The Effect of Bacterial High Temperature Protein G on Intestinal Epithelial Cells in the Context of Pediatric Crohn's Disease

by

Emma Finlayson-Trick

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
July 2019

Table of Contents

List of Tables

List of Figures

Abstract

Crohn's disease (CD), a type of inflammatory bowel disease, is caused by environmental, microbial, genetic, and immunological factors. Increased expression of specific heat shock proteins (HSPs) in CD patients protects intestinal epithelial cells (IECs) from death. *High temperature protein G* (*HtpG*), a gene encoding a bacterial HSP, is less abundant in pediatric CD patients than in healthy individuals. The role of HtpG in CD is poorly understood, thus I sought to describe HtpG activity on the innate immune response of IECs. I used metagenomic sequencing data to determine that only certain HtpG lineages appear in the intestines. By measuring IEC inflammatory mediators, I observed HtpG to be inflammatory, but with the capacity to dampen the effects of another inflammatory molecule, tumor necrosis factor. I also identified, by co-affinity purification, that HtpG interacts with three common CD-associated gene variants. These results suggest HtpG acts extra- and intra-cellularly to mediate inflammatory signaling.

List of Abbreviations Used

Amp	Ampicillin
ANOVA	Analysis of variance
AP-1	Activator protein-1
ATG16L1	Autophagy-related 16 like 1
ATP	Adenosine triphosphate
B. fragilis	*Bacteroides fragilis*
BSA	Bovine serum albumin
C	Carboxy
CARD	Caspase-recruitment domain
CCL	CC-motif chemokine ligand
CD	Crohn's disease
CD*n*	Cluster of differentiation
cDNA	Complementary deoxyribonucleic acid
cRAP	common Repository of Adventitious Proteins
CRC	Colorectal cancer
CXCL	CXC-motif chemokine ligand
DAMP	Damage associated molecular pattern
DMEM	Dulbecco's Modified Eagles Medium
DNA	Deoxyribonucleic acid
ECL	Enhanced chemiluminescence
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EEN	Exclusive enteral nutrition
ELISA	Enzyme-linked Immunosorbent Assay
FBS	Fetal bovine serum
HCl	Hydrochloric acid
HRP	Horseradish peroxidase
HSP	Heat shock protein
HtpG	High temperature protein H
GFP	Green fluorescent protein
GI	Gastrointestinal tract
grp94	Glucose-related protein 94
GWAS	Genome-wide association studies
IBD	Inflammatory bowel disease
IEL	Intraepithelial lymphocyte
IFN	Interferon
IκB	NF-κB inhibitor
IKK	IκB Kinase
IL	Interleukin
ILC	Innate lymphoid cell
IPTG	Isopropyl β-D-1-thiogalactopyranoside
JNK	c-Jun N-terminal kinases
Kan	Kanamycin
KEGG	Kyoto Encyclopedia of Genes and Genomes

LB	Lysogeny broth
LC-MS/MS	Liquid chromatography tandem mass-spectrometry
LPS	Lipopolysaccharide
LRR	Leucine-rich repeat
MAFFT	Multiple Alignment using Fast Fourier Transform
MAMP	Microbe associated molecular pattern
MAPK	Mitogen-activated protein kinases
M-cell	Microfold cells
MDP	Muramyl dipeptide
MDP_{DD}/MDP_{LL}	Inactive forms of muramyl dipeptide
miRNA	MicroRNA
M. mazei	*Methanosarcina mazei*
mRNA	Messenger RNA
N	Amino
NaCl	Sodium chloride
Na_2PO_4	Disodium phosphate
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NLR	Nucleotide-binding oligomerization domain-like receptor
NLRC	NLRs containing CARD
NLSR	NOD-like signaling receptor
NK cell	Natural killer cell
NKT cell	Natural killer T cell
NOD2	Nucleotide-binding oligomerization domain-containing protein 2
OMV	Outer membrane vesicle
PAMP	Pathogen associated molecular pattern
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
P. gingivalis	*Porphyromonas gingivalis*
PRR	Pattern recognition receptor
PVDF	Polyvinylidene difluoride
RAxML	Randomized Axelerated Maximum Likelihood
RIPA	Radioimmunoprecipitation assay
RIPK2	Receptor-interacting serine/threonine-protein kinase 2
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
SCFAs	Short-chain fatty acids
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEM	Standard error of the mean
SNPs	Single-nucleotide polymorphisms
TBS	Tris-buffered saline
TBST	Tris-buffered saline with Tween-20
TGF-β	Transforming growth factor-β
T_H cells	T helper cells
TIMP2	Tissue inhibitors of metalloproteinases
TLR	Toll-like receptor
TNF	Tumor necrosis factor

