy mice, another species was found only in the BTBR/standard diet group, and another only in the BTBR/ketogenic diet group. The identities of these three bacterial species could not be classified due to error in DNA sample preparation and time constraints, but a protocol was created to check the identities of these three species using PCR and gel electrophoresis. PCR primers were created for 10 bacterial species based on Finegold et al.'s work with the gut bacteria in autistic individuals. In the future, the polymerase chain reactions should be conducted on the original DNA samples based the groups in which they were present in each study and the sequencing defect should be worked out in order to classify the three species of interest in a more direct manner. This protocol would use PCR to help determine the identities of these three species of interest.

Introduction

The normal flora refers to the diverse array of microbial species that regularly inhabit mammalian skin and mucous membranes. Certain bacterial genera populate various regions of the human body at different stages of life and health (Davis, 1996). The composition of the normal flora is also influenced by environment, diet, and antibiotic history (Ardeshir et al., 2014). Certain diseases, like autism spectrum disorders (ASD), have been associated with an altered flora compared to healthy gastrointestinal floras. It has been shown that a ketogenic, high-fat diet significantly improves the behavioral symptoms of a mouse analog of ASD (Ruskin et al., 2013). The purpose of this study was to determine whether the gut flora of the mouse model of autism would also be altered by the ketogenic diet.

In humans, the gastrointestinal tract is a particularly bacteria-rich environment. Because the gut microbiome is implicit in many physiological functions, its role in diseased states has been widely studied (Schreiner, 2015). The gastrointestinal flora has been associated with many states of disease including cardiovascular disease, inflammatory bowel disease, obesity, and autism spectrum disorders (Adams et al., 2011 and Schreiner, 2015). Although there are numerous studies showing correlations of altered gut microbiome composition with these diseases, it is difficult to determine whether these differences in the flora (between healthy and diseased individuals) are the causative agents or effects of the diseased state; or even significant at all (Schreiner, 2015). This uncertainty stems from the fact that the microbiome differs between individuals and species based on various environmental and physiological factors, as well as the individual's stage of life.

The normal flora, or the microbiome, are organisms that maintain an important relationship with their host by assisting with metabolic processes and by outcompeting pathogenic bacteria for the host's resources (Davis, 1996 and Schreiner et al., 2015). A 1985 study demonstrated that mice whose normal flora was compromised by antibiotics were more susceptible to colonization by *P. aeruginosa*, a pathogenic organism, than were untreated mice (Hentges et al., 1985).

The use of germ free animals in laboratory experiments has further enhanced our understanding of the importance of the normal flora in terms of immunity and metabolism (Davis, 1996). Germ free animals are surgically delivered and immediately placed into a sterile environment so that the normal flora is unable to colonize the host. Germ free animals have exhibited decreased intestinal motility, reduced epithelial cell renewal rates, and a lack of immune stimulation as compared to animals with the normal micro-ecological system, indicating that the bacteria in and on the body are involved in these physiological processes (Davis, 1996).

In contrast to the normal flora, dysbiosis refers to an imbalance between the beneficial and pathogenic microbiota and their host, and is widely believed to be a factor in inflammatory bowel disease among the other diseases that are associated with altered gut flora presence (Adams et al., 2011 and Comito et al., 2014). The composition and the quantity of the bacteria may be changed in a state of dysbiosis. For example, species of the *Firmicutes* and *Bacteroides* genera appear to be present in reduced amounts in individuals with inflammatory bowel disease as compared with healthy individuals (Comito et al., 2014). This microbial imbalance activates the inflammatory immune response and causes inflammatory cytokines to be released,

potentially causing the symptoms of the disease (Comito et al., 2014). However, there is no true consensus on whether the alterations in the flora are a cause or result of inflammatory bowel disease.

Autism spectrum disorders (ASD) have also been linked to dysbiosis of the gastrointestinal flora. Individuals with an ASD can experience gastrointestinal issues like diarrhea, constipation, and bloating at a much higher rate and at an increased severity than rest of the population (Adams et al., 2011). The changes in the gut bacteria of individuals with ASD may be contributing to these gastrointestinal symptoms of ASD, or the imbalance of bacteria present may be due to other causes; for now its role in ASD remains uncertain. A 2005 study found that there are increased levels of certain *Clostridium* clusters I and II species (toxin-producing species) in children with ASD.

The gut bacteria are likely involved in the gastrointestinal symptoms of ASD but possibly the behavioral symptoms as well. The gastrointestinal distress, no matter its cause, may also be a factor leading to the behavioral symptoms of ASD. The gastrointestinal and behavioral symptoms of ASD typically present themselves simultaneously, indicating a clear connection between the two (Parracho et al., 2005). The unpleasantness of the gastrointestinal issues may increase the severity of the behavioral symptoms of ASD, in that the pain and discomfort may increase frustration and aggression, and decrease ability to concentrate and communicate effectively (Adams, 2011).

Even if the gut bacteria are not involved in the gastrointestinal issues, they may be contributing to the behavioral aspects of ASD by way of the gut-brain axis theory.

The gut-brain axis (GBA) refers to the bidirectional signaling between the gastrointestinal flora and the brain (Cryan and O'Mahony, 2010). A theory proposed by one study is that the toxins produced by the increased *Clostridium* levels communicate with afferent nerves, which factor into the autistic symptoms by communicating ineffectively with the central nervous system (Parracho et al., 2005).

The potential pathway for the communication from the gut bacteria to the central nervous system was identified by a 1998 study. The vagus nerve runs from the abdomen to the brainstem, allowing neurotoxins and other secretions produced by gut bacteria to affect the brain, specifically toxins like tetanus toxin, produced by *Clostridium tetani* (Bolte, 1998). This correlation between the gastrointestinal symptoms and the behavioral symptoms widely seen in ASD individuals makes the need to study the gastrointestinal microbiota of these individuals extremely important.

The development of a mouse analog for autistic symptoms has become an invaluable tool in studying ASD. The strain of mouse is the BTBR T+tf/J. These mice are scientifically accepted as experimental models that may be used to study ASD because they exhibit the core behavioral symptoms of autism: low sociability, repetitive behaviors, and communication deficits (McFarlane, 2007).

