2018

Investigating the Role of the Gut Microbiome in Huntington Disease

Casey G. Hart
University of Central Florida

INVESTIGATING THE ROLE OF THE GUT MICROBIOME IN HUNTINGTON DISEASE

by

CASEY G. HART
B.S. University of Central Florida, 2016

A thesis submitted in partial fulfillment of the requirements
for the Honors in the Major Program in Biomedical Sciences
in the College of Medicine
and in the Burnett Honors College
at the University of Central Florida
Orlando, Florida

Fall Term, 2018

Thesis Chair: Amber Southwell, Ph.D.

ABSTRACT

Huntington disease (HD) is an inherited neurodegenerative disease caused by a trinucleotide repeat expansion in the huntingtin (*HTT*) gene. Metabolic dysfunction is a feature of HD that is recapitulated in HD mouse models. Our lab has shown that circadian feeding rhythms are disrupted in humanized HD mice and restored by suppression of brain HTT. Furthermore, when circadian feeding rhythm is artificially restored, in addition to normalization of metabolic function, liver and striatal HTT is temporarily reduced, demonstrating that HTT is involved in gut-brain feedback. The gut microbiome, which can regulate gut-brain feedback, has been implicated in the pathogenesis of other central nervous system disorders and we hypothesize it also plays a role in HD. The objective of this study is to investigate alterations in relative abundance of HD gut microbiota using existing plasma metabolomics data to identify candidate bacteria. If distinct microbiota profiles are demonstrated, this would provide the basis for future unbiased studies to investigate the complete HD microbiome.

ACKNOWLEDGMENTS

I would like to thank my thesis chair, Dr. Amber Southwell, for her mentorship, invaluable insight, and patience throughout my thesis and entire undergraduate research experience. I'm incredibly thankful for the opportunities and guidance she has given me. My research experience continuously strengthens my love for science and reinforces my decision to pursue medicine.

Thanks to my committee members, Dr. Shibu Yooseph and Dr. Yoon-Seong Kim for their time and patience throughout my thesis journey.

Thank you Dr. William Self and Ms. Jessica Wilson for discussion and for taking the time to grow me various positive and negative control bacterial cultures.

Thank you past and current members of the Southwell lab. Your instruction and teamwork have shaped me into the scientist I am today. You guys made it a pleasure to come to lab each day.

I would also like to thank the UCF Office of Undergraduate Research for funding my research through the Summer Undergraduate Research Fellowship and Student Research Grants.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AL	*Ad libitum*
ASO	Antisense oligonucleotide
ATCC	American Type Culture Collection
BMI	Body mass index
bp	Base pairs
CAG	Cytosine-adenine-guanine
CCM	Cerebral cavernous malformation
CO_2	Carbon dioxide
C_t	Cycle threshold
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
FVB	Friend leukemia virus B
H_2	Hydrogen
HD	Huntington disease
Hdh	Mouse HD homolog
HTT	Huntingtin
Hu18/18	Humanized CAG18/CAG18
Hu97/18	Humanized CAG97/CAG18
ICV	Intracerebroventricular
IGF-1	Insulin-like growth factor 1
ILA	Indole-3-lactic acid

IPA	Indole-3-propionic acid
muHTT	Mutant Huntingtin
N_2	Nitrogen
NC	Normal control
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PHD	Premanifest HD
qPCR	Quantitative polymerase chain reaction
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
SF	Scheduled feeding
TBE	Tris/boric acid/EDTA buffer
TSE	TSE metabolic cage monitoring
WT	Wild-type
wtHTT	Wild-type Huntingtin

CHAPTER ONE: INTRODUCTION

Huntington disease (HD) is a progressive neurodegenerative disease caused by a trinucleotide repeat expansion of CAG within exon 1 of the huntingtin (*HTT*) gene on the short arm of chromosome 4[1]. Patients with 36 – 39 CAG repeats display reduced penetrance—either developing HD late in life or never becoming symptomatic. However, in patients with 40 or more repeats, the disease is fully penetrant, and they will develop HD at some point in their lives. Although symptom onset may begin at any age, it usually occurs between 30 and 50 years of age and is inversely correlated with the number of CAG repeats[2-4]. Patients who develop HD symptoms before the age of 20 have juvenile HD, which is usually caused by 60 or more CAG repeats[2-4]. Although HD is known as an inherited autosomal dominant disease, approximately 10% of HD cases are caused by new mutations[5].

HD is characterized by progressive motor, cognitive, and behavioral decline. Symptom development varies and may not be the same in every patient. Typically a patient is formally diagnosed with HD when overt motor deficits begin to manifest, but cognitive and psychiatric symptoms may develop during the premanifest stage years prior[6]. These early cognitive and psychiatric symptoms are usually subtle and may come in the form of difficulty multitasking or planning and disinhibition or anxiety, respectively[7]. At any time throughout progression, a number of psychiatric disorders may emerge, including depression, apathy, or obsessive and compulsive behavior[6]. The psychiatric symptoms of HD are often the most distressing for patients and their caregivers[6]. Personality changes, irritability, and aggression may occur and vary on a day-to-day basis[6]. Early cognitive decline usually affects executive functioning and short-term memory, while sparing long-term memory, but may ultimately lead to advanced

dementia[6]. Motor decline usually begins as minor, suppressible fidgeting that may go unnoticed by the patient[8]. Bradykinesia is another common symptom, causing slowed muscle movements, including reduced eye movement and facial expressions[8]. Loss of postural reflexes and developing an ataxic gait may immobilize patients, reducing their independence[8]. Motor decline may eventually culminate in involuntary, excessive movements known as chorea, the hallmark symptom of HD, but is not developed by all patients[7]. HD progression typically lasts 15 – 20 years after diagnosis before resulting in death[8].

Two of the most common non-neurological symptoms of HD are progressive weight loss[9] and metabolic dysfunction[10]. However, anecdotal evidence suggests a small proportion of HD patients are pathologically overweight, despite having the same CAG repeat lengths as normal weight and pathologically underweight patients. Furthermore, patients with a higher body mass index (BMI) at disease onset have a slower rate of HD progression[11]. Although HD patients have increased sedentary energy expenditure despite adequate feeding[12], a mechanism of action has yet to be characterized that explains this spectrum of weight abnormalities.

Pathological weight loss is recapitulated in HD model mice expressing fragments of human *HTT*, either transgenically[13,14] or knocked-in to exon 1 of the mouse HD homolog (*Hdh*) gene[15]. Conversely, pathological weight gain is recapitulated in transgenic HD model mice expressing full-length human *HTT*, either along with the full complement of *Hdh*[16,17] or in *Hdh*-null backgrounds, which are referred to as humanized mice[18,19].

While studying metabolic dysfunction of the pathologically overweight humanized HD model mice[18], our lab discovered that circadian feeding is disrupted, despite maintenance of naturally nocturnal circadian activity (Figure 1). Interestingly, circadian feeding patterns are restored by suppression of brain HTT (Figure 2), suggesting that circadian feeding may be

regulated by HTT in the brain. Furthermore, when circadian feeding patterns are artificially restored with scheduled feeding, body weight (Figure 3) and metabolic markers (Figure 4) are normalized, while liver and striatal HTT is temporarily suppressed (Figure 5). This demonstrates that HTT is involved in gut-brain feedback. Since HTT suppression during scheduled feeding is only temporary, while metabolic effects are lasting, we suspect that HTT is not the master regulator of this feedback loop. Instead, we hypothesize that the gut microbiome may influence this pathway, possibly contributing to HD pathogenesis.

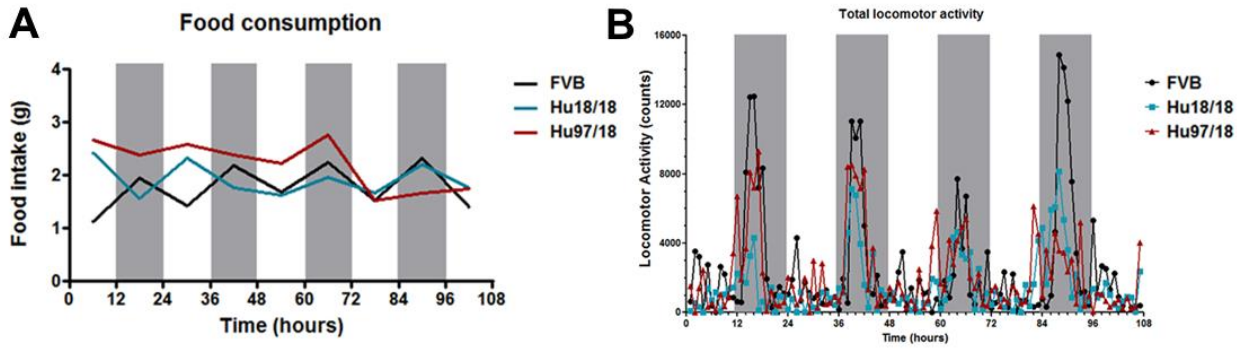

Figure 1. Circadian feeding is disrupted but circadian activity is maintained in humanized HD mice. (A) Measurement of food intake over time for FVB (wild-type (WT) mice), Hu18/18 (homozygous for WT HTT (wtHTT)), and Hu97/18 mice (heterozygous for WT and mutant HTT (muHTT)). FVB mice display typical nocturnal feeding patterns with increased food consumption during the dark phase, while Hu18/18 and Hu97/18 mice display disrupted circadian feeding patterns. (B) Hu18/18 and Hu97/18 mice display similar circadian activity to FVB mice with increased locomotor activity during the dark phase.

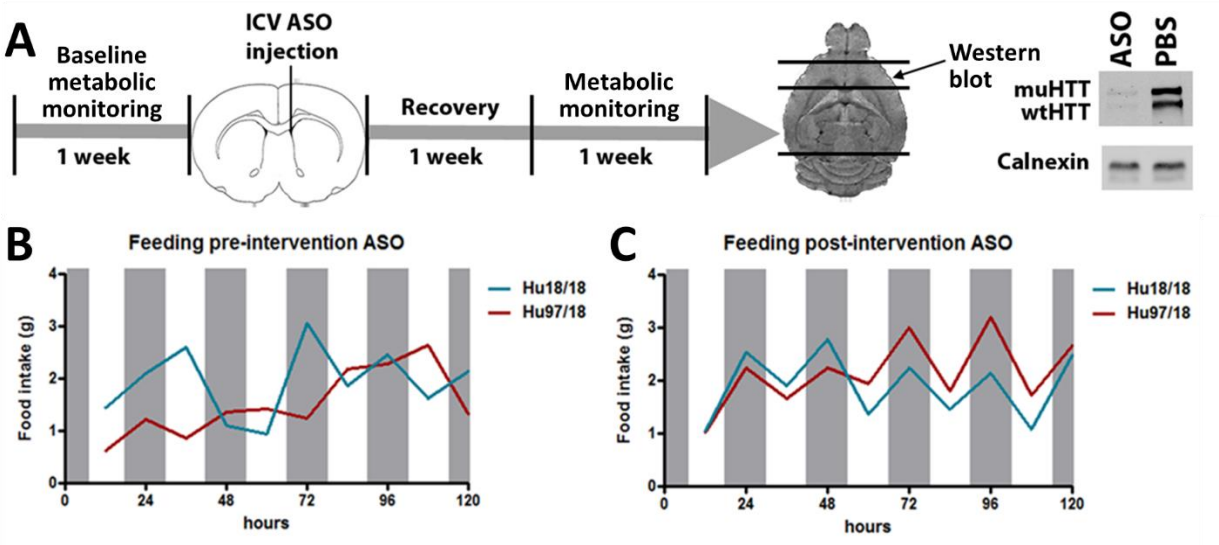

Figure 2. Circadian feeding patterns are restored by suppression of brain HTT. (A) Overview of experimental design: mice undergo baseline metabolic phenotyping, receive an intracerebroventricular (ICV) injection of HTT antisense oligonucleotide (ASO), and undergo follow up metabolic phenotyping after a 1-week recovery period. muHTT and wtHTT suppression in brain is then verified via Western blotting. (B) Hu18/18 and Hu97/18 mice display disrupted circadian feeding patterns at baseline. (C) Hu18/18 and Hu97/18 mice display typical circadian feeding patterns after brain HTT suppression.