Trap1	TNF receptor associated protein 1
T_{reg} cells	T regulatory cells
V8	*NOD2* variant R702W
V12	*NOD2* variant G908R
V13	*NOD2* variant L1007fs
wPCDAI	Weighted Pediatric Crohn's Disease Activity Index
WT	Wildtype

Acknowledgements

A big thank you to all members of the MIRA lab past and present who made this work possible. Especially to Dr. Jessica Connors who patiently and enthusiastically answered every question I posed to her. Thanks to Scott Whitehouse who trained me in my early days in the lab, to Mushfiqir Rahman who will hopefully continue this project forward, and to the crew of undergraduate students who always made work fun.

Thank you to my committee members Dr. Joseph Bielawski and Dr. John Rohde for always being available to talk about my latest theories and results. Thank you as well to Dr. Katherine Dunn for teaching me how to make gene trees (using script no longer seems quite as scary).

The biggest thank you to my co-supervisors Drs. Andrew Stadnyk and Johan Van Limbergen. I seriously appreciate the time and effort you both have put into supporting me and this project. To Andy—thank you for our weekly meetings where we not only discussed my work, but also birds, cinnamon buns, pizza, and medical school. To Johan—thank you for taking me on as your first graduate student.

Thank you to the Natural Sciences and Engineering Research Council of Canada and the Cancer Research Training Program (through Dalhousie Medical Research Foundation and CIBC) for financially supporting my MSc work.

Lastly, I would like to thank my family, and my partner Lucas for being my loudest cheerleaders.

Chapter 1: Introduction

Spanning from the mouth to the anus, the digestive tract functions to ingest and digest food, absorb nutrients, and eliminate waste. Within the digestive tract, the gastrointestinal tract (GI), which includes the stomach and intestines, also functions to maintain homeostasis by differentiating between helpful (i.e., innocuous food, commensal microbes, and self-antigens) and harmful (i.e., toxins, invading pathogens) environmental signals. The mucosa, the innermost layer of the intestines that interacts with the environment, can be divided into the epithelium; a single layer of epithelial cells impregnated with intraepithelial lymphocytes (IELs); the lamina propria, containing connective tissue, fibroblasts, myofibroblasts, various immune cells, and blood and lymph vessels; and the muscularis mucosa, a thin layer of smooth muscle cells. The intestines, otherwise known as the guts, are a dynamic environment, relying on various sensory and response mechanisms to maintain homeostasis. Unfortunately, when homeostasis is disrupted (e.g., the immune system, malnutrition, bacterial overgrowth, etc.), diseases can arise. Inflammatory bowel disease (IBD), chiefly ulcerative colitis and Crohn's disease (CD), is an increasingly common condition of the GI tract. CD will be discussed in greater detail in section 1.2.

1.1: The Intestines from Macro to Micro

The small intestine, measuring 6-7 m, is divided into three distinct regions known as the duodenum, jejunum, and ileum (Mowat & Agace, 2014). The duodenum, in conjunction with bile from the gallbladder and enzymes from the pancreas, first neutralizes and then digests chyme that arrives from the stomach. The jejunum and ileum primarily absorb nutrients from digested food. The large intestine, measuring approximately 1.5 m, is divided into two regions called the cecum and colon (ascending, transverse, descending, and sigmoid) (Mowat & Agace,

2014). The large intestine mediates fluid balance by controlling water and salt reabsorption and elimination. Furthermore, the large intestine propels feces towards the rectum for defecation.