A 2013 study by Ruskin et al. used the BTBR strain to determine the effect of a ketogenic diet on the three aforementioned core symptoms of autism. Ketogenic diets are high in fat and low in carbohydrates and have been shown to dramatically improve seizures, a common comorbidity of ASD, in children and adults (Neal et al., 2008; Sirven et al., 1999). The ketogenic diet redirects the body's primary fuel from glucose to ketone bodies, thereby increasing blood ketone levels and mitochondrial functioning

while decreasing blood glucose levels (Ruskin et al., 2013). Ruskin et al. found that this diet type significantly improved all three of the core behavioral symptoms of BTBR mice (2013).

ASD are extremely prevalent disorders; currently 1 in 68 children suffer from them and the rate is rising (Autism Speaks, 2016). Individuals with ASD experience ongoing difficulties with social functioning and while behavioral therapy and diet restrictions help to assimilate individuals with ASD into society, they are not cures nor do they work for every individual. Our inability to cure ASD makes these disorders mysterious: researchers and the public alike are desperate to learn more about its causes and comorbidities, which makes the use of BTBR strain mice an indispensable tool.

A recent study also made use of the BTBR strain and the ketogenic diet to show that diet rather than genotype was the major force behind gut microbial changes and that the ketogenic diet decreases total microbial levels (Klein et al., 2016). The same study also identified *Clostridium leptum* as the bacterial species most involved in the host metabolism (Klein et al., 2016). A thesis study at Trinity College used BTBR mice and a ketogenic diet to observe differences in the gut flora between healthy and BTBR mice, and with or without a ketogenic diet (Labe, 2016). Both of these studies isolated bacterial DNA from mice fecal samples and used PCR to amplify the 16S rRNA gene (Klein et al. and Labe). The 16S rRNA gene is the most common genetic marker for studying bacterial phylogeny due to its presence in all bacterial species, conserved gene function, and variable regions that help differentiate genus and species (Janda and Abbott, 2007).

Based on previous work, there is an apparent connection between ASD and the gut flora and the ketogenic diet has shown promise in alleviating the core behavioral symptoms of ASD in BTBR mice. This study looks at the potential connection between the gut flora of BTBR mice and ketogenic diet, and continues the work of Trinity College alumna Shelby Labe. Labe isolated the bacterial DNA from fecal samples from healthy mice on standard diets, healthy mice on ketogenic diets, BTBR mice on standard diets, and BTBR mice on ketogenic diets. The introduction of the ketogenic diet increased the amount of bacterial species diversity in the fecal samples as compared to mice on the standard diet and increased the overall abundance of bacteria in the healthy mice only as compared to the BTBR mice.

Comparisons were made between the four experimental groups with tRFLP analysis, which showed the distinct (though unidentified) bacterial species present in the fecal DNA samples. There were 28 unidentified bacterial species that were common between at least two of the experimental groups. Ten species were present in all experimental groups, nine were present in the control group (healthy mice, standard diet), five were present in the BTBR mice on either diet, and one was present in both mice strains on the standard diet. There was a single bacterial species found each in the BTBR/ketogenic diet group, the BTBR/standard diet group, and the normal mice on both diets. These final three unidentified bacterial species are potentially significant because they were unique to only one of the mice groups.

Species M (arbitrarily labeled by Labe) was found in the healthy mice on both diets, but not significantly in the BTBR mice fecal samples (Figure 1). This occurrence indicates a possible significance in that its absence from the gut flora of the BTBR mice.

Since it was found significantly in the healthy mice, species M appears to a part of the normal flora. Since the BTBR mice groups are lacking this species, it may be important in normal gut-brain interactions.

Species G was present in BTBR mice on the standard diet and not significantly in the healthy mice fecal samples (Figure 1). This indicates its presence in the gastrointestinal tract is positively correlated with the BTBR mice, which display the autistic behavioral symptoms. Whether the correlation between the presence of this bacterium and the autistic symptoms is a cause or result of unhealthy gut-brain reactions or whether it is simply coincidence is unknown.

Species D was present in BTBR mice on the ketogenic diet, but not significantly present in the healthy mice (Figure 1). The BTBR mice on the ketogenic diet had significantly improved behavioral symptoms (Ruskin, 2013). This points to the possibility that species D may be a bacteria that is not a regular member of the normal flora, but may act like one or contribute in some way to re-stabilizing healthy gut-brain interactions from the state of dysbiosis.

Figure 1: Bacterial species present in mice experimental groups

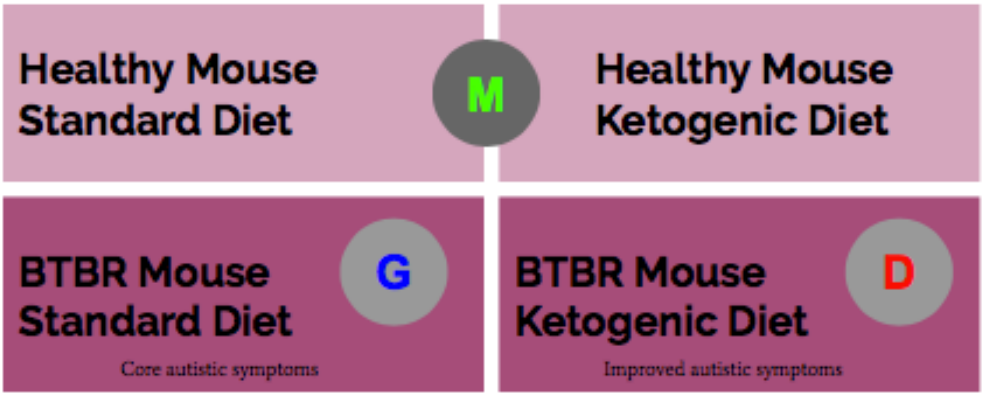

Fig. 1: Bacterial species M is significantly present in healthy mice, independent of diet type. Species G is significantly present in BTBR mice on the standard diet. Species D is significantly present in the BTBR mice on the ketogenic, high-fat diet.

The hope of this study was to discover what species D, G, and M are and to learn more about how they exist and metabolize in their host environment. Learning this information will help narrow down whether those species have the potential to affect gut-brain interactions that may be prevalent in autism spectrum disorders.