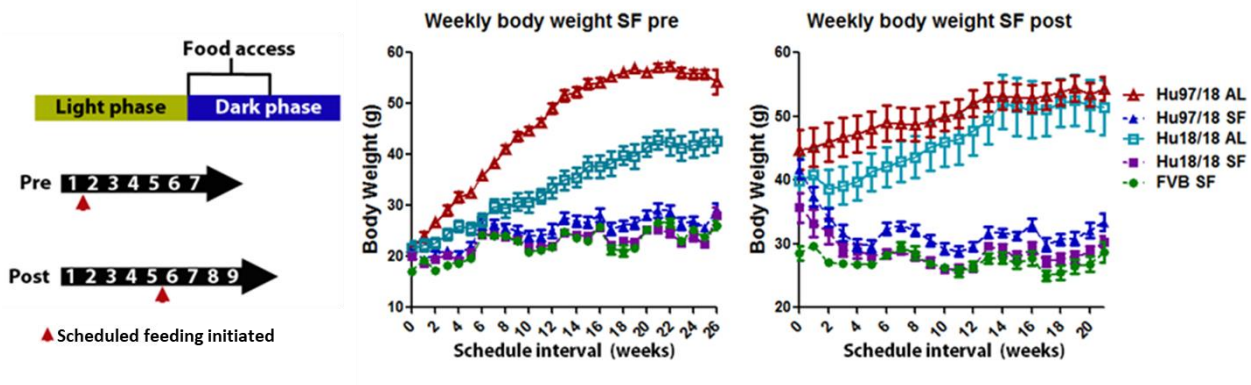

Figure 3. Body weight normalizes with scheduled feeding. SF was initiated pre- or post-pathological weight gain and mice were weighed weekly. When initiated pre-symptomatically, SF prevents pathological weight gain. When initiated post-symptomatically, SF induces normalization of humanized HD mouse body weight within approximately 4 weeks.

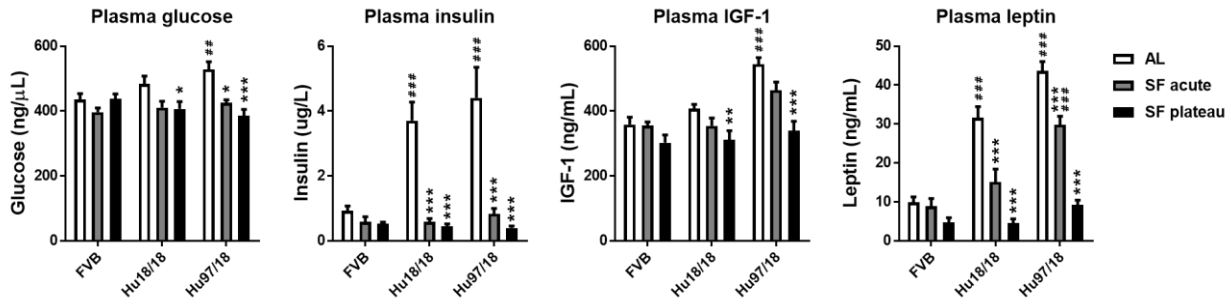

Figure 4. Metabolic markers normalize with scheduled feeding. Compared to FVB controls, humanized HD mice display elevated circulating glucose, insulin, insulin-like growth factor 1 (IGF-1), and leptin. SF rapidly normalizes glucose and insulin levels during the acute weight loss phase and eventually normalizes IGF-1 and leptin levels by the plateau phase. * = different from AL fed for the same genotype. # = different from FVB for the same feeding paradigm. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

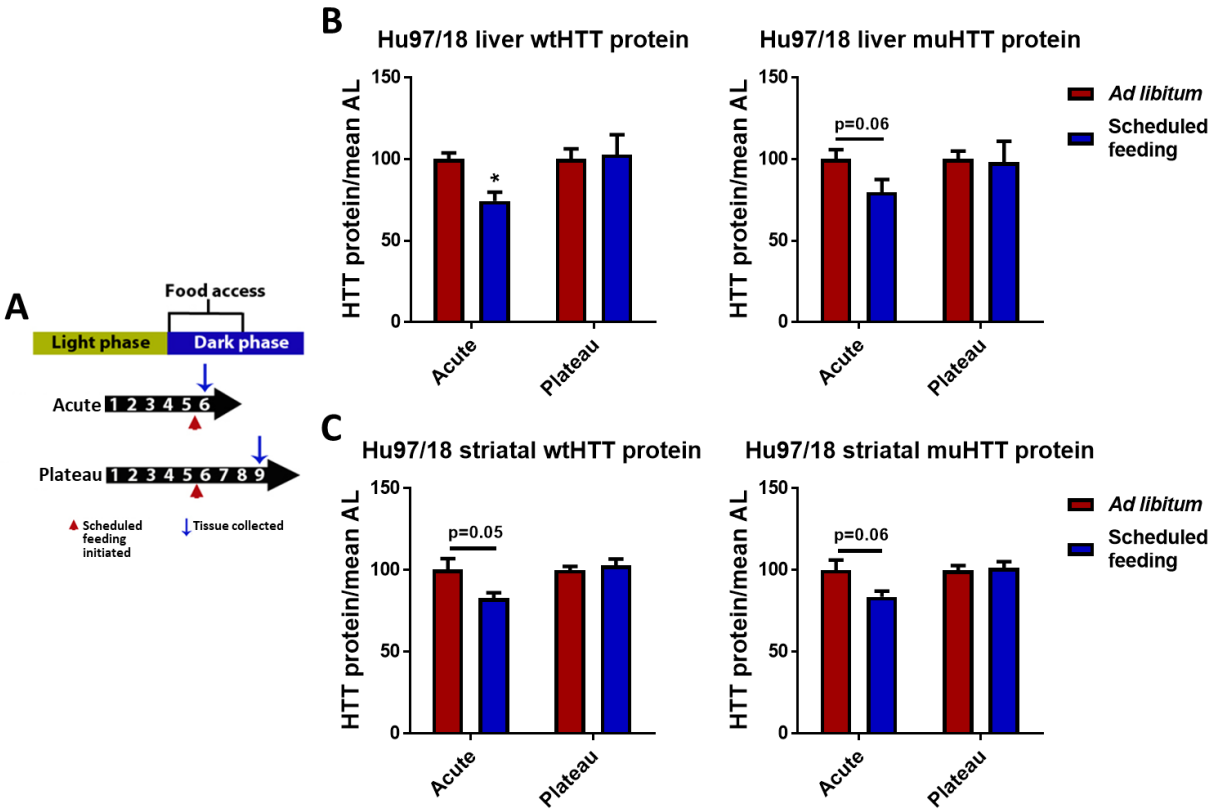

Figure 5. Striatal and liver HTT is temporarily suppressed by scheduled feeding. (A) Experimental design: Scheduled feeding (SF) i.e. food access during only the first 6 hours of the dark phase, was initiated at 5.5 months of age, a time point when humanized HD model mice have already undergone significant pathological weight gain. Acute: tissues are collected during the acute weight loss phase. Plateau: tissues are collected in the plateau phase after body weight has normalized. (B,C) SF induces reduction of (B) liver and (C) striatal HTT protein during the acute weight loss phase. Once the plateau phase is reached, HTT returns to basal levels. * = different from *ad libitum* (AL) at $p < 0.05$.

A microbiome is the collective genome of all microorganisms found within a specific microbial community. The microorganisms within a microbiome, commonly referred to as microbiota, can consist of viral, unicellular, and multicellular microbes. However, the most abundant and studied microbiota is bacteria[20], which will be the focus of this study. The gut microbiome is the most clinically relevant microbiome because it houses the greatest number of bacteria within the host[21]. Bacteria have co-evolved a mutualistic relationship with their host. The host provide bacteria with vital nutrients necessary for growth, while the bacteria develop the host's immune system and synthesize beneficial metabolites[22].

Bacteria inhabiting the gut are involved in many host metabolic pathways, and changes in bacterial composition can result in metabolic dysregulation and disorders[22], including neurological disorders[23]. For instance, the presence of gram-negative bacteria or lipopolysaccharide accelerates cerebral cavernous malformation (CCM) by over-activating endothelial toll-like receptor 4. Meanwhile, germ-free mice are protected from CCM and even a single course of antibiotics permanently reduces CCM susceptibility[24]. Another study found that probiotic treatments with *Bacteroides fragilis* improves gastrointestinal abnormalities and behavioral deficits in autism model mice. Additionally, these probiotic treatments altered gut microbiota composition and levels of several serum metabolites[25].

It is well studied that microbiota-derived metabolites produced in the gut are absorbed in the intestines, where they cross the intestinal epithelium. Metabolites then enter systemic circulation after passing through the liver, where they can reach or exceed typical drug dose concentrations[26]. For instance, gut bacteria that overexpress the protein α-synuclein promote Parkinson disease-related motor deficits in mice, while antibiotic treatments decrease these symptoms[27], demonstrating that microbiota-derived metabolites can influence host physiology.

Hoping to develop a noninvasive way of identifying active processes behind HD pathogenesis, a plasma metabolomics study previously identified metabolite alterations within several pathways of HD patients[28]. The tryptophan, tyrosine, purine, and antioxidant pathways were specifically targeted for investigation because they have been previously implicated as relevant to HD neurodegeneration[29-31]. The plasma metabolomic profile of 52 premanifest HD, 102 early symptomatic HD, and 140 healthy controls were quantified using liquid chromatography. Of the 29 known metabolites within the analyzed pathways, 21 had significant concentration alterations (Table 1). Dysregulation of any of these metabolites could have wide-ranging effects on the host.

Table 1. Metabolites significantly altered in HD plasma.

Pathway	Metabolite	NC vs PHD	NC vs HD	PHD vs HD
Tryptophan	Serotonin	↓	↓	
	N-acetylserotonin	↑	↑	
	5-hydroxytryptophan			↓
	5-hydroxyindoleacetate	↓		↑
	Kynurenine		↓	
	3-hydroxyanthranilic acid		↓	↓
	Indole-3-lactic acid	↓		
	N-methyltryptamine	↓		
	Indole-3-propionic acid	↓	↓	
Tyrosine	Tyrosine	↑		
	Homovanillic acid		↑	↑
	Homogentisic acid	↑		
	2-hydroxyphenylacetate		↓	↓
	3-hydroxyphenylacetate		↑	↑
	4-hydroxyphenyllactic acid		↓	↓
Purine	Xanthine	↓	↓	
	Xanthosine		↓	
	Urate			↓
Antioxidant	Gamma tocopherol		↑	↑
	Alpha tocopherol	↑	↑	

Methionine	Methionine	↑	↓

NC: normal control, PHD: premanifest HD, HD: early symptomatic HD. Arrows indicate increase or decrease in concentration. Adapted from[28].

Gut microbiota concentration is measured by quantifying the microbial DNA present in stool or cecal samples[32]. Quantification of stool bacteria is more representative of the large intestine's microbiota profile, while the cecum gives a more accurate representation of the small intestine's microbiota profile[33]. By targeting the variable regions of the 16S ribosomal RNA (rRNA) gene with specific primers pairs, all bacterial genomes within a sample can be amplified via polymerase chain reaction (PCR). If a fluorescent probe or dye is used, fluorescence can be measured in real-time during amplification via quantitative PCR (qPCR), allowing quantification of the amplicon. 16S rRNA gene amplification is a non-targeted approach to quantify bacteria, while quantification of specific target bacteria can be performed via qPCR using a genus- or species-specific primer pair. By comparing quantification between multiple samples and a reference gene, the relative abundance of each sample can be obtained[34].

This study aims to investigate the gut microbiome of an HD model mouse through analysis of stool and cecal samples. Q175FDN mice, which have human *HTT* exon 1 with ~190 CAG repeats knocked-in to *Hdh*, display the most robust HD-like phenotypes of any knock-in or full-length HD mouse model and recapitulate pathological weight loss[15].

The host-microbe relationship between a mouse and its gut microbiome have co-evolved differently than that of humans, and will therefore never fully recapitulate the human gut microbiome. However, a comparison of murine and human gut microbial diversity has shown 90% and 89% similarity between bacterial phyla and genera, respectively[35]. These strong similarities justify the use of mice as a human experimental model. In addition, studies have

shown that the biggest contributors to microbiota diversity are genetics and diet. Since facility-housed mice are inbred and receive the same food, microbiota diversity should be minimal within each cohort, increasing validity.

If distinct microbiota profiles are demonstrated between WT and HD mice, future studies to investigate the gut microbiome in HD patients would be warranted. Finding a relationship between the gut microbiome and HD could identify novel, noninvasive biomarkers for disease onset and progression, as well as potential therapeutic targets.