1.1.1: The Intestinal Epithelium

The intestinal epithelium is the largest mucosal surface within the body (Mowat & Agace, 2014; Peterson & Artis, 2014). Consequently, the epithelium acts as a physical barrier with chemical reinforcements, such as detergents, mucus, proteolytic enzymes (trypsin), cell-wall degrading enzymes (lysozyme), and antimicrobial peptides (defensins, etc.) (Chassaing, Kumar, Baker, Singh, & Vijay-Kumar, 2014). The epithelium consists of mucosal invaginations called crypts of Lieberkühn (Spence, Lauf, & Shroyer, 2011). In general, the epithelium is composed of several specialized cell lineages that arise from intestinal stem cells located at the base of crypts (Crosnier, Stamataki, & Lewis, 2006). Absorptive enterocytes, the most common cell type, are involved in metabolic and digestive processes (Peterson & Artis, 2014). Enteroendocrine cells, goblet cells, and Paneth cells maintain the epithelial barrier through secretion of hormone regulators, mucins, and antimicrobial peptides, respectively (Allaire et al., 2018). Microfold (M)-cells, located in lymphoid aggregates called Peyer's patches, indiscriminately sample luminal contents and transport intact antigens to underlying dendritic cells for antigen presentation (Ohno, 2016; Peterson & Artis, 2014). Less abundant are chemosensory cells (tuft cells) that regulate immune responses in the lamina propria by translating luminal signals into effector functions (Ting & von Moltke, 2019). Finally, IELs—comprised mainly of T-cells—bind to epithelial cells through integrins and contribute to pathogen resistance and barrier homeostasis (van Konijnenburg & Mucida, 2017).

The small and large intestines have architecturally distinct epitheliums with unique cellular compositions (Figure 1). Firstly, the small intestine has villi that protrude into the lumen.

Villi greatly increase surface area and therefore allow for optimal contact with and absorption of nutrients (Allaire et al., 2018). Secondly, the large intestine has a loose mucus layer that covers the firm mucus layer found in both intestines. The firm layer shields the epithelium from luminal microbes but not diffusing content, such as antimicrobial peptides. The bacterial-infused loose mucus layer moves with fecal pellets as they pass through the colon (Kamphuis, Mercier-Bonin, Eutamène, & Theodorou, 2017). Thirdly, Paneth cells are unique to the small intestine (Allaire et al., 2018; Bevins & Salzman, 2011). Finally, M-cells are found in greater abundance in the small intestine and goblet cells are found in greater abundance in the large intestine (Okumura & Takeda, 2017).