The objective of this study was to further our knowledge of the gastrointestinal flora in autism spectrum disorders, the gut-brain axis, and the effect of the ketogenic diet on ASD by identifying the presence of particular bacterial species in the gastrointestinal tract of a mice model of ASD as compared to healthy mice. Sequencing the 16s rRNA gene of species M, G, and D, along with further investigation and literature review, would classify the species and help to discover whether their metabolic products have an effect on gut-brain interactions or the gastrointestinal symptoms seen in ASD.

Materials and Methods

Sample selection

Bacterial DNA was isolated from fecal samples of BTBR mice on standard and ketogenic diets and normal mice on standard and ketogenic diets using MO BIO PowerFecal® DNA Isolation Kit (Labe, 2016). Labe used restriction enzyme *Hae*III to cut the isolated bacterial DNA and used Terminal Restriction Fragment Length Polymorphisms (tRFLP) to determine similarities and differences between the experimental mouse groups. DNA samples were selected for PCR amplification by determining which samples from each experimental group contained species D, G, and M as defined by Labe's thesis (2016). The D, G, and M bacterial DNA found in the mice

fecal samples were also chosen based on their concentration of DNA (found using the Trinity College Nanodrop) and their appearance on a 0.8% agarose gel. There were four mice fecal samples studied that were from BTBR mice on ketogenic diets (species D); four from BTBR mice on standard diets (species G); and four from healthy mice to study species M (two on standard diets, two on ketogenic diets).

Polymerase Chain Reaction

The section of DNA amplified was the 16s rRNA gene. The primers were unlabeled; the forward primer used was 27F- 5'AGAGTTTGATCCTGGCTCAG-3' and the reverse primer used was 926R-5'-CCGTCAATTCMTTTRAGTTT-3' (Figure 2). The polymerase chain reactions were placed in a thermocycler with settings for 95°C initial denaturation, 50°C primer annealing, and 72°C extension for 35 cycles. The results were checked on a 0.8% agarose gel. The QIAquick® PCR Purification Kit was used to remove unused primers from the PCR products. The Trinity College Biology Department Nanodrop was used to quantify the DNA present.

Figure 2: Universal primers

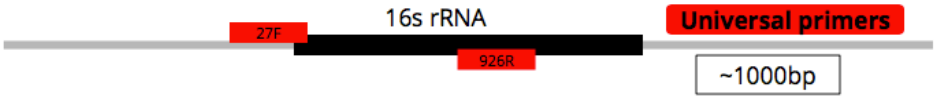

Fig. 2: The approximate binding locations of the 27F and 926R primers are shown in red on a model of the 16s rRNA gene. The expected PCR product size with these primers is approximately 1000 base pairs.

Cloning of 16s rRNA to plasmids

The cloning was done with the TOPO® TA Cloning® Kit for Sequencing. The purified PCR product was mixed with water, a salt solution and the TOPO® plasmid

vector. This reaction was incubated at room temperature for 5 minutes then incubated on ice.

Transformation of TOPO® vector

The cells transformed were One Shot R Mach1™ - T1® chemically competent *Escherichia coli* cells. A portion of the cloning reaction was put into a vial of the competent E. coli cells. Following a 5-minute incubation period on ice, the cells were heat shocked for 30 seconds and transferred immediately back to ice. The cells were shaken for 1 hour at 37°C.

Luria-Bertani broth (LB) plates with 50 μg/μL kanamycin were warmed in a 37°C incubator for 30 minutes prior to plating the transformation reactions. The kanamycin presence ensured that only cells with the plasmid vector would grow, due to a kanamycin resistance gene being present on the TOPO® vector, but not in the One Shot R Mach1™ - T1R cells. Ten microliters of the transformation reaction were diluted in 200 μL of LB, and 100 μL of this dilution was spread onto the pre-warmed plates.

The plates were incubated at 37°C for 8-12 hours. Isolated colonies from each transformation reaction were selected and grown further on a grid LB/kanamycin plate.

Plasmid minipreps

Isolated colonies were grown overnight in 5 mL of LB broth with 50 μg/μL kanamycin. The cultures were shaken at 37°C for approximately 12 hours at 80rpm. Four milliliters of the overnight cultures were used for each plasmid preparation. The

cells were pelleted by centrifuging at 8000 rpm for two minutes. The samples were then prepared using the QIAprep® Spin Miniprep kit.

Plasmid samples were concentrated using Savant Speed Vac®Plus (SC110A), then diluted in 20 μL of TE buffer. The DNA in the samples was quantified using the Trinity College Biology Department Nanodrop.

Sanger sequencing of 16s rRNA gene

The plasmid samples (n=44) were sent to the Yale University DNA Analysis Facility to be sequenced. There were 17 plasmid samples containing the 16s rRNA gene of species G, 17 plasmids samples containing the 16s rRNA gene of species M, and 10 plasmid samples containing the 16s rRNA gene of species D. The samples were placed into individual wells of a 96-well plate, at low concentrations of approximately 500-700 nanograms of DNA and high concentrations of approximately 700-1000 nanograms of DNA. The forward M13 primer, which has a site slightly upstream from the 16s rRNA integration site, was also put in the wells. The samples were to be analyzed using Sanger sequencing techniques.

Sequence analysis

Had the Yale Facility's Sanger sequencing been successful, the sequences would have been analyzed using the nucleotide Basic Local Alignment Search Tool (nBLAST) on the National Center for Biotechnology Information (NCBI).

Species selection for PCR screening

Since sequencing the 16s rRNA genes in the TOPO® vector was unsuccessful on two separate occasions, the next step was to search previous related studies for

bacterial species that could align with the characteristics of species D, G, and M. Once chosen, the 16s rRNA genes of the selected species were found using the Center for Microbial Ecology's Hierarchy Browser (Michigan State University, 2014).

PCR primer design

PCR analysis was attempted to determine whether the 16s rRNA gene segments in the plasmids were those of certain bacteria species deemed significant by previous studies. Using PCR amplification with primers specific to those significant bacteria species would determine whether those species were present in the samples I had prepared initially for sequencing. Primers were designed for 11 bacteria species. The primers were designed with help from the OligoAnalyzer 3.1 from Integrated DNA Technologies.