CHAPTER TWO: RESULTS

Aim 1: To Identify Bacterial Candidates of Exclusively Microbiota-derived Metabolites

By identifying bacterial producers of exclusively microbiota-derived metabolites, we can identify candidate microbes that may be altered in the HD gut microbiome. Metabolites that are produced by both humans and microbiota were not pursued because it would not be possible to determine which source contributed to each metabolic alteration. Using an online metabolic pathway database[36], potential exclusively microbiota-derived metabolites were identified. Further searching through NCBI's gene database and relevant literature review revealed that out of the 21 metabolites from Table 1, only four are produced exclusively by microorganisms (Table 2).

Table 2. Microbiota-derived metabolites significantly altered in HD plasma

Pathway	Metabolite	NC vs PHD	NC vs HD	PHD vs HD
Tryptophan	Indole-3-lactic acid	↓		
	Indole-3-propionic acid	↓	↓	
Antioxidant	Gamma tocopherol		↑	↑
	Alpha tocopherol	↑	↑	

NC: normal control, PHD: premanifest HD, HD: early symptomatic HD. Arrows indicate increase or decrease in concentration. Derived from Table 1.

Gamma and alpha tocopherol are antioxidants within the vitamin E biosynthesis pathway, which is only found in plants and photosynthetic microorganisms[37]. *Synechocystis sp*. PCC 6803 is a freshwater cyanobacterium and is the only bacterium currently known to produce gamma and alpha tocopherol[38]. These metabolites were shown to be increased in HD plasma[28] (Table 1, 2). Although cyanobacterial infections are known to occur in humans[39], our literature review did not find evidence that cyanobacteria are normally present in the human gut microbiome. However, non-photosynthetic relatives of cyanobacteria have been found in the gut[40]. Although

these cyanobacterial relatives have been found to synthesize vitamins B and K[40], there is currently no evidence that they are capable of also synthesizing vitamin E. There are also no studies showing *Synechocystis sp*. PCC 6803 presence in the human gut. Thus, any observed alterations in tocopherol concentration are most likely not the result of bacterial production, but instead the result of dietary or physiological changes.

Both indole-3-lactic acid (ILA) and indole-3-propionic acid (IPA) are products of reductive tryptophan catabolism. Since ILA production has been observed within multiple phyla[41], ILA may be produced from several unique pathways, none of which are well-elucidated. Additionally, since significant ILA alteration was only found when comparing premanifest HD patients to controls and was not found when comparing manifest HD patients to premanifest HD patients or controls[28] (Table 1, 2), this change may not be highly relevant to HD pathogenesis.

On the other hand, IPA is significantly altered within both premanifest and manifest HD patients when compared to controls[28] (Table 1, 2). IPA has a well-characterized pathway. It's produced by phenyllactate dehydratase, which is encoded by the *fldC* gene[42]. IPA has shown potent antioxidant activity *in vitro* and *in vivo* within rat brains[43]. IPA even protected primary neurons and neuroblastoma cells from the oxidative stress caused by amyloid beta, the primary pathology of Alzheimer disease[44]. Oxidative stress plays a role in HD-related neurodegeneration, although it's still unclear whether oxidative stress is causal or a downstream result[45]. Antioxidant trials in HD patients have been generally unsuccessful[46-50], so development of novel ways to reduce oxidative stress would be beneficial.

Furthermore, a recent study has shown that in the absence of IPA, mice have increased activation of immune cells[42]. Similarly, HD patients have increased activation of the innate immune system, both centrally and peripherally[51]. Monocytes in HD patients display a

hyperactive response to stimulation, as well as macrophages and microglia in HD model mice[51]. An increase in circulating interleukin 6 levels has also been observed in HD patients an average of sixteen years before the onset of clinical symptoms[51]. Additionally, IPA has been shown to fortify the intestinal barrier[52], and its absence, increases intestinal permeability[42]. With these documented effects, alteration of IPA concentration presents a viable candidate for contributing to HD-related metabolic dysfunction and potentially, neurodegeneration. For these reasons, we prioritized quantification of IPA-producing bacteria. An *in silico* bacterial target search identified two genera with four species each that have been shown to produce IPA (Table 3).

Table 3. Known bacterial producers of the exclusively microbiota-derived metabolite IPA

Indole-3-propionic acid	*Clostridium botulinum*[53] *Clostridium cadaveris*[42] *Clostridium paraputrificum*[54] *Clostridium sporogenes*[42,53,55] *Peptostreptococcus anaerobius*[42,56] *Peptostreptococcus asaccharolyticus*[54] *Peptostreptococcus russellii*[56] *Peptostreptococcus stomatis*[56]

Aim 2: To Develop Bacterial Candidate Qualitative and Quantitative Assays

DNA was extracted from five gram negative and three gram positive bacterial cultures (Gram negative: *E. coli*, *P. aeruginosa*, *A. faecalis*, *E. aerogenes*, and *S. marcescens*; Gram positive: *C. sporogenes*, *C. difficile*, and *S. aureus*) for use as positive and negative controls during PCR assay development. DNA concentration and purity was measured via spectrometry and is summarized in Table 4. A DNA concentration of >10 ng/µL was preferred to ensure enough DNA for PCR. Absorbance 260/280 and 260/230 ratios of approximately 1.8 – 2.0 and 1.8 – 2.2, respectively, were preferred to ensure an adequately pure DNA sample. Gram positive

bacteria are more difficult to lyse due to their thicker peptidoglycan layer. This most likely explains why two of the gram positive bacterial cultures showed the greatest deviations from the preferred DNA values. However, these deviations did not seem to affect downstream applications.

Table 4. Spectrophotometry of positive and negative control bacterial DNA samples

Bacteria	Concentration (ng/µL)	260/280	260/230
Escherichia coli ATCC 8739	507.40	2.01	1.80
Pseudomonas aeruginosa ATCC 27853	761.97	2.02	2.12
Alcaligenes faecalis ATCC 8750	907.73	2.06	2.09
Enterobacter aerogenes ATCC 13048	408.80	2.01	1.88
Serratia marcescens ATCC 14756	381.80	2.00	2.07
Clostridium sporogenes ATCC 3584	14.20	1.64	0.94
Clostridium difficile ATCC 9689	94.23	1.90	2.69
Staphylococcus aureus ATCC 25923	30.03	2.31	0.68

Positive and negative control DNA samples were validated using the 16S rRNA gene-specific primer pair 27F – 1492R[57] (Figure 6). All eight bacterial samples displayed the expected amplicon length at approximately 1,500 bp, confirming that bacterial DNA was present in each sample.

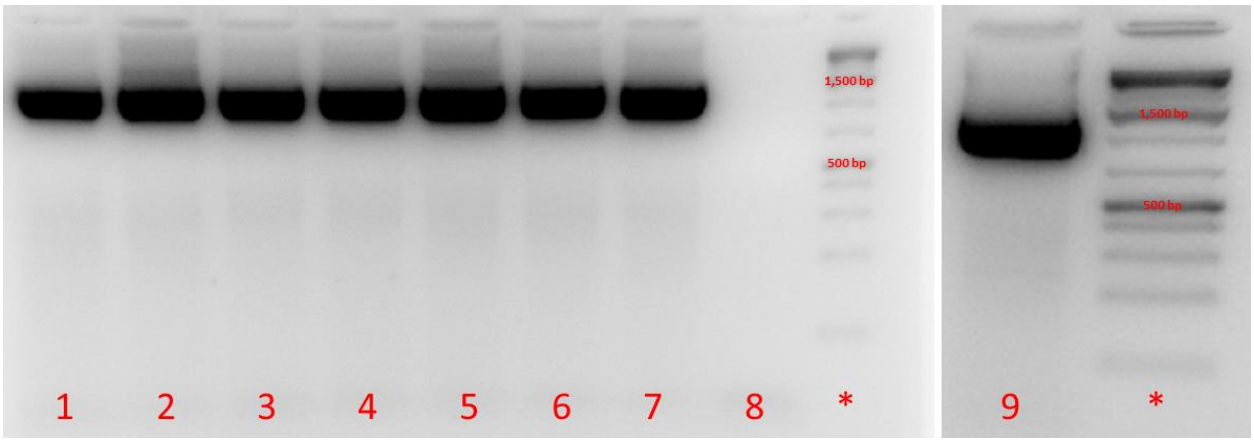

Figure 6. Positive and negative control DNA samples contain bacterial DNA. Total DNA extracted from positive and negative control bacterial cultures show a DNA band at approximately 1,500 bp. Expected amplicon length for 16S rRNA gene primer pair: approximately 1,500 bp. Lane 1: *C. sporogenes*. Lane 2: *E. coli*. Lane 3: *P.*

aeruginosa. Lane 4: *A. faecalis*. Lane 5: *E. aerogenes*. Lane 6: *S. aureus*. Lane 7: *S. marcescens*. Lane 8: Negative control. Lane 9: *C. difficile*. Lane *: GeneRuler 1 kb DNA Ladder.

The specificity of the *fldC* primer pair was validated using the same eight bacterial DNA samples (Figure 7). As expected, *C. sporogenes* was the only bacterial DNA sample that displayed the expected amplicon length at 600 bp. Although *C. difficile* is closely related to *C. sporogenes*, it does not produce IPA, and thus, should not display a DNA band. This control was important to ensure that the *fldC* primer pair is specific to *fldC*-producing bacteria and does not target closely related species.

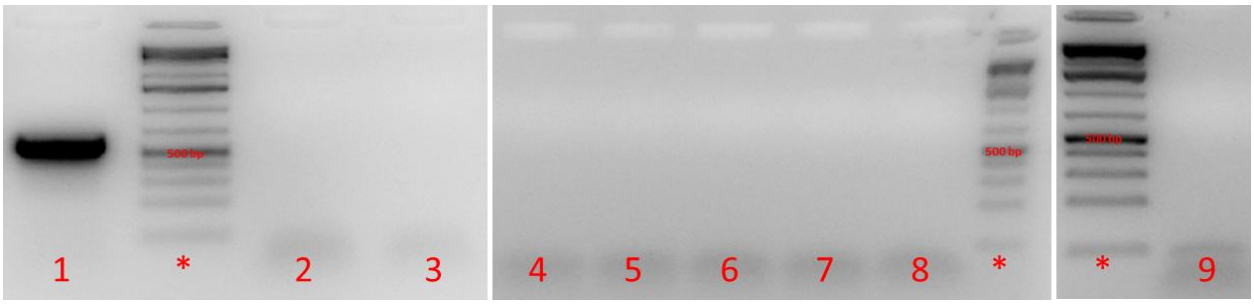

Figure 7. Positive control DNA samples contain the *fldC* gene. Positive control displays a DNA band at 600 bp, while negative controls do not display amplification. Expected amplicon length for *fldC* primer pair: 600 bp. Lane 1: *C. sporogenes*. Lane 2: *E. coli*. Lane 3: *P. aeruginosa*. Lane 4: *A. faecalis*. Lane 5: *E. aerogenes*. Lane 6: *S. aureus*. Lane 7: *S. marcescens*. Lane 8: Negative control. Lane 9: *C. difficile*. Lane *: GeneRuler 1 kb DNA Ladder.

To determine the lower limit of detection for the *fldC* primer pair within a mixed microbiota DNA sample, 50 ng of WT mouse stool DNA was spiked with decreasing amounts of *C. sporogenes* DNA, ranging from 2.84 ng (10^{-9} grams) to 2.33 ag (10^{-18} grams) (Figure 8). *FldC* PCR was performed with 30 cycles, which displayed bands as low as the 1/125 dilution (Figure 8A). *FldC* PCR was repeated, this time with 40 cycles, which displayed bands in all lanes, including the lowest dilution of 1/15,625 (Figure 8B). Non-specific binding also increased. Each subsequent lane displays a reduction in band intensity, suggesting the serial dilution was performed correctly.