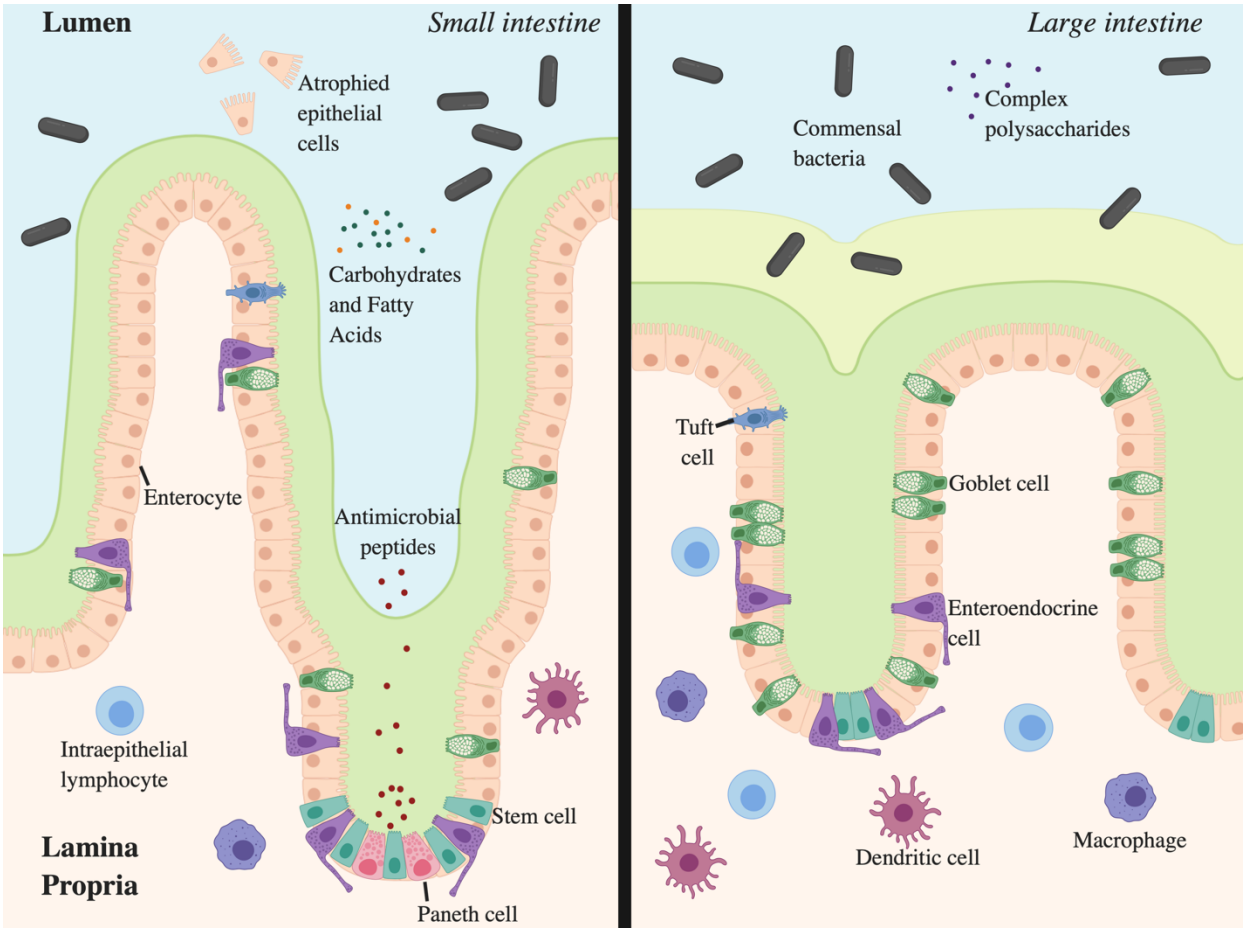


Figure 1. Schematic of small and large intestine. The epithelium of both intestines is composed of various cell types including enterocytes, Paneth cells, stem cells, tuft cells, goblet cells, and enteroendocrine cells. Immune cells lie within the lamina propria. The epithelium of the small and large intestine is covered in a layer of firm mucus that is penetrable by soluble factors, such as antimicrobial peptides. Within the large intestine, the firm mucus layer is covered by an additional bacterial-infused loose mucus layer that moves with the fecal pellet. The intestinal lumen not only contains microorganisms, but also nutrients derived from digested food and atrophied epithelial cells. Figure created with BioRender.

1.1.1.1: The Role of Intestinal Epithelial Cells in Immune Homeostasis

Coupled with acting as a barrier, the intestinal epithelium maintains homeostasis by sensing the external environment through pattern recognition receptors (PRRs). PRRs recognize pathogen (or microbe) associated molecular patterns (P[M]AMPS), such as glycans, as well as damage-associated molecular patterns (DAMPS), such as nuclear and cytoplasmic proteins. Epithelial cells can express a range of PRRs including membrane-bound toll-like receptors (TLRs) and cytoplasmic nucleotide-binding oligomerization domain-like receptors (NLRs).

In healthy individuals, PRRs interact with commensal microbial stimuli to educate the immune system and create a tolerogenic state. Upon interaction with pathogenic microbial stimuli, however, PRRs induce a pro-inflammatory immune response that activates immune cells such as dendritic cells, macrophages, and IELs (Geremia & Arancibia-Cárcamo, 2017; Rimoldi et al., 2005; Zeuthen, Fink, & Frokiaer, 2008). These cells have their own repertoire of PRR and can respond once they arrive at the site where the epithelium may be leaking or breached. In general, innate immune cells act as regulators of tissue homeostasis, inflammation, and the early response to infection. How epithelial cells mediate their tailored hypo-responsiveness, so as to avoid continuous inflammatory signaling, is a topic of extensive research (Fukata & Arditi, 2013). Epithelial cell PRRs appear to become hypo-responsive shortly after birth upon stimulation with appropriate ligands (Lotz et al., 2006). Additionally, PRR responsiveness is regulated by cellular localization (Allaire et al., 2018). As an example, basolateral TLR9 ligand exposure activates a pro-inflammatory response, whereas apical TLR9 ligand exposure induces tolerance to subsequent TLR stimulation (Lee et al., 2006).

Finally, the intestinal epithelium and lamina propria are major locations for adaptive immune cell accumulation (Agace & McCoy, 2017). Adaptive immune cell composition varies

greatly between intestinal sites with cluster of differentiation (CD*n*)8+ T cells dominating the epithelium and CD*n*4+ and CD*n*8+ T cells dominating the lamina propria (Agace & McCoy, 2017). Furthermore, Peyer's patches—organized into large B cell follicles, small T cell zones, and overlying follicle associated epithelium—are sites of lymphocyte activation and differentiation (Reboldi & Cyster, 2016). The adaptive immune system plays a crucial role in maintaining tolerance to commensal bacteria by establishing resident epithelial effector memory T cell populations (Agace & McCoy, 2017).