Using PCR to screen for specific bacteria

The PCR reactions would utilize the designed primers, *Taq* polymerase, dNTPs, and the previously prepared plasmids. The reactions would have been conducted in the Thermocycler for 30 rounds with the settings in Table 1.

Table 1: Polymerase Chain Reaction set up for specific primer pairs

Bacteria	**Denaturing Temp.**	**Annealing Temp.**	**Primer Extension Temp.**
C. boltea	95°C	51°C	72°C
C. leptum	95°C	54.7°C	72°C
C. perfringens	95°C	46.8°C	72°C
C. disporicum	95°C	47.5°C	72°C
C. tertium	95°C	48°C	72°C
B. longum	95°C	46.7°C	72°C
B. angulatum	95°C	48.5°C	72°C
R. ablus	95°C	52°C	72°C
D. piger	95°C	50.9°C	72°C
B. vulgatus	95°C	47.3°C	72°C

Unfortunately, testing for the presence of these species in the 44 plasmid samples was never completed due to time constraints. If there had been sufficient time to do the polymerase chain reactions, the PCR products would have been analyzed using a 0.8% agarose gel with a 1 kb (Table 1). The gel would display an approximate size of the PCR product that could be compared to the expected product size as determined by the primer selection.

Results

Sequence analysis

As noted in the "Methods," the 16s rRNA genes were unable to be sequenced by the Yale DNA Analysis Facilities, for reasons explored in the discussion section. Therefore, I was unable to analyze the sequences of the genes belonging to species D, G, and M.

Species selection for PCR screening

By delving into related studies done by researchers over the past decade or so, I was able to determine several bacterial species that could potentially be the identities of species D, G, and M. The species that have been previously tagged as being different between healthy and ASD individuals, and therefore potentially important in ASD, are shown in Table 2.

Table 2: Potentially important bacteria investigated

Present only in ASD individuals	**Source**	**Present only in healthy individuals**	**Source**
Bacteroides vulgatus	Finegold et al., 2010	*Clostridium disporicum*	Finegold et al., 2002
Clostridium boltea	Finegold et al., 2010	*Clostridium tertium*	Finegold et al., 2002
C. perfringens	Finegold et al., 2017	*Clostridium leptum*	Finegold et al., 2010
Desulfovibrio piger	Finegold et al., 2010	*Ruminococcus albus*	Finegold et al., 2002
----------	----------	*Bifidobacterium longum*	Finegold et al., 2010
----------	----------	*B. angulatum*	Finegold et al., 2010

Two possible bacterial pathways for affecting ASD are production of toxins and secretion of metabolic wastes (Finegold, 2002). A 2010 study by Finegold et al. identified *Clostridium boltea* as a species that is overabundant in the gut flora of children with ASD (Table 2). In the same study, *Bacteroides vulgatus* and *Desulfovibrio piger* were also found in a significantly greater abundance in the severely autistic group (Table 2). In a more recent study by Finegold et al., *Clostridium perfringens* was found to be significantly more abundant in autistic children than the control children (Table 2).

Clostridium is a known toxin-producing genus, which could contribute to species of it being in the ASD groups and not in the healthy groups (Fingegold et al., 2010). *Clostridium perfringens* in particular is known to produce at least four different toxins that have been shown to be implicit in other gastrointestinal diseases (Finegold et al., 2017). Both *Bacteroides* and *Desulfovibrio* species can produce lipopolysaccharide, a virulent factor not found in the gram-positive species seen in the gastrointestinal tract (Finegold et al., 2010). In addition, *Desulfovibrio* species can produce hydrogen sulfide, a metabolic waste, which could account for or contribute to some ASD symptoms (Finegold et al, 2010).

In the 2010 Finegold et al. study, *Bifidobacterium longum, Bifidobacterium angulatum*, and *Clostridium leptum* were found to be significantly more prominent in the healthy group than the ASD group (Table 2). *Clostridium* is a highly varied genus, which accounts for species of it being in both groups (Song et al., 2002). The *Bifidobacterium* genus produces exopolysaccharides, which are fermentable substrates that can be used by other members of the normal flora (Finegold et al., 2002). Another study showed that *Clostridium disporicum*, *Clostridium tertium, and Ruminococcus albus* were all missing from the guts of autistic children as well (Finegold et al., 2002; Table 2). *R. albus* seems to be an anomaly for the *Ruminococcus* genus, which is also a known toxin-producing genus not normally found in the normal flora (Finegold et al, 2002).

Each of the aforementioned species would have been a good candidate for screening the prepared plasmid samples using the PCR screening technique.

PCR Screening primer design

Primers were designed for the PCR amplification of the 16s rRNA genes of the species shown in Table 2. Conducting PCR with the DNA of species D, G, and M as the template DNA and the primers being specific to species already identified as being different between healthy and autistic groups would help to identify or rule out the bacteria in Table 2 as being the identities of species D, G, and M. Gel electrophoresing the PCR products and comparing the band size with the expected product size (determined by the primer set used) would allow this affirmation or denial of the species in Table 2 as being the identities of species D, G, and M.

Primers were designed based on sequence similarity to bacteria of the same genus, length, hairpin likelihood, self-dimer likelihood, heterodimer likelihood, melting temperature agreement, GC-content, GC-clamp presence, and product length (Table 3).

Table 3: Primer Design Criteria

Criterion	Value
Sequence similarity to bacteria in same genus	> 10% difference
Length	15-25 base pairs
Hairpin	$\Delta G < 3.0$
Self Dimer	$\Delta G < 5.0$
Heterodimer	$\Delta G < 6.0$
Melting temperature	Within 2°C of the other
GC Content	40-65%
Presence of GC Clamp	Full or partial
PCR Product Length	500-1500 base pairs

PCR Primers

The primers were designed for PCR on the freshly prepared plasmid samples, but it was difficult finding primers that were within the sequence that the original PCR produced with the universal primers. The forward universal primer, 27F, binds slightly upstream of the 16s rRNA so all of the designed primers were within its boundaries. The reverse universal primer, however, binds to the 800-900 base region of the gene. The 16s rRNA gene is approximately 1500 base pairs in total, so looking for primers in the region inside the reverse universal primer limited the options significantly. Each reverse primer that I designed lies outside of the segment created by the initial PCR, so the reactions would have had to be done on the original DNA samples, although they were older. The designed primers and their expected PCR product lengths are shown in Table 4.