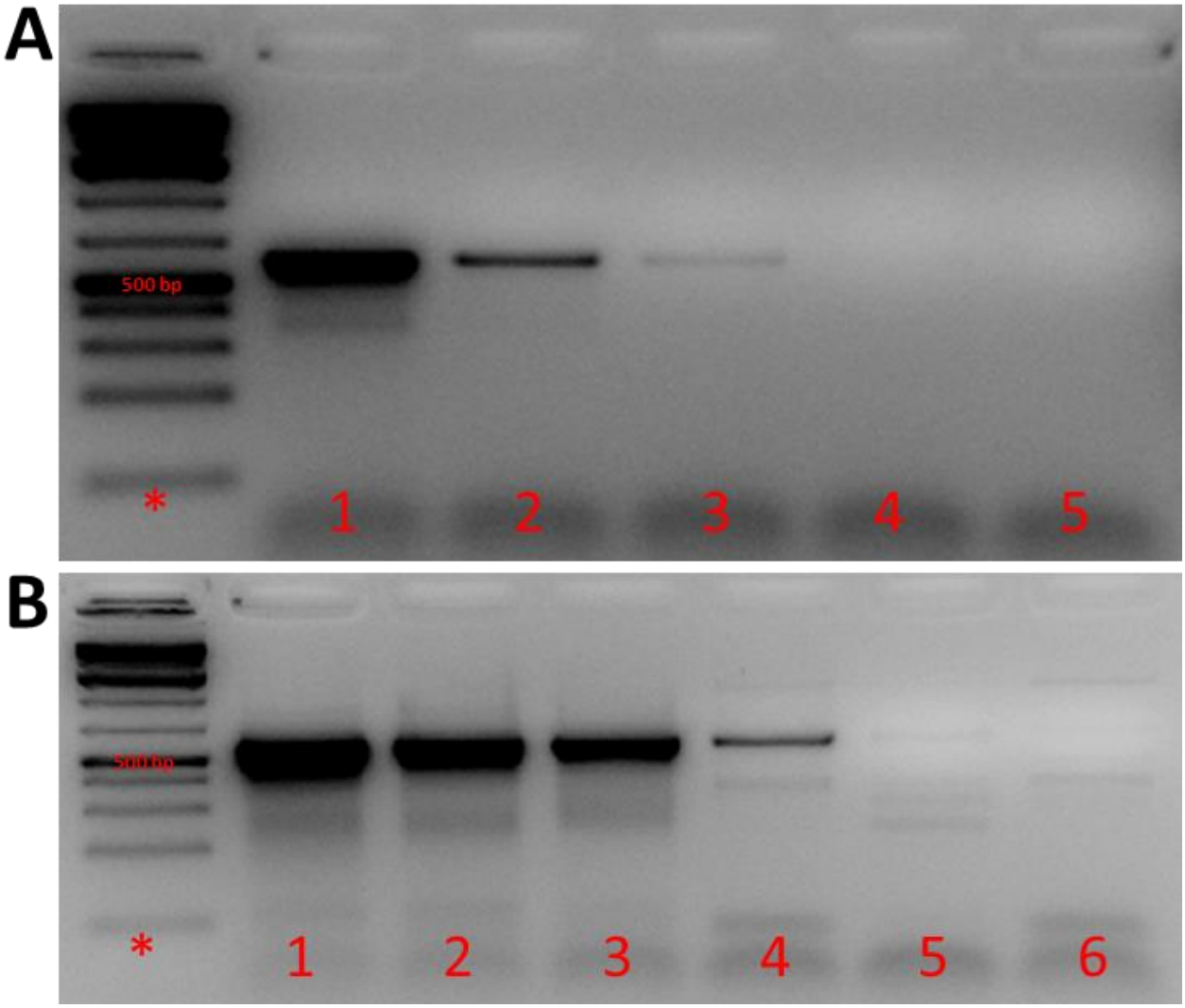

Figure 8. Determining the lower limit of detection of the *fldC* primer pair. 50 ng of mouse stool DNA was spiked with a 5x dilution series of *C. sporogenes* DNA. (A) 30 cycle PCR. Lane *: GeneRuler 1 kb DNA Ladder. Lane 1: 1/5 (~2.84 ng). Lane 2: 1/25 (~568.13 pg). Lane 3: 1/125 (~113.63 pg). Lane 4: 1/625 (~22.73 pg). Lane 5: 1/3,125 (~4.55 pg). (B) 40 cycle PCR. The exposure in lanes 5 and 6 was increased to better visualize the bands. Lanes * – 5 are the same as (A). Lane 6: 1/15,625 (~909.01 fg).

To increase assay sensitivity and specificity, a second, nested PCR reaction was performed using a primer pair targeting a region within the first 600 bp PCR product. Two nested primer pairs were developed, with expected amplicon lengths of 81 bp and 480 bp, respectively. Using nested PCR reduces non-specific binding and increases amplicon yield, so

performing nested PCR should increase the amplification of *fldC* if it is within the DNA sample, even if it is present at very low abundance.

The two nested PCR reactions using PCR product from Figure 8B displayed the expected DNA bands at 81 bp (Figure 9A) and 480 bp (Figure 9B), respectively. The lowest DNA dilution evaluated was approximately 909 fg ($\times 10^{-15}$) of *C. sporogenes* DNA spiked into 50 ng of WT stool DNA, which represents approximately 0.00182% of the total DNA within the sample. We chose to continue using the 480 bp nested primer pair in subsequent assays because it appears to have greater specificity than the 81 bp nested primer pair.

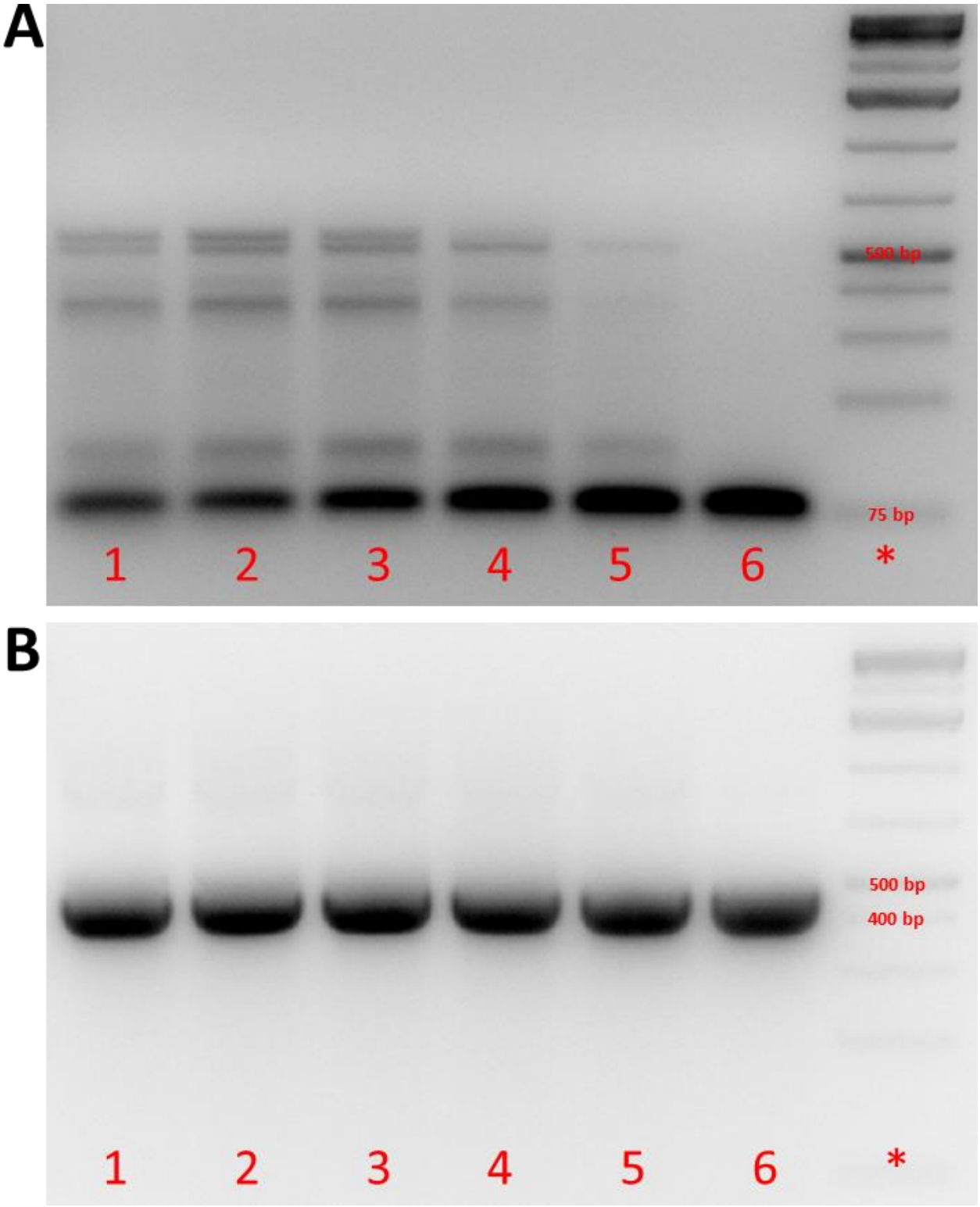

Figure 9. Nested PCR increases *fldC* amplification. Second nested PCR reactions using the 600 bp *fldC* amplicon as DNA template. Lane 1: 1/5 (~2.84 ng). Lane 2: 1/25 (~568.13 pg). Lane 3: 1/125 (~113.63 pg). Lane 4: 1/625

(~22.73 pg). Lane 5: 1/3,125 (~4.55 pg). Lane 6: 1/15,625 (~909.01 fg). Lane *: GeneRuler 1 kb DNA Ladder. (A) Expected amplicon length: 81 bp. (B) Expected amplicon length: 480 bp.

Subsequent attempts to examine smaller *C. sporogenes* dilutions were varied (Figure 10). For the first nested PCR reaction, the DNA bands are no longer detectable below the 1/625 dilution, despite the gel in Figure 8B showing bands two dilutions below this. Also, the gel's highest intensity band appears in the 1/78,125 dilution. For the second nested PCR reaction, high intensity bands are seen in the first seven dilutions, where it begins to fade in the 1/1,953,125 dilution, only to appear again in the two lowest dilutions. The variability in band detection may be caused by the increased difficulty of pipetting small amounts of DNA as the dilutions decreased. For this reason, a conservative estimate of the nested PCR's limit of detection is approximately 1.45 fg of *C. sporogenes* in 50 ng of WT stool DNA, representing approximately 2.91×10^{-6}% of the total DNA in the sample. These results demonstrate that this nested PCR assay increases the *fldC* primer pair limit of detection considerably. Thus, we are encouraged that the nested PCR reaction worked as expected to amplify PCR signal.

We have determined that nested PCR is unnecessary for amplification of the 16S rRNA gene due to the large proportion of 16S DNA found within each stool DNA sample.

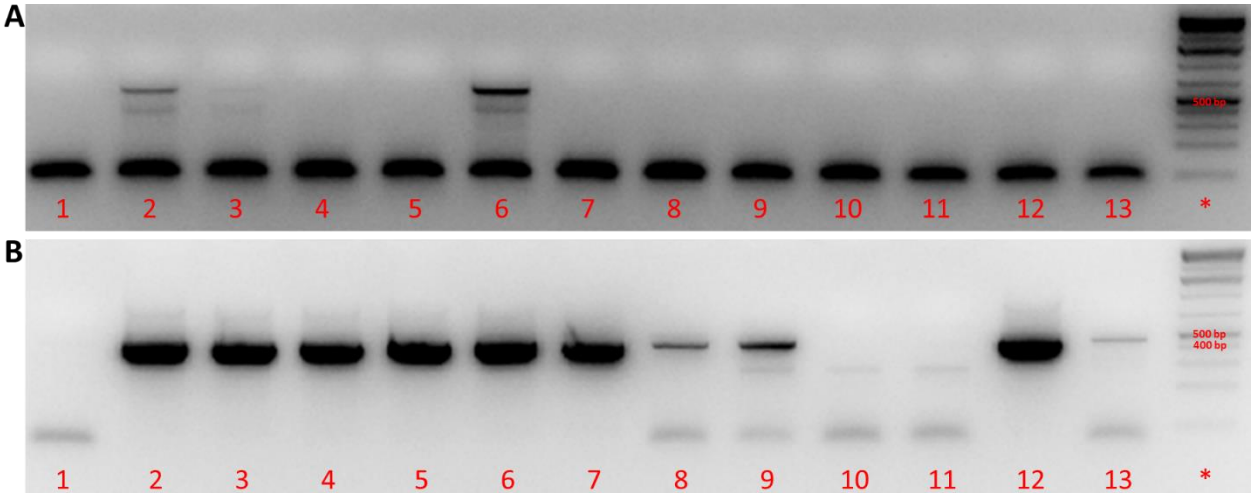

Figure 10. Nested PCR serial dilution. Both reactions of the nested PCR assay using a 5x serial dilution of *C. sporogenes* DNA spiked into WT stool DNA. Lane 1: Negative control. Lane 2: 1/125 (~113.63 pg). Lane 3: 1/625 (~22.73 pg). Lane 4: 1/3,125 (~4.55 pg). Lane 5: 1/15,625 (~909.01 fg). Lane 6: 1/78,125 (~181.80 fg). Lane 7: 1/390,625 (~36.36 fg). Lane 8: 1/1,953,125 (~7.27 fg). Lane 9: 1/9,765,625 (~1.45 fg). Lane 10: 1/48,828,125 (~290.88 ag). Lane 11: 1/244,140,625 (~58.18 ag). Lane 12: 1/1,220,703,125 (~11.64 ag). Lane 13: 1/6,103,515,625 (~2.33 ag). Lane *: GeneRuler 1 kb DNA Ladder. (A) Nested PCR reaction 1. Expected amplicon length: 600 bp. (B) Nested PCR reaction 2. Expected amplicon length: 480 bp.