1.1.2: The Intestinal Microbiota

Within a given environmental niche, the collection of microorganisms and their cumulative genetic material are referred to as the microbiota and microbiome, respectively. Microbial cells are at least as abundant as somatic cells within the body and contribute far more genes than the human genome (Sender, Fuchs, & Milo, 2016). The human microbiota is constituted with bacteria, viruses, archaea, fungi, phages, and protists. It is estimated that 500-1000 bacterial species are found in or on the human body, with the majority of bacteria living in the intestines (Sender et al., 2016; Turnbaugh et al., 2007). Intestinal bacteria play critical roles in maintaining host physiology by educating the immune system, preventing colonization and growth of pathogenic bacteria, and maintaining epithelial integrity (Ahern, Faith, & Gordon, 2014; Ashida, Ogawa, Kim, Mimuro, & Sasakawa, 2011; Diaz Heijtz et al., 2011). Bacterial colonization is chaotic during the first few years of life but stabilizes by adulthood (Gilbert et al., 2018). While each individual harbours a unique microbiome, functional redundancy is an intrinsic property of the gut (Moya & Ferrer, 2016). Microbial composition is subject to change in response to a variety of factors including diet, antibiotic usage, co-habitation, exercise, and stress (Gilbert et al., 2018).

Several distinct bacterial habitats occur along the length of the intestines as a result of physiological gradients (e.g., pH, oxygen, and nutrient) and immune activity (e.g., antimicrobial proteins). Moving from the proximal end of the small intestine to the distal end of the large intestine, antimicrobial peptide and oxygen concentrations decrease, whereas pH increases (Donaldson, Lee, & Mazmanian, 2016). Furthermore, the types of nutrients available to bacteria change from simple carbohydrates and fatty acids in the small intestine to complex polysaccharides in the large intestine. Consequently, bacterial density and diversity increase distally along the length of the intestines (Donaldson et al., 2016; Islam et al., 2011). Fast-growing facultative anaerobes from the phyla Firmicutes and Proteobacteria dominate the small intestine whereas fermentative polysaccharide-degrading anaerobes from the phyla Firmicutes and Bacteroidetes dominate the large intestine (Alhagamhmad, Day, Lemberg, & Leach, 2016).

Bacteroides species (phylum Bacteroidetes) have adapted mechanisms to establish stable, long-term, commensal associations with their hosts (Ley et al., 2008). When germ-free mice are colonized by *Bacteroides*, many of the defects associated with an under-developed immune system are corrected (Ivanov et al., 2008). Nevertheless, when intestinal homeostasis is disrupted, *Bacteroides* can become pathogenic. Individuals with CD have an increased risk of developing intra-abdominal abscesses, which tend to be colonized by anaerobic bacteria such as *Bacteroides fragilis* (*B. fragilis*) (Gibson, Onderdonk, Kasper, & Tzianabos, 1998). Around the world, *Bacteroides* species are dominant colonizers of the human gut, making them perfect candidates to study as windows into the microbiota (Wexler & Goodman, 2017).

1.1.2.1: Sequencing Technologies

Advances in sequencing technologies, namely 16S ribosomal ribonucleic acid (rRNA) gene sequencing and metagenomic sequencing, have vastly improved characterization of the human microbiota. The 16S rRNA gene is ubiquitous amongst bacteria and contains hypervariable regions that can be used for identification (Martinez-Porchas, Villalpando-Canchola, Ortiz Suarez, & Vargas-Albores, 2017). Metagenomic sequencing comprehensively samples all genes from all organisms present within a sample. Metagenomic sequencing therefore tends to offer better taxonomic and functional resolution than 16S rRNA gene sequencing (Jovel et al., 2016). Even so, there are some important concepts to remember when interpreting intestinal microbiota metagenomic studies. Firstly, the majority of studies utilize fecal samples which may be convenient to collect, but may not accurately reflect microbial diversity at the mucosal surface (Watt et al., 2016). As such, intestinal biopsy is the gold-standard; however, it is a highly invasive procedure and subject to sample bias unless numerous samples are taken. Secondly, approximately half of the sequencing reads from human fecal samples cannot be mapped to bacterial reference genomes (Qin et al., 2010). Recent work by Zou et al. (2019) has begun to ameliorate this issue, having added 264 previously unpublished bacterial genomes to the Culturable Genome Reference database (Zou et al., 2019). Finally, sequenced genes are not necessarily expressed. Therefore, metagenomic sequencing cannot provide insight into which microbial functions are actually occurring/changing within a population (Tanca et al., 2017). Recently the Human Microbiome Project 2, otherwise known as the Integrative Human Microbiome Project, followed 90 IBD patients for one year collecting taxonomic, metagenomic, metatranscriptomic, metaproteomic, and metabolic data on the intestinal microbiota. This open-source dataset provides

information on microbial presence, predicted and actual microbial function, and microbial metabolic contribution to the host.