Figure 3: Designed primers

Fig. 3: Unlike the universal primers (Fig. 2), the designed primers are specific to certain sequences in the 16s rRNA gene that correspond with a specific bacterial species. The designed primers bind to different parts of the 16s rRNA gene, and each primer set has a different PCR product size.

Table 4: Designed Primer Sequence and Product Length

Species	Forward Primer (5' →3')	Reverse Primer (5' →3')	Product Length
C. boltea	CCAAAGCGACGATCAGTAGC	GCGTTGCTGACTCCCATG	1097 bp
C. leptum	CCGCATAAGACCTCAGTACCGC	GGGATTTGCTTGCCTTCACAGGG	1042 bp
C. perfringens	CATCATTCAACCAAAGGAGC	CAAGGGATGTCAAGTGTAGG	743 bp
C. disporicum	CAGGGACGATAATGACGG	CCAGTTACGGGTAATTCAGG	514 bp
C. tertium	CCGCATAACATTACATTTTCGC	GGTTCTCCTACGGCTACC	1453 bp
B. longum	GGCACTTTGTGTTGAGTG	CTCGACTGCGTGAAGG	837 bp
B. angulatum	CTGGGAAAGATTTTATCGGTATGG	CGCCTTCATGGAGTCG	1079 bp
R. ablus	CGAGCGAAAGAGTGCTTGC	CGACTGCTTCCTCCTTGC	1332 bp
D. piger	GCGGCGTGCTTAACACAT	CCTCACGGTATCGCTGC	1204 bp
B. vulgatus	CGTCTACTCTTGGACAGC	CGATGGCAGTCTTGTCAG	986 bp

The DNA of species G and D would have been the template DNA in polymerase chain reactions done with the primers designed for *Clostridium boltea, Bacteroides vulgatus, Desulfovibrio piger,* and *Clostridium perfringens* (Table 4). The characteristics of the primer sets specific to these four bacteria are shown in Table 5.

Table 5: Primer characteristics for bacteria previously seen in ASD individuals

Species	*Clostridium boltea*	*Bacteroides vulgatus*	*Desulfovibrio piger*	*Clostridium perfringens*
Sequence difference %	15% from *C. disporicum*	33.33% from *B. pyogenes*	0% from *D. intestinalis*	35% from *C. disporicum*
GC %, clamp	55%, yes	55.6%, yes	55.6%, yes	45%, partial
Length	20 bp	18 bp	18 bp	20 bp
Hairpin	ΔG=-1.07	ΔG=-0.75	ΔG=-1.84	ΔG=-0.19
Self Dimer	ΔG=-4.62	ΔG=-3.61	ΔG=-4.85	ΔG=-3.14
Melting Temp	56.0°C	52.3°C	56.6°C	51.8°C
Sequence difference	27.8% from *C. disporicum*	22.22% from *B. pyogenes*	17.6% from *D. intestinalis*	20% from *C. disporicum*
GC %, clamp	61.1%, yes	55.6%, yes	64.7%, yes	50%, yes
Length	18 bp	18 bp	17 bp	20 bp
Hairpin	ΔG=0.22	ΔG=-0.19	ΔG=-0.78	ΔG=0.64
Self Dimer	ΔG=-5.38	ΔG=-3.61	ΔG=-3.61	ΔG=-1.95
Melting temp	56.7°C	53.2°C	55.9°C	52.7°C
Hetero-dimer	ΔG=-5.31	ΔG=-4.87	ΔG=-6.91	ΔG=-5.00

Table 5: The grey rows indicate the primer characteristics of the forward primers of the species indicated. The blue rows indicate the primer characteristics of the reverse primers of the species indicated. The final row indicates the primer characteristic that is dependent on both the forward and reverse primer.

The DNA of species M and D would have been the template DNA in polymerase chain reactions done with the primers designed for *Bifidobacterium longum, Bifidobacterium angulatum, Clostridium leptum, Clostridium disporicum, Clostridium tertium,* and *Ruminococcus albus* (Table 4). The characteristics of the primer sets specific to these six bacteria are shown in Table 6.

Table 6: Primer characteristics for bacteria previously seen in non-ASD individuals

Species	*B. longum*	*B. angulatum*	*C. leptum*	*C. disporicum*	*C. tertium*	*R. ablus*
Sequence difference %	27.8% from *B. angulatum*	41.7% from *B. longum*	32% from *C. tertium*	17% from *C. perfringens*	32% from *C. leptum*	42.1% from *R. flaveflaciens*
GC %, clamp	50%, yes	41.7%, partial	59%, yes	55.6%, partial	40.9%, yes	57.9%, yes
Length	18 bp	24 bp	22 bp	18 bp	22 bp	19 bp
Hairpin	ΔG=-1.29	ΔG=0.45	ΔG=0.46	ΔG=-0.64	ΔG=1.08	ΔG=-0.69
Self Dimer	ΔG=-4.89	ΔG=-3.89	ΔG=-3.65	ΔG=-3.61	ΔG=-3.61	ΔG=-4.74
Melting Temp	51.7°C	53.5°C	59.7°C	52.5°C	53°C	57.0°C
Sequence difference %	18.75% from *B. angulatum*	12.5% from *B. longum*	52% from *C. tertium*	35% from *C. perfringens*	28% from *C. leptum*	27.8% from *R. flaveflaciens*
GC %, clamp	62.5%, yes	62.5%, yes	57%, yes	50%, yes	61.1%, yes	61.1%, yes
Length	16 bp	16 bp	23 bp	20 bp	18 bp	18 bp
Hairpin	ΔG=-1.41	ΔG=-1.17	ΔG=-1.09	ΔG=0.84	ΔG=-0.02	ΔG=0.25
Self Dimer	ΔG=-6.76	ΔG=-5.38	ΔG=-4.67	ΔG=-5.36	ΔG=-4.41	ΔG=-3.61
Melting temp	52.9°C	53.0°C	60.9°C	52.8°C	54.4°C	56.1°C
Hetero-dimer	ΔG=-5.09	ΔG=-3.61	ΔG=-5.09	ΔG=-3.61	ΔG=-4.41	ΔG=-4.74

Table 6: The grey rows indicate the primer characteristics of the forward primers of the species indicated. The blue rows indicate the primer characteristics of the reverse primers of the species indicated. The final row indicates the primer characteristic that is dependent on both the forward and reverse primer.