Since DNA from the eight bacterial culture controls was extracted using a DNeasy Blood & Tissue Kit, while DNA from the stool and cecal samples was extracted using a QIAamp Fast DNA Stool Mini Kit, variations in the DNA extraction processes may affect results. To address this, DNA from pure *C. sporogenes* bacteria or a WT mouse stool sample spiked with pure *C. sporogenes* bacteria were extracted using the stool kit. Both of these samples displayed the expected 600 bp DNA band when using the *fldC* primer pair (Figure 11), confirming that the QIAamp Fast DNA Stool Mini Kit can extract DNA from gram positive bacteria outside of and within a mixed microbiota sample.

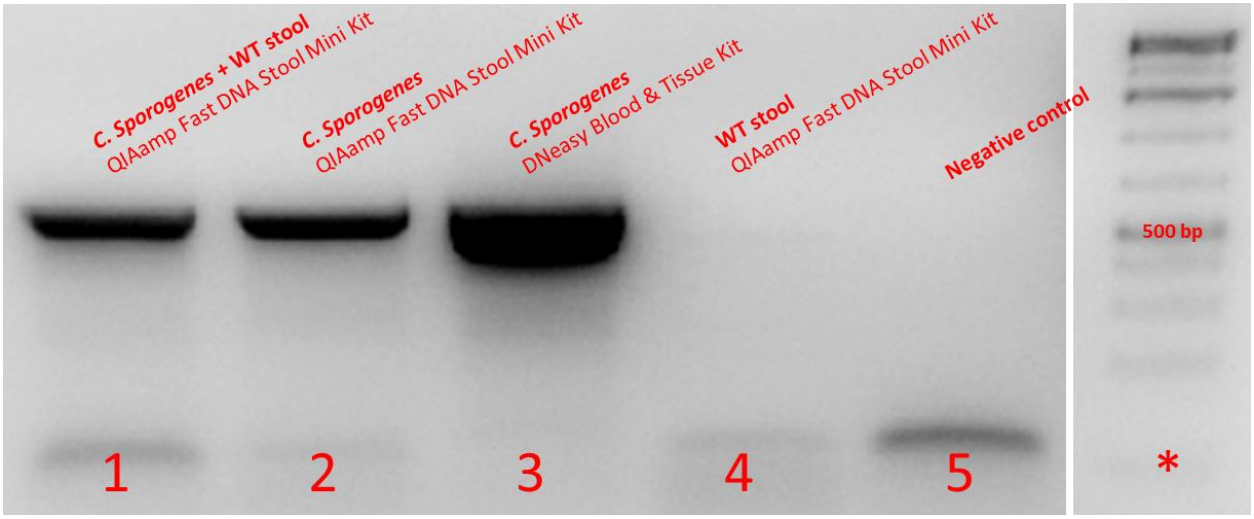

Figure 11. The QIAamp Fast DNA Stool Mini Kit can extract *C. sporogenes* DNA from stool. DNA extracted from stool spiked with *C. sporogenes* bacteria using a QIAamp Fast DNA Stool Mini Kit displays the expected DNA band at 600 bp when using the *fldC* primer pair. Lane 1: *C. sporogenes* + WT stool extracted using a QIAamp Fast DNA Stool Mini Kit. Lane 2: *C. sporogenes* extracted using a QIAamp Fast DNA Stool Mini Kit. Lane 3: *C. sporogenes* extracted using a DNeasy Blood & Tissue Kit. Lane 4: WT mouse stool DNA without spiked *C. sporogenes*. Lane 5: *fldC* primer pair negative control. Lane *: GeneRuler 1 kb DNA Ladder.

We next sought to develop a qPCR assay using a genus-specific *Clostridium spp.* primer pair[58], a genus-specific *Lactobacillus spp.* primer pair[59], a *Leuconostoc mesenteroides* primer pair[60], and an actin-specific primer pair as a reference gene (Table 8). Actin was chosen because the primer set and reaction has been previously optimized in our lab for use as a reference gene in mouse tissue qPCR. Cecal DNA samples from 3 WT mice were pooled to provide a DNA template with a concentration of 785 ng/µL. Each reaction was run in triplicate and four different template DNA dilutions (1, 1/4, 1/16, and 1/64) were used with each primer pair to determine if template concentration affects amplification in a dose-dependent manner.

The qPCR C_t results displayed no amplification for the *Clostridium spp*. and *Leuconostoc mesenteroides* primer pairs (Table 5). Both the actin and *Lactobacillus spp*. primer pairs displayed no or inconsistent amplification in the pre-diluted triplicate samples, suggesting too much template DNA was present. As the DNA template became more diluted, less variation

between the C_t values was observed. The 1/64 dilution had the least amount of variation between C_t values, suggesting a DNA template concentration of approximately 10 ng/µL would be ideal for this qPCR assay. The 1/64 dilution C_t values for actin and *Lactobacillus spp*. were used to calculate the ΔC_t and normalize amplification (Table 6). Now that this assay has been validated using a single DNA template, it can be used with samples from WT and HD mice to determine the relative abundance of *Lactobacillus spp*. in the gut using the $\Delta\Delta C_t$ method.

Table 5. qPCR assay development C_t results

Dilution	Actin			*Clostridium spp.*			*Lactobacillus spp.*			*Leuconostoc mesenteroides*		
1	N/A	27.14	25.34	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1/4	24.36	28.15	23.48	N/A	N/A	N/A	29.98	31.48	31.46	N/A	N/A	N/A
1/16	26.11	25.41	25.04	N/A	N/A	N/A	27.31	27.72	26.70	N/A	N/A	N/A
1/64	27.28	26.98	26.87	N/A	N/A	N/A	27.94	28.03	27.60	N/A	N/A	N/A

Table 6. qPCR assay development ΔC_t calculations

	C_t	Average C_t	ΔC_t
Actin	27.28		
	26.98	27.04	
	26.87		
			0.82
Lactobacillus spp.	27.94		
	28.03	27.86	
	27.60		

Aim 3: To Identify and Quantify Bacterial Candidates in Mouse Stool and Cecal Samples

With a PCR and qPCR assay developed, bacterial targets can be identified and quantified in mouse stool and cecum samples. The nested PCR assay was run on stool samples from an N of 4 WT and HD mice at time point 1 (Figure 12). Each HD stool sample displays a DNA band in the second nested PCR reaction's gel, demonstrating the presence of *fldC*-containing bacteria

in these samples, while each WT stool sample appears empty. DNA bands are absent in the first gel, but present in the second gel, suggesting that our bacterial targets are low in quantity, requiring the use of the second nested PCR reaction to visualize them.

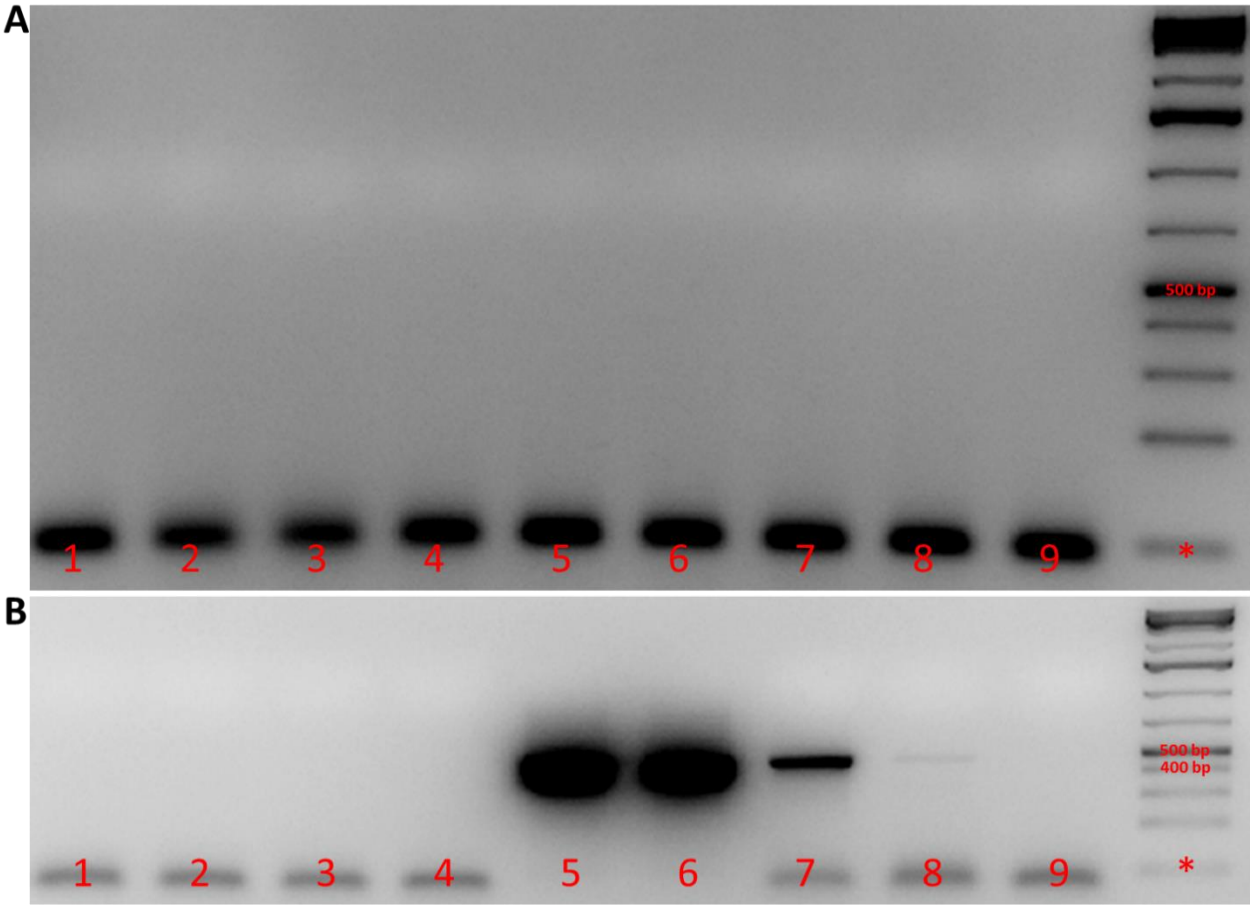

Figure 12. Nested PCR results for time point 1 stool samples. Both reactions of the nested PCR assay using WT and HD mouse stool DNA from time point 1. Lanes 1 – 4: WT mice. Lanes 5 – 8: HD mice. Lane 9: Negative control. Lane *: GeneRuler 1 kb DNA Ladder. (A) Nested PCR reaction one. Expected amplicon: 600 bp. (B) Nested PCR reaction two. Expected amplicon: 480 bp.

The nested PCR assay was also performed on seven healthy human stool samples (Figure 13) to validate its use in human stool. Two of the seven samples tested positive for *fldC*-containing bacteria. Again, the first PCR reaction did not produce visible product, suggesting that our bacterial targets are also present in low quantity within humans.