1.2: Crohn's disease

CD is a chronic relapsing inflammatory disorder that can involve any part of the digestive tract, predominately the terminal ileum and colon, and presents with transmural, patchy, and asymmetrical inflammation (Peterson & Artis, 2014). Using the weighted Pediatric Crohn's Disease Activity Index (wPCDAI), CD is categorized as inactive, mild, or moderate to severe based on pain, weight, growth, laboratory examination, and general well-being (Silverberg et al., 2005). The main symptoms of CD include diarrhea, abdominal pain, fatigue, weight loss, and nausea. Due to the chronic nature of CD, patients are at risk of developing other localized (i.e., strictures, fistulas, abscesses, or cancers) or extra-intestinal manifestations (i.e., in the joints, skin, or eyes) (Kaplan et al., 2018).

1.2.1: Epidemiology

Prevalence refers to the number of people with a disease whereas incidence refers to number of newly diagnosed cases. CD has a global reach with historically high prevalence rates in Europe and Canada (at 322 and 319 per 100,000 persons, respectively) (Kamm, 2017). Nevertheless, prevalence rates of CD are increasing rapidly in newly industrialized countries in Asia, Africa, and South America (Ng et al., 2017).

CD can be diagnosed at any age; however, adolescents and young adults are most susceptible (Torres et al., 2017). Between 2000 and 2008, pediatric CD prevalence rates increased 3.9% per year in Canada (Benchimol, Kaplan, et al., 2017). Amongst the provinces, Nova Scotia and Quebec have the highest incidence rates of pediatric CD at 9.3 and 8.8 per 100,000 people, respectively (Benchimol, Bernstein, et al., 2017). Within the pediatric

population, boys are more likely to be diagnosed with CD (Kaplan et al., 2018). CD is more common in individuals of Ashkenazi Jewish origin than in non-Jews (Levine et al., 2016).

1.2.2: Etiology

CD has a complex etiology involving the interaction of various environmental, immunological, microbial, and genomic factors. CD is thought to arise in genetically predisposed individuals when gut microbial changes trigger immune activation and chronic inflammation (de Souza, Fiocchi, & Iliopoulos, 2017). In this section, each factor will be reviewed separately.

1.2.2.1: Environmental Factors

The role of environmental factors in CD development are highlighted by Canadian studies showing that first and second generation children of immigrants from countries with low IBD rates are at an increased risk of developing IBD (Benchimol et al., 2015; Carroll et al., 2016). Risk is further increased if children spend their first five years of life in urban environments (Benchimol, Kaplan, et al., 2017). The relationship between urbanization and chronic inflammatory diseases has previously been explored through the hygiene hypothesis. This theory proposes that children who grow up in relatively "clean" environments (i.e., low bacterial contamination) develop abnormal immune responses because their immune systems are not properly educated to microorganisms. Several CD studies indirectly support the hygiene hypothesis as CD is less likely to develop in individuals who had pets during childhood, lived on a farm, had a large family, or drunk unpasteurized milk (Cholapranee & Ananthakrishnan, 2016; Frolkis et al., 2013).

Diet is also associated with CD risk as certain dietary components can modify the intestinal microbiota and epithelium, disrupting homeostasis. Increasing rates of global CD

parallel increasing consumption rates of western diets (high in fats and refined sugar, and low in fiber) (Frank M. Ruemmele, 2016). Numerous studies have attempted to identify risky and protective food factors; however, the results are relatively inconsistent as food frequency questionnaires are susceptible to recall bias (Penagini et al., 2016). Of the food factors examined, dietary fiber intake consistently appears to lower risk of developing IBD (Ananthakrishnan et al., 2013). Alternatively, commonly used food additives, such as maltodextrin, carboxymethyl cellulose, carrageenan, and xanthan gum make the intestinal environment more pervasive to bacterial colonization and increase risk of IBD development (Darfeuille-Michaud et al., 2004). While diet can increase risk for developing CD, certain dietary components have also been harnessed into successful therapeutics for