The PCR samples would have been run on an agarose gel to determine approximate PCR product size. If the PCR product size seen on the gel aligned with the expected product size as shown in Table 3, then the 16s rRNA gene in that plasmid is likely the bacteria species for which the primers were designed. If my data were to be consistent with Finegold et al.'s prior research, species G (the bacterial species found in BTBR mice on standard diets) would be one of *Clostridium boltea, Bacteroides vulgatus, Desulfovibrio piger,* or *Clostridium perfringens*. I would expect species M and D to be one of *Bifidobacterium longum, Bifidobacterium angulatum, Clostridium leptum, Clostridium disporicum, Clostridium tertium,* or *Ruminococcus albus*.

Discussion

It is unclear as to why the Sanger sequencing was ineffective. The procedures, vectors, and primer type and amount were all consistent with prior successful sequencing at the Yale facilities for another Trinity College research project. There did appear to be enough DNA in the samples because the Nanodrop confirmed each sample that was sent to Yale had at least 500 ηg of DNA.

The freshly made plasmid samples that were sent to the Yale DNA Analysis Facilities in December 2016 were unsuccessful because they did not seem to have enough DNA in them. In January and February 2017, I remade the samples with double the volume of initial culture to start the procedure in order to increase the amount of DNA. I also concentrated the DNA with a SpeedVac and resuspended the plasmids in TE buffer in hopes of increasing DNA amounts. The primers were changed from T3 and T7 in the first batch to M13 forward and reverse in the second batch. Both sets of primers have binding sites to the vector used. Finally, I changed the broth in the overnight cultures from Tryptic Soy broth to LB broth as suggested by the Yale facility employees. For unknown reasons, the second round of plasmids that were sent to Yale also could not be sequenced. The new plasmid preps were sent to the Yale facilities in March of 2017, after which the Yale technicians and I discussed possible reasons for the failure but found no definitive explanation.

There are a number of possible reasons for the lack of success with the Sanger sequencing. The primers may have degraded through various freeze-thaw cycles. The same could be true for the plasmid samples, but is less likely because the plasmids were prepared two months or sooner before the sequencing reactions were attempted.

Another possible reason is that the samples contained some inhibitory contaminants like salts or ethanol, which would have been due to the rushed nature of the sample preparations.

In order to salvage the project and reach the end goal of learning more about the effects of the ketogenic diet on BTBR mice gut bacteria, a new procedure was created. Unfortunately, there was not enough time to execute the new procedure. Using PCR to check whether pre-determined bacteria species were present in the plasmid samples I had already prepared would be a roundabout but rapid way to determine whether my results were consistent with previous studies' findings. In the initial stages of this project, the 16s rRNA genes were amplified using universal primers (27F and 926R) that were designed to amplify the gene in all bacterial species. Gel electrophoreses of polymerase chain reactions using the universal primers all generate approximately the same size band (~1 kilobase). Designing more specific primers for the polymerase chain reactions will produce different sized bands that will allow those different species to be distinguishable on a gel electrophoresis.

The primers designed with *Clostridium boltea* in mind were strong candidates because they were 20 and 18 base pairs long (forward and reverse primers, respectively) and had at least 15% of base pairs differ from *Clostridium disporicum*. These factors are important because other *Clostridium* species are being studied and I wanted the primers to be able to distinguish the 16s rRNA genes at the species level. The high GC content (55% and 61%) and presence of a GC clamp for both the forward and reverse primers are good indicators because guanine and cytosine use three hydrogen bonds to connect DNA strands as opposed to two, which will help in primer

annealing. Each primer had low risk of a hairpin formation, which is when the primer loops around and forms hydrogen bonds between base pairs on its own strand, preventing proper PCR results. The probability of self-dimer formation (two of the same primer forming hydrogen bonds) and heterodimer formation (the two primers in the pair forming hydrogen bonds with each other) were slightly higher but the ΔG was less than -5.5 and the other positive characteristics of the primer pair outweighed these mediocre ones. The expected PCR product size of the 16s rRNA gene of *Clostridium boltea* is 1097 base pairs, differing from the other *Clostridium* species used in this study by greater than 50 base pairs.

The primers designed for *Bacteroides vulgatus* were good because they each have 18 base pairs with 55.6% of them being guanine or cytosine, along with GC clamps. The risks of hairpin and self-dimer formation were low (ΔG < -3.7) and the melting temperatures were within 0.9°C of the other primer. The one concern about this primer pair was that the ΔG for heterodimer formation is -4.87, which is not ideal, but the aforementioned characteristics outweigh this slightly increased possibility of PCR dysfunction. The expected PCR product size for this primer pair is 986 base pairs.

The primers designed for amplifying the 16s rRNA gene of *Desulfovibrio piger* are 18 and 17 base pairs long (forward and reverse, respectively). They each have GC clamps and GC content between 55 and 65%. The risk of hairpin formations are low (ΔG < -2); the risk of self-dimer formation for the reverse primer is low (ΔG = -3.61); the risk of self-dimer formation for the forward primer is intermediate (ΔG = -4.85); and the risk of heterodimer formation is slightly higher than ideal with the ΔG being -6.91. Though the high risk of heterodimer formation is worrisome, the other possible primer

pairs for this species had even higher risks or had issues fulfilling the other criteria. The expected product size for this primer pair is 1204 base pairs.

The primers designed for testing *Clostridium perfringens* presence were each 20 base pairs long, and had at least 20% of base pairs differ from *C. disporicum*. There is low to intermediate risk for hairpin, self-dimer, and heterodimer formation ($\Delta G < -5$). The forward primer has a GC content of 45% and a partial GC clamp while the reverse primer has a GC content of 50% and a full GC clamp. The expected product size of this primer set is 743 base pairs, which differs from the other *Clostridium* species used in this study by at least 200 base pairs.