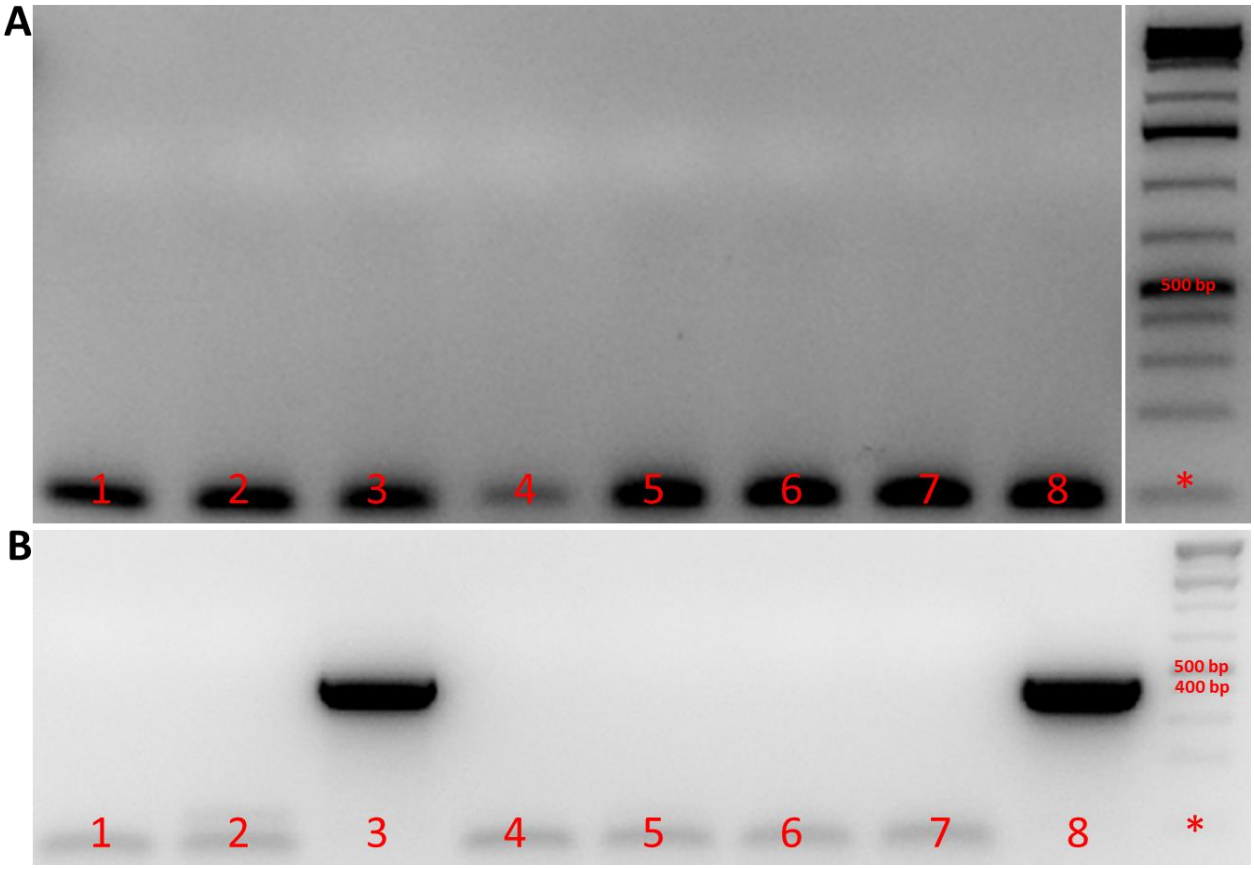

Figure 13. Nested PCR results for human stool samples. Both reactions of the nested PCR assay using human stool DNA. Lane 1: Negative control. Lanes 2 – 8: Healthy human stool. Lane *: GeneRuler 1 kb DNA Ladder. (A) Nested PCR reaction one. Expected amplicon: 600 bp. (B) Nested PCR reaction two. Expected amplicon: 480 bp.

CHAPTER THREE: DISCUSSION

Our lab has previously shown that HTT plays a role in gut-brain feedback. We hypothesize that the gut microbiome may regulate this feedback loop, and thus play a role in HD pathogenesis. We sought to investigate this by identifying exclusively microbiota-derived metabolites previously shown to be altered in HD plasma[28]. By identifying and quantifying bacterial producers of said metabolites, we aimed to measure genotype-associated alterations of microbial concentration within the gut of an HD model mouse.

In an effort to do this, we collected stool and cecal samples from WT and HD mice either before or after the onset of HD-like signs and neurodegeneration. Identification and quantitative assays were developed and validated to target the known bacterial producers of the exclusively microbiota-derived metabolite, IPA. Our literature review uncovered eight currently known bacterial producers of IPA, all from the *Clostridium* and *Peptostreptococcus* genera, both commonly found in the human gut microbiome[21].

With these two genera being so highly represented in the gut, we expected *fldC*-containing bacteria to be within a greater number of stool samples. The limit of detection for the *fldC* nested PCR assay was conservatively determined to be as low as 1.45 fg. It has been shown that approximately 8.4 fg of DNA is present within each bacterial cell in soil samples[61]. In addition, if we use *C. sporogenes* as an example, which has a genome of approximately 4.1 Mb[62], we can estimate that the *C. sporogenes* genome weighs approximately 4.47 fg, based on an average base pair weighing 650 daltons. Assuming our bacterial targets in stool and cecum have a similar amount of DNA, then the nested PCR assay developed here could be expected to detect as little as one bacterial cell.

Furthermore, the *fldC* gene is 1,125 bp, making it 0.027% (1.21 ag) of the *C. sporogenes* genome. Despite variable results, a DNA band in the ~2.33 ag serial dilution was observed. If this amount is accepted as the limit of detection for the *fldC* nested PCR assay, then it might be possible to detect as little as one copy of the *fldC* gene within a sample. With that being said, any samples showing a negative result are most likely due to lack of *fldC*-containing bacteria in the sample and not because of a detection failure of the assay.

We were able to detect *fldC* in mouse stool and can now work toward quantitative studies. The developed qPCR assay will be performed on the currently identified samples, as well as samples collected at additional time points and from a greater number of animals to quantify any relative abundance alterations across genotype and age. Cecum samples will also be analyzed to determine the effect of microbiota source on bacterial target composition, though this approach will not be able to be employed in humans due to the invasiveness of cecal sample collection.

A targeted approach to identify and quantify bacteria was selected for this study, but a non-targeted approach via 16S rRNA gene sequencing could also be used to determine the relative abundance of all bacterial phyla or genera within each sample. Targeted studies are beneficial when specific genus- or species-level specificity is desired, but this approach is low throughput and requires prior target microbe identification. A non-targeted approach mass sequences the entire gut microbiome, however that comes at the cost of specificity. Usually only phylum-level alterations in abundance can be assessed.

In this study, potential target metabolites were chosen from a previous HD plasma metabolomics paper[28], and was therefore limited by the number of significantly altered metabolites found within HD plasma. Our HD model mouse may not perfectly recapitulate the

HD metabolome or gut microbiome. In future studies, the Q175FDN HD model mouse plasma metabolome can be explored to determine if HD-like metabolite alterations are present.

Future research could also investigate the role of the gut microbiome in HD pathogenesis by comparing the disease progression of HD model mice with normal gut microbiomes versus germ-free or antibiotic-treated HD model mice. Any change in the rate of HD progression among the microbiota-depleted HD model mice could be attributed to the gut microbiome. The effects of fecal transplants on HD progression could also be investigated. For instance, comparing the onset and progression rate of HD-related phenotypes in HD model mice that receive a fecal transplant from an HD patient to mice that receive fecal transplants from a healthy controls.

It is too early to draw conclusions from the preliminary results gathered here, but the assays developed show promise in investigating IPA-producing bacteria in the gut. In future studies we will use these and other methods to further investigate the role of the gut microbiome in HD pathogenesis.

CHAPTER FOUR: METHODOLOGY

In Silico Bacterial Target Search

Metabolites within a pathway can be very similar and nomenclature may vary, so each metabolite from Table 1 was verified using PubChem[63], a chemical database, and HMDB[64], a human metabolite database. Once the exact metabolite was confirmed, each compound was found within MetaCyc[36], a metabolic pathway database. Selecting a reaction known to produce the target metabolite displays known organisms that possess that pathway, as well as suspected genes expressing key pathway enzymes. This provides a starting point to find relevant literature related to these pathways and organisms using PubMed. Potential genes found through literature review were verified using NCBI's gene database, which yields more relevant literature. Studies on each candidate gene, pathway, organism, and metabolite were searched to verify their absence in humans. All candidate bacterial targets must have a primary study validating its ability to produce the target metabolite.

Bacterial Culture Growth

The following bacterial cultures were grown: *Clostridium sporogenes* (ATCC 3584), *Clostridium difficile* (ATCC 9689), *Escherichia coli* (ATCC 8739), *Pseudomonas aeruginosa* (ATCC 27853), *Alcaligenes faecalis* (ATCC 8750), *Enterobacter aerogenes* (ATCC 13048), *Staphylococcus aureus* (ATCC 25923), and *Serratia marcescens* (ATCC 14756).

C. sporogenes and *C. difficile* were inoculated in tryptic soy broth (BD Biosciences) within an anaerobic chamber and grown at 37°C with approximately 80% N_2, 10% CO_2, and 10% H_2 for 48 hours. The broth was centrifuged at 10,000 × g for 10 min. The supernatant was discarded, and the pellet was reabsorbed in at least 5 volumes of RNA*later* (Invitrogen). *E. coli*

was grown on endo agar at 37°C for 48 hours. *P. aeruginosa*, *A. faecalis*, *E. aerogenes*, *S. aureus*, and *S. marcescens* were grown on brain heart infusion agar at 37°C for 48 hours. Bacterial colonies grown on agar were collected and placed in at least 5 volumes of RNA*later*. Bacterial samples were stored at 4°C overnight to allow full penetration of the RNA*later* before being moved to -80°C for long-term storage.

Total DNA was extracted from bacterial cells using the DNeasy Blood & Tissue Kit (Qiagen), following the manufacturer's instructions. Extracted DNA was stored at -20°C until analysis. DNA concentration and purity were measured via spectrophotometry using a NanoDrop 8000 (Thermo Scientific). DNA concentrations of >10 ng/µL and absorbance 260/280 and 260/230 ratios of ~1.8 – 2.2 were preferred, but due to the scarcity of samples, only largely abnormal samples were discarded.

Mouse Stool and Cecal Collection

All animals were handled in accordance with the University of Central Florida Institutional Animal Care and Use Committee guidelines under an approved animal use protocol. Mice were fed *ad libitum* and group-housed in a specific-pathogen-free facility with a 12 hour light 12 hour dark light cycle.

Stool and cecal samples were collected from an N of 8 HD or WT mice aged to < 3 months (pre-symptomatic) and > 9 months (post-symptomatic). To collect stool, mice were temporarily isolated in a sterilized cage until they defecated. Stool was collected and immediately placed in at least 5 volumes of RNA*later* to preserve DNA integrity. To collect cecal contents, a laparotomy was-performed. The cecum was surgically removed and the cecal contents squeezed into a 2.0 mL microcentrifuge tube containing at least 5 volumes of RNA*later*.

Samples were stored at 4°C overnight to allow complete RNA*later* penetration of samples before long-term storage at -80°C. Total genomic DNA was extracted from all stool and cecal samples using the QIAamp Fast DNA Stool Mini Kit (Qiagen), following the manufacturer's instructions. Extracted DNA was stored at -20°C until analysis. DNA concentration and purity were measured via spectrophotometry using a NanoDrop 8000 with the same DNA concentration and purity criteria as the bacterial culture DNA.

PCR

PCR reactions were performed in a total volume of 25 µL using a SimpliAmp Thermal Cycler (Applied Biosystems) and GoTaq G2 Green Master Mix (Promega). Reactions contained a final concentration of 1.25 U of Taq DNA polymerase, 200 µM of each deoxynucleoside triphosphate (G, A, T, and C), 1.5 mM magnesium chloride, 0.4 µM of each primer, and approximately 50 ng of DNA template. Amplification consisted of an initial denaturation step at 94°C for 5 min, followed by 40 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 1 min for the 16S rRNA gene primer pair or 60°C for the *fldC* primer pair, and extension at 72°C for 90 sec. Cycling included a final extension step at 72°C for 10 min, followed by cooling to 4°C. PCR products were separated via agarose gel electrophoresis at 5 V/cm for 20 minutes. Gels contained 0.5 µg/mL of ethidium bromide in 1.5% (w/v) agarose/1X TBE buffer (100 mM tris, 100 mM boric acid, and 20 mM EDTA). DNA bands were visualized under UV light using a Gel Doc XR+ Imager (Bio-Rad). Expected amplicon lengths for the *fldC* and 16S primer pairs of 600 bp and ~1,500 bp, respectively, were assessed using the GeneRuler 1 kb DNA Ladder (Thermo Scientific). All primer pairs used are summarized in Table 8.