The primers designed for *Bifidobacterium longum*, the first of the bacteria shown to be lacking in ASD individuals, were at least 18% different from the related *B. angulatum.* They each contain GC clamps and 50-62.5% guanine and cytosine bases. There are low to intermediate risks for hairpin and heterodimer formation, as well as self-dimer formation for the forward primer ($\Delta G < -5.1$) but a slightly higher risk for self-dimer formation for the reverse primer ($\Delta G = -6.76$). The positive aspects of the primer design outweighed this elevated risk because other possibilities for primer pairs had much higher risks in other categories. The expected product size for these primers is 837, which is over 200 base pairs different from the other *Bifidobacterium* species investigated in this study.

The forward primer designed for *Bifidobacterium angulatum* is 24 bases and have 25% of bases different from *B. longum* while the reverse primer is 16 bases and differs 12.5% from *B. longum.* The risks of hairpin, self-dimer, and heterodimer formations are all low to intermediate ($\Delta G < 5.40$). Each primer has a GC content above

40% and below 63%, and the reverse primer has a GC clamp. The expected product size is 1079 base pairs, which is over 200 base pairs different from the other *Bifidobacterium* species in this study.

The primers designed for *Clostridium leptum* had low to intermediate risks for hairpin, self-dimer, and heterodimer formation ($\Delta G < -5.10$). The forward primer has 22 bases and a GC clamp while the reverse primer has 23 bases and a GC clamp. The GC content of each primer in the set is between 57 and 59% and the expected product size is 1042 base pairs. The expected product size is close to that of *C. boltea* but differs from the other *Clostridium* species in this study by over 200 base pairs.

The forward and reverse primers designed for *Clostridium disporicum* were 18 and 20 bases, respectively; and had 55.6% and 50% GC content, respectively. The risk for hairpin, self-dimer, and heterodimer formation are all low-intermediate ($\Delta G < 5.40$). The forward primer is 17% different from *C. perfringens,* while the reverse is 35% different from *C. perfringens.* The expected product size is 514 base pairs, which is over 200 base pairs different from the other *Clostridium* species in the study.

The primers designed for *Clostridium tertium* both have low risk for hairpin, self-dimer, and heterodimer formation ($\Delta G < -4.42$). The forward primer is 22 bases long and 32% of those bases differ from *C. leptum* while the reverse primer is 18 bases long and 28% of those bases differ from *C. leptum.* The GC contents are 40.9% and 61.1% and each primer has a GC clamp. The expected product size is 1453 base pairs, which differs from the other *Clostridium* species in this study by over 300 base pairs.

The primers designed for *Ruminococcus albus* each have low probability of forming hairpins, self-dimers, or heterodimers ($\Delta G < -4.75$). The GC content of each is

between 57 and 62%, and each have GC clamps. The forward primer is 19 bases long and the reverse is 18 bases. The expected product size of this primer set is 1332 base pairs.

The primers were designed with PCR efficiency in mind, but also with species identification in mind. It was important to have some base pair differences within the primer sequence from the related species that are also apparently implicit in ASD. In addition, the expected product size was made to differ from the related bacteria for ease of distinguishing gel electrophoresis results. Although the PCR and gel electrophoresis analysis could not be completed in the time frame, I feel confident that primers I designed would make it easy to rule the aforementioned bacteria in or out as the identities of species D, G, and M.

There are 44 DNA samples that could be used as the template in PCR reactions with these primer sets. There are 17 samples that contain the 16s rRNA gene of species G, which was shown to only be significantly present in the BTBR/ketogenic diet mouse group. This experimental group was the only mouse group to exhibit the core autistic symptoms, which is why I expect species G to be one of the bacterial species in Table 4. There are also 17 plasmid samples that contained the 16s rRNA gene for species M, which was shown to be present in the healthy mice on both diet types. This bacteria is likely part of the normal flora and contributes to normal gut-brain interactions because it is not in the mouse group that displays the core autistic symptoms. For this reason, I suspect that species M is likely one of the bacteria displayed in Table 5, which were all shown to be absent from autistic individuals in previous studies.

There are 10 plasmid samples that contain the 16s rRNA gene of species D, which was found in the BTBR mouse group that was on the ketogenic diet. Since the ketogenic diet was shown to improve the autistic symptoms of these BTBR mice, I expect that species D is one of the bacteria shown in Table 5 as well. However, this would only be true if the gut bacteria are implicit in actually causing the core autistic symptoms. In addition, the bacteria in Tables 4 and 5 are just a small sample of bacteria present in mouse gastrointestinal tracts. Species D, G, and M may not be any of the bacterial species listed in Tables 4 and 5.

In the future, the polymerase chain reactions should be done with the primers I have designed. In addition, the sequencing error should be figured out, as sequencing the plasmids is still a much more straightforward pathway to figuring out the identities of species D, G, and M. Whether or not these three species are affecting ASD will not be known until they are classified and more is known about how they operate in the body.

If the species had been identified, the next step would be to study the way the bacteria behave in the body. The effects of the ketogenic diet on individuals with ASD should be monitored both behaviorally and in terms of the gastrointestinal flora. Probiotics that can be taken orally that contain the bacteria missing from ASD individuals' guts should then be created and their effects on the gastrointestinal distress and behavioral symptoms tested.

Acknowledgments

I would like to thank my thesis advisor, Dr. Lisa-Anne Foster for guiding me through my research and giving me the opportunity to take on this project. Additionally, I thank Dr. Claire Fournier and Dr. Kathleen Archer for being on my thesis committee and for their helpful suggestions on my work, and to the rest of the Biology department for their continued support. Finally, thank you to the Neuroscience department and former laboratory member Shelby Labe (Trinity '16) for beginning this project.

Literature Cited

Adams, James B., Johansen, L. J., Powell, L. D., Quiq, D., and Rubin, R. A. 2011. Gastrointestinal flora and gastrointestinal status in children with autism – comparisons to typical children and correlation with autism severity. BMC Gastroenterology. 11 (22).