To enhance detection of *fldC* even if present at low abundance, a second, nested PCR reaction was performed using 0.5 µL of the first round PCR product as the DNA template and a 55°C annealing temperature. Two nested primer pairs were developed with expected amplicon lengths of 81 bp and 480 bp, respectively. All other reagents, cycling conditions, and visualization procedures remained the same.

C. *sporogenes* Serial Dilution

A 5x dilution series of *C. sporogenes* DNA was generated beginning with 14.2 ng spiked into 50 ng of WT mouse stool DNA. Dilutions from 1/5 (~2.84 ng) to 1/6,103,515,625 (~2.33 ag) were used. Both nested PCR reactions were performed as stated above.

Gram Positive DNA Extraction Verification

Using a 10 µL stem-loop, pure *C. sporogenes* culture was spiked into a WT HD model mouse stool sample. DNA was extracted from this spiked stool sample using a QIAamp Fast DNA Stool Mini Kit, following the manufacturer's instructions. DNA from a pure *C. sporogenes* culture sample was also extracted as a positive control. DNA concentration and purity were measured via spectrophotometry. Only the first nested PCR reaction was required to visualize amplification.

qPCR Assay Development

A qPCR assay was developed using a genus-specific *Clostridium spp.* primer pair[58], a genus-specific *Lactobacillus spp.* primer pair[59], a *Leuconostoc mesenteroides* primer pair[60], and an actin-specific primer pair previously optimized and validated by our lab as a reference gene

(Table 8). Three separate cecal DNA samples from WT mice were pooled together to create a DNA template with a concentration of 785 ng/µL. Each reaction was run in triplicate and four different dilutions (1, 1/4, 1/16, and 1/64) were used with each primer pair. The amplification cycle consisted of an initial denaturation step at 95°C for 10 min, followed by 40 cycles of denaturation at 95°C for 15 sec and a combined annealing/elongation step at 60°C for 1 min. C_t values were determined using SDS Software v2.4 for 7900HT Fast.

Table 7. Primers used throughout study

Primer Target	Sequence (5' – 3')	Amplicon
16S rRNA gene[57]	F: AGA GTT TGA TCM* TGG CTC AG R: CGG TTA CCT TGT TAC GAC TT	~1,500 bp
fldC gene[42]	F: TGG GGA ATA TGA TAT GTT GTC TGG CAT GAT G R: TGT TCA GCT AAT CTA TCC ATT GGT GTA TTC GC	600 bp
fldC amplicon	F: ACC AGT AAC AGT TCC ACA AAA CAG R: TCC ATT GGT GTA TTC GCA TCT G	480 bp
fldC amplicon	F: TGG GAC AAA ATT TTA AAC TAA CCG T R: CAG CTT CCA TTT TTC TGT TTT GTG G	81 bp
Actin#	F: GGA GAC GGG GTC ACC CAC AC R: AGC CTC AGG GCA TCG GAA CC	450 bp
Clostridium spp.[58#]	F: AAA GGA AGA TTA ATA CCG CAT AA R: TTC TTC CTA ATC TCT ACG CA	536 bp
Lactobacillus spp.[59#]	F: CTC AAA ACT AAA CAA AGT TTC R: CTT GTA CAC ACC GCC CGT CA	209 bp
Leuconostoc mesenteroides[60#]	F: AAC TTA GTG TCG CAT GAC R: AGT CGA GTT GCA GAC TAC AA	1,152 bp

* M denotes either an A or C.
# Primer only used for assay validation.

$\Delta\Delta C_t$ analysis

Comparing the amplification of the 16S rRNA reference gene to the gene of interest will normalize the quantification of *fldC* to all bacteria within a sample. The amplification of the 16S rRNA gene and *fldC* via qPCR will produce a sigmoidal graph. The cycle number where the fluorescence of the amplified target DNA is higher than the background noise is known as the cycle threshold (C_t). Since each sample is run in triplicate, there will be three C_t values for each gene of each sample, which will be averaged together and excluding any outliers. Calculating the difference between the average C_t of the gene of interest and the average C_t of the reference gene (ΔC_t) will normalize the gene of interest. This is performed for both HD and WT samples. Then the difference between the HD ΔC_t and the WT ΔC_t is calculated, creating the $\Delta\Delta C_t$. Finally, calculating $2^{-(\Delta\Delta C_t)}$ will provide the relative quantification of the IPA-producing bacteria present within the samples.

Relative quantifications will be group analyzed as the dependent variable, alongside age and genotype as the independent variables. Two-way ANOVA will be performed, followed by Bonferroni post-hoc analysis.

LIST OF REFERENCES

1. The Huntington's Disease Collaborative Research Group. (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell, 72*(6), 971-983.

2. Snell, R. G., MacMillan, J. C., Cheadle, J. P., Fenton, I., Lazarou, L. P., Davies, P., MacDonald, M. E., Gusella, J. F., Harper, P. S., and Shaw, D. J. (1993). Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease. *Nat Genet, 4*(4), 393-397.

3. Duyao, M., Ambrose, C., Myers, R., Novelletto, A., Persichetti, F., Frontali, M., Folstein, S., Ross, C., Franz, M., Abbott, M., Gray, J., Conneally, P., Young, A., Penney, J., Hollingsworth, Z., Shoulson, I., Lazzarini, A., Falek, A., Koroshetz, W., Sax, D., Bird, E., Vonsattel, J., Bonilla, E., Alvir, J., Bickham Conde, J., Cha, J. H., Dure, L., Gomez, F., Ramos, M., Sanchez-Ramos, J., Snodgrass, S., de Young, M., Wexler, N., Moscowitz, C., Penchaszadeh, G., MacFarlane, H., Anderson, M., Jenkins, B., Srinidhi, J., Barnes, G., Gusella, J., and MacDonald, M. (1993). Trinucleotide repeat length instability and age of onset in Huntington's disease. *Nat Genet, 4*, 387.

4. Andrew, S. E., Paul Goldberg, Y., Kremer, B., Telenius, H., Theilmann, J., Adam, S., Starr, E., Squitieri, F., Lin, B., Kalchman, M. A., Graham, R. K., and Hayden, M. R. (1993). The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. *Nat Genet, 4*, 398.

5. Falush, D., Almqvist, E. W., Brinkmann, R. R., Iwasa, Y., and Hayden, M. R. (2001). Measurement of Mutational Flow Implies Both a High New-Mutation Rate for Huntington Disease and Substantial Underascertainment of Late-Onset Cases. *The American Journal of Human Genetics, 68*(2), 373-385.

6. Ghosh, R., and Tabrizi, S. J. (2015). Clinical Aspects of Huntington's Disease. In H. H. P. Nguyen and M. A. Cenci (Eds.), *Behavioral Neurobiology of Huntington's Disease and Parkinson's Disease* (pp. 3-31). Berlin, Heidelberg: Springer Berlin Heidelberg.

7. Walker, F. O. (2007). Huntington's disease. *The Lancet, 369*(9557), 218-228.

8. Dayalu, P., and Albin, R. L. (2015). Huntington disease: pathogenesis and treatment. *Neurol Clin, 33*(1), 101-114.

9. Sanberg, P. R., Fibiger, H. C., and Mark, R. F. (1981). Body weight and dietary factors in Huntington's disease patients compared with matched controls. *Med J Aust, 1*(8), 407-409.

10. Aziz, N. A., Swaab, D. F., Pijl, H., and Roos, R. A. (2007). Hypothalamic dysfunction and neuroendocrine and metabolic alterations in Huntington's disease: clinical consequences and therapeutic implications. *Rev Neurosci, 18*(3-4), 223-251.

11. Myers, R. H., Sax, D. S., Koroshetz, W. J., Mastromauro, C., Cupples, L. A., Kiely, D. K., Pettengill, F. K., and Bird, E. D. (1991). Factors associated with slow progression in Huntington's disease. *Arch Neurol, 48*(8), 800-804.

12. Pratley, R. E., Salbe, A. D., Ravussin, E., and Caviness, J. N. (2000). Higher sedentary energy expenditure in patients with Huntington's disease. *Ann Neurol, 47*(1), 64-70.

13. Schilling, G., Becher, M. W., Sharp, A. H., Jinnah, H. A., Duan, K., Kotzuk, J. A., Slunt, H. H., Ratovitski, T., Cooper, J. K., Jenkins, N. A., Copeland, N. G., Price, D. L., Ross, C. A., and Borchelt, D. R. (1999). Intranuclear Inclusions and Neuritic Aggregates in Transgenic Mice Expressing a Mutant N-Terminal Fragment of Huntingtin. *Hum Mol Genet, 8*(3), 397-407.

14. Mangiarini, L., Sathasivam, K., Seller, M., Cozens, B., Harper, A., Hetherington, C., Lawton, M., Trottier, Y., Lehrach, H., Davies, S. W., and Bates, G. P. (1996). Exon 1 of the HD Gene with an Expanded CAG Repeat Is Sufficient to Cause a Progressive Neurological Phenotype in Transgenic Mice. *Cell, 87*(3), 493-506.

15. Southwell, A. L., Smith-Dijak, A., Kay, C., Sepers, M., Villanueva, E. B., Parsons, M. P., Xie, Y., Anderson, L., Felczak, B., Waltl, S., Ko, S., Cheung, D., Dal Cengio, L., Slama, R., Petoukhov, E., Raymond, L. A., and Hayden, M. R. (2016). An enhanced Q175 knock-in mouse model of Huntington disease with higher mutant huntingtin levels and accelerated disease phenotypes. *Hum Mol Genet, 25*(17), 3654-3675.

16. Hodgson, J. G., Agopyan, N., Gutekunst, C.-A., Leavitt, B. R., LePiane, F., Singaraja, R., Smith, D. J., Bissada, N., McCutcheon, K., Nasir, J., Jamot, L., Li, X.-J., Stevens, M. E., Rosemond, E., Roder, J. C., Phillips, A. G., Rubin, E. M., Hersch, S. M., and Hayden, M. R. (1999). A YAC Mouse Model for Huntington’s Disease with Full-Length Mutant Huntingtin, Cytoplasmic Toxicity, and Selective Striatal Neurodegeneration. *Neuron, 23*(1), 181-192.

17. Gray, M., Shirasaki, D. I., Cepeda, C., André, V. M., Wilburn, B., Lu, X.-H., Tao, J., Yamazaki, I., Li, S.-H., Sun, Y. E., Li, X.-J., Levine, M. S., and Yang, X. W. (2008). Full-Length Human Mutant Huntingtin with a Stable Polyglutamine Repeat Can Elicit Progressive and Selective Neuropathogenesis in BACHD Mice. *The Journal of Neuroscience, 28*(24), 6182-6195.

18. Southwell, A. L., Warby, S. C., Carroll, J. B., Doty, C. N., Skotte, N. H., Zhang, W., Villanueva, E. B., Kovalik, V., Xie, Y., Pouladi, M. A., Collins, J. A., Yang, X. W., Franciosi, S., and Hayden, M. R. (2013). A fully humanized transgenic mouse model of Huntington disease. *Hum Mol Genet, 22*(1), 18-34.

19. Southwell, A. L., Skotte, N. H., Villanueva, E. B., Østergaard, M. E., Gu, X., Kordasiewicz, H. B., Kay, C., Cheung, D., Xie, Y., Waltl, S., Dal Cengio, L., Findlay-Black, H., Doty, C. N., Petoukhov, E., Iworima, D., Slama, R., Ooi, J., Pouladi, M. A., Yang, X. W., Swayze, E. E., Seth, P. P., and Hayden, M. R. (2017). A novel humanized mouse model of Huntington disease for preclinical development of therapeutics targeting mutant huntingtin alleles. *Hum Mol Genet, 26*(6), 1115-1132.

20. O'Hara, A. M., and Shanahan, F. (2006). The gut flora as a forgotten organ. *EMBO Rep, 7*(7), 688-693.

21. Guarner, F., and Malagelada, J.-R. (2003). Gut flora in health and disease. *The Lancet, 361*(9356), 512-519.

22. Nicholson, J. K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., and Pettersson, S. (2012). Host-gut microbiota metabolic interactions. *Science, 336*(6086), 1262-1267.

23. Wang, Y., and Kasper, L. H. (2014). The role of microbiome in central nervous system disorders. *Brain Behav Immun, 38*, 1-12.

24. Tang, A. T., Choi, J. P., Kotzin, J. J., Yang, Y., Hong, C. C., Hobson, N., Girard, R., Zeineddine, H. A., Lightle, R., Moore, T., Cao, Y., Shenkar, R., Chen, M., Mericko, P., Yang, J., Li, L., Tanes, C., Kobuley, D., Võsa, U., Whitehead, K. J., Li, D. Y., Franke, L., Hart, B., Schwaninger, M., Henao-Mejia, J., Morrison, L., Kim, H., Awad, I. A., Zheng, X., and Kahn, M. L. (2017). Endothelial TLR4 and the microbiome drive cerebral cavernous malformations. *Nature, 545*, 305.