Ardeshir, Amir, Narayan, N. R., Mendez-Lagares, G., Lu, D., Rauch, M., Huang, Y., Van Rompay, K. K. A., Lynch, S. V., and Hartigan-O'Connor, D. J. 2014. Breast-fed and bottle-fed infant rhesus macaques develop distinct gut microbiotas and immune systems. Science Translational Medicine. 6 (252): 1-22.

Bolte, E. R. 1998. Autism and Clostridium tetani. Medical Hypotheses. 51 (2): 133-144.

Comito, Donatella, Cascio, A., Romano, C. 2014. Microbiota biodiversity in inflammatory bowel disease. Italian Journal of Pediatrics. 40 (32): 1-6.

Cryan, John F. and O'Mahony, S. M. 2010. The microbiome-gut-brain axis: from bowel to behavior. Neurogastroenterology and Motility. 23 (3): 187-192.

Davis, Charles Patrick. 1996. Medical Microbiology. 4th edition. Galveston, TX: University of Texas Medical Branch at Galveston.

Facts about autism. 2017. Autism Speaks. <https:www.autismspeaks.org/what-autism/facts-about-autism>, accessed 24 Apr. 2017.

Finegold, Sydney M., Dowd, S. E., Gontcharova V., Liu, C., Henley, K. E., Wolcott, R. D., Youn, E., Summanen, P. H., Granpeesheh, D., Dixon, D., Liu, M., Molitoris, D. R., & Green, J. A. 2010. Pyrosequencing study of fecal microflora of autistic and control children. Anaerobe, 16, 444–453.

Finegold, Sydney M., Molitoris, D., Song, Y., Liu, C., Vaisanen, M. L., Bolte, E., McTeague, M., Sandler, R., Wexler, H., Marlowe, E., Collins, M., Lawson, P., Summanen, P.,Baysallar, M., Tomzynski, T., Read, E., Johnson, E., Rolfe, R., Nasir, P., Shah, H., Haake, D., Manning, P., Kaul, A. 2002. Gastrointestinal Microflora Studies in Late-Onset Autism. Clinical Infectious Disease. 35 (1): 6-16.

Finegold, Sydney M., Summanem, P.H., Downes, J., Corbett, K., Komoriya, T. 2017. Detection of Clostridium perfringens toxin genes in the gut microbiota of autistic children. Anaerobe. 43 (1): 1-5.

Janda, Michael J. and Abbott, Sharon L. 2007. 16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls. Journal of Clinical Microbiology. 45 (9): 2761-2764.

Klein, Matthias S., Newell, C., Bomhof, M. R., Reimer, R. A., Hittel, D. S., Rho, J. M., Vogel, H. J., and Shearer, J. 2016. Metabolomic Modeling to Monitor Responsiveness to Gut Microbiota Manipulation in the BTBRT+tf/j Mouse. Journal of Proteome Research. 15: 1143-1150.

Hentges, David J., Stein, A. J., Casey, S. W., Que, J. U. 1985. Protective Role of Intestinal Flora Against Infection with Pseudomonas aeruginosa in Mice: Influence of Antibiotics on Colonization Resistance. Infection and Immunity. 47 (1): 118-122.

Kang, Seungha, Denman, S., Morrison, M., Yu, Z., Dore, J., Leclerc, M., McSweeney, C. 2010. Dysbiosis of fecal microbiota in Crohn's disease patients as revealed by a custom phylogenetic microarray. Inflammatory Bowel Diseases. 16 (12): 2034-2042.

Labe, Shelby. 2016. “Does gut flora change in a mouse model of autism spectrum disorders on a ketogenic diet?” Hartford, Conn. Trinity College Digital Repository.

McFarlane, H. G., Kusek, G. K., Yang, M., Phoenix, J. L., Bolivar, V. J., Crawley, J. N. 2007. Autism-like behavioral phenotypes in BTBR T+tf/J mice. Genes, Brain and Behavior. 7 (2): 152-163.

Neal, E. G., Chaffe, H., Schwartz, R. H., Lawson, M. S., Edwards, N., Fitzsimmons, G., Whitney, A., and Cross, J. H. 2008. The ketogenic diet for the treatment of childhood epilepsy: a randomised controlled trial. The Lancet Neurology. 7 (6): 500-506.

Parracho, Helena, Bingham, M. O., Gibson, G. R., McCartney, A. L. 2005. Differences between the gut microflora of children with autism spectrum disorders and that of healthy children. Journal of Medical Microbiology. 54: 987-991.

Pequegnet, Brittany, Sagerman, M., Moez, V., Toh, M., Chow, H., Allen-Vercoe, E., Monteiro, M. 2013. A vaccine and diagnostic target for Clostridium bolteae, an autism-associated bacterium. Vaccine. 31 (26): 2787-2790.

Ruskin, David N., Svedova, J., Cote, J. L., Sandau, U., Rho, J. M., Kawamura Jr., M., Bowsoin, D., Masino, S. A. 2013. Ketogenic diet improves core symptoms of autism in BTBR mice. PLoS ONE. 8 (6): e65021.

Schreiner, Andrew B., Kao, J. Y., Young, V. B. 2015. The gut microbiome in health and disease. Current Opinion in Gastroenterology. 31 (1): 69-75.

Sirven, J., Whedon, B., Caplan, D., Liporace, J., Glosser, D., O'Dwyer, J., and Sperling, M. R. 1999. The ketogenic diet for intractable epilepsy in adults: preliminary results. Epilepsia 40 (12): 1721-1726.

Song, Yuli, Liu, C., Finegold, S. M. 2004. Real-Time PCR Quantitation of Clostridia in Feces of Autistic Children. Applied Environmental Microbiology. 70 (11): 6459-6465

Song, Yuli, Liu, C., Molitorus, D., Tomzynski, T., McTeague, M., Read, E., and Finegold, S. 2002. Use of 16S–23S rRNA Spacer-Region (SR)-PCR for Identification of Intestinal Clostridia. Systematic and Applied Microbiology. 25 (4): 528-535.

The National Institute of Mental Health. Autism Spectrum Disorder. October, 2016.