25. Hsiao, E. Y., McBride, S. W., Hsien, S., Sharon, G., Hyde, E. R., McCue, T., Codelli, J. A., Chow, J., Reisman, S. E., Petrosino, J. F., Patterson, P. H., and Mazmanian, S. K. (2013). Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. *Cell, 155*(7), 1451-1463.

26. Sharon, G., Garg, N., Debelius, J., Knight, R., Dorrestein, P. C., and Mazmanian, S. K. (2014). Specialized metabolites from the microbiome in health and disease. *Cell Metab, 20*(5), 719-730.

27. Sampson, T. R., Debelius, J. W., Thron, T., Janssen, S., Shastri, G. G., Ilhan, Z. E., Challis, C., Schretter, C. E., Rocha, S., Gradinaru, V., Chesselet, M. F., Keshavarzian, A., Shannon, K. M., Krajmalnik-Brown, R., Wittung-Stafshede, P., Knight, R., and Mazmanian, S. K. (2016). Gut Microbiota Regulate Motor Deficits and Neuroinflammation in a Model of Parkinson's Disease. *Cell, 167*(6), 1469-1480.e1412.

28. Rosas, H. D., Doros, G., Bhasin, S., Thomas, B., Gevorkian, S., Malarick, K., Matson, W., and Hersch, S. M. (2015). A systems-level "misunderstanding": the plasma metabolome in Huntington's disease. *Ann Clin Transl Neurol, 2*(7), 756-768.

29. Beal, M. F., Matson, W. R., Swartz, K. J., Gamache, P. H., and Bird, E. D. (1990). Kynurenine Pathway Measurements in Huntington's Disease Striatum: Evidence for Reduced Formation of Kynurenic Acid. *Journal of Neurochemistry, 55*(4), 1327-1339.

30. Stoy, N., Mackay, G. M., Forrest, C. M., Christofides, J., Egerton, M., Stone, T. W., and Darlington, L. G. (2005). Tryptophan metabolism and oxidative stress in patients with Huntington's disease. *Journal of Neurochemistry, 93*(3), 611-623.

31. Pearson, S. J., and Reynolds, G. P. (1992). Increased brain concentrations of a neurotoxin, 3-hydroxykynurenine, in Huntington's disease. *Neuroscience Letters, 144*(1), 199-201.

32. Kumar, R., Eipers, P., Little, R. B., Crowley, M., Crossman, D. K., Lefkowitz, E. J., and Morrow, C. D. (2014). Getting started with microbiome analysis: sample acquisition to bioinformatics. *Curr Protoc Hum Genet, 82*, 18.18.11-29.

33. Pang, W., Vogensen, F. K., Nielsen, D. S., and Hansen, A. K. (2012). Faecal and caecal microbiota profiles of mice do not cluster in the same way. *Laboratory Animals, 46*(3), 231-236.

34. Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative CT method. *Nature Protocols, 3*, 1101.

35. Krych, L., Hansen, C. H. F., Hansen, A. K., van den Berg, F. W. J., and Nielsen, D. S. (2013). Quantitatively Different, yet Qualitatively Alike: A Meta-Analysis of the Mouse Core Gut Microbiome with a View towards the Human Gut Microbiome. *PLoS One, 8*(5), e62578.

36. Caspi, R., Altman, T., Dreher, K., Fulcher, C. A., Subhraveti, P., Keseler, I. M., Kothari, A., Krummenacker, M., Latendresse, M., Mueller, L. A., Ong, Q., Paley, S., Pujar, A., Shearer, A. G., Travers, M., Weerasinghe, D., Zhang, P., and Karp, P. D. (2012). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res, 40*(Database issue), D742-d753.

37. Hofius, D., and Sonnewald, U. (2003). Vitamin E biosynthesis: biochemistry meets cell biology. *Trends in Plant Science, 8*(1), 6-8.

38. Sattler, S. E., Cahoon, E. B., Coughlan, S. J., and DellaPenna, D. (2003). Characterization of Tocopherol Cyclases from Higher Plants and Cyanobacteria. Evolutionary Implications for Tocopherol Synthesis and Function. *Plant Physiology, 132*(4), 2184-2195.

39. Stewart, I., Schluter, P. J., and Shaw, G. R. (2006). Cyanobacterial lipopolysaccharides and human health – a review. *Environmental Health, 5*(1), 7.

40. Di Rienzi, S. C., Sharon, I., Wrighton, K. C., Koren, O., Hug, L. A., Thomas, B. C., Goodrich, J. K., Bell, J. T., Spector, T. D., Banfield, J. F., and Ley, R. E. (2013). The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria. *Elife, 2*, e01102.

41. Roager, H. M., and Licht, T. R. (2018). Microbial tryptophan catabolites in health and disease. *Nat Commun, 9*(1), 3294.

42. Dodd, D., Spitzer, M. H., Van Treuren, W., Merrill, B. D., Hryckowian, A. J., Higginbottom, S. K., Le, A., Cowan, T. M., Nolan, G. P., Fischbach, M. A., and Sonnenburg, J. L. (2017). A gut bacterial pathway metabolizes aromatic amino acids into nine circulating metabolites. *Nature, 551*(7682), 648-652.

43. Poeggeler, B., Pappolla, M. A., Hardeland, R., Rassoulpour, A., Hodgkins, P. S., Guidetti, P., and Schwarcz, R. (1999). Indole-3-propionate: a potent hydroxyl radical scavenger in rat brain. *Brain Res, 815*(2), 382-388.

44. Chyan, Y. J., Poeggeler, B., Omar, R. A., Chain, D. G., Frangione, B., Ghiso, J., and Pappolla, M. A. (1999). Potent neuroprotective properties against the Alzheimer beta-amyloid by an endogenous melatonin-related indole structure, indole-3-propionic acid. *J Biol Chem, 274*(31), 21937-21942.

45. Kumar, A., and Ratan, R. R. (2016). Oxidative Stress and Huntington's Disease: The Good, The Bad, and The Ugly. *Journal of Huntington's disease, 5*(3), 217-237.

46. Feigin, A., Kieburtz, K., Como, P., Hickey, C., Abwender, D., Zimmerman, C., Steinberg, K., and Shoulson, I. (1996). Assessment of coenzyme q10 tolerability in huntington's disease. *Movement Disorders, 11*(3), 321-323.

47. The Huntington Study Group. (2001). A randomized, placebo-controlled trial of coenzyme Q_{10} and remacemide in Huntington's disease. *Neurology, 57*(3), 397-404.

48. Hersch, S. M., Gevorkian, S., Marder, K., Moskowitz, C., Feigin, A., Cox, M., Como, P., Zimmerman, C., Lin, M., Zhang, L., Ulug, A. M., Beal, M. F., Matson, W., Bogdanov, M., Ebbel, E., Zaleta, A., Kaneko, Y., Jenkins, B., Hevelone, N., Zhang, H., Yu, H., Schoenfeld, D., Ferrante, R., and Rosas, H. D. (2006). Creatine in Huntington disease is safe, tolerable, bioavailable in brain and reduces serum 8OH2′dG. *Neurology, 66*(2), 250-252.

49. Huntington Study Group TREND-HD Investigators. (2008). Randomized controlled trial of ethyl-eicosapentaenoic acid in huntington disease: The trend-hd study. *Arch Neurol, 65*(12), 1582-1589.

50. Ranen, N. G., Peyser, C. E., Coyle, J. T., Bylsma, F. W., Sherr, M., Day, L., Folstein, M. F., Brandt, J., Ross, C. A., and Folstein, S. E. (1996). A controlled trial of idebenone in Huntington's disease. *Movement Disorders, 11*(5), 549-554.

51. Bjorkqvist, M., Wild, E. J., Thiele, J., Silvestroni, A., Andre, R., Lahiri, N., Raibon, E., Lee, R. V., Benn, C. L., Soulet, D., Magnusson, A., Woodman, B., Landles, C., Pouladi, M. A., Hayden, M. R., Khalili-Shirazi, A., Lowdell, M. W., Brundin, P., Bates, G. P., Leavitt, B. R., Moller, T., and Tabrizi, S. J. (2008). A novel pathogenic pathway of immune activation detectable before clinical onset in Huntington's disease. *J Exp Med, 205*(8), 1869-1877.

52. Venkatesh, M., Mukherjee, S., Wang, H., Li, H., Sun, K., Benechet, A. P., Qiu, Z., Maher, L., Redinbo, M. R., Phillips, R. S., Fleet, J. C., Kortagere, S., Mukherjee, P., Fasano, A., Le Ven, J., Nicholson, J. K., Dumas, M. E., Khanna, K. M., and Mani, S. (2014). Symbiotic bacterial metabolites regulate gastrointestinal barrier function via the xenobiotic sensor PXR and Toll-like receptor 4. *Immunity, 41*(2), 296-310.

53. Elsden, S. R., Hilton, M. G., and Waller, J. M. (1976). The end products of the metabolism of aromatic amino acids by Clostridia. *Arch Microbiol, 107*(3), 283-288.

54. Smith, E. A., and Macfarlane, G. T. (1996). Enumeration of human colonic bacteria producing phenolic and indolic compounds: effects of pH, carbohydrate availability and retention time on dissimilatory aromatic amino acid metabolism. *J Appl Bacteriol, 81*(3), 288-302.

55. Wikoff, W. R., Anfora, A. T., Liu, J., Schultz, P. G., Lesley, S. A., Peters, E. C., and Siuzdak, G. (2009). Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc Natl Acad Sci U S A, 106*(10), 3698-3703.

56. Wlodarska, M., Luo, C., Kolde, R., d'Hennezel, E., Annand, J. W., Heim, C. E., Krastel, P., Schmitt, E. K., Omar, A. S., Creasey, E. A., Garner, A. L., Mohammadi, S., O'Connell, D. J., Abubucker, S., Arthur, T. D., Franzosa, E. A., Huttenhower, C., Murphy, L. O., Haiser, H. J., Vlamakis, H., Porter, J. A., and Xavier, R. J. (2017). Indoleacrylic Acid Produced by Commensal Peptostreptococcus Species Suppresses Inflammation. *Cell Host & Microbe, 22*(1), 25-37.e26.

57. Weisburg, W. G., Barns, S. M., Pelletier, D. A., and Lane, D. J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol, 173*(2), 697-703.

58. Hung, C.-H., Cheng, C.-H., Cheng, L.-H., Liang, C.-M., and Lin, C.-Y. (2008). Application of Clostridium-specific PCR primers on the analysis of dark fermentation

hydrogen-producing bacterial community. *International Journal of Hydrogen Energy, 33*(5), 1586-1592.

59. Dubernet, S., Desmasures, N., and Gueguen, M. (2002). A PCR-based method for identification of lactobacilli at the genus level. *FEMS Microbiol Lett, 214*(2), 271-275.

60. Lee, H. J., Park, S. Y., and Kim, J. (2000). Multiplex PCR-based detection and identification of Leuconostoc species. *FEMS Microbiol Lett, 193*(2), 243-247.

61. Torsvik, V. L., and Goksoyr, J. (1978). Determination of bacterial DNA in soil. *Soil Biology and Biochemistry, 10*(1), 7-12.

62. Poehlein, A., Riegel, K., König, S. M., Leimbach, A., Daniel, R., and Dürre, P. (2015). Genome sequence of Clostridium sporogenes DSM 795T, an amino acid-degrading, nontoxic surrogate of neurotoxin-producing Clostridium botulinum. *Standards in Genomic Sciences, 10*(1), 40.

63. Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S., Shoemaker, B. A., Wang, J., Yu, B., Zhang, J., and Bryant, S. H. (2016). PubChem Substance and Compound databases. *Nucleic Acids Res, 44*(D1), D1202-D1213.

64. Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vazquez-Fresno, R., Sajed, T., Johnson, D., Li, C., Karu, N., Sayeeda, Z., Lo, E., Assempour, N., Berjanskii, M., Singhal, S., Arndt, D., Liang, Y., Badran, H., Grant, J., Serra-Cayuela, A., Liu, Y., Mandal, R., Neveu, V., Pon, A., Knox, C., Wilson, M., Manach, C., and Scalbert, A. (2018). HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res, 46*(D1), D608-d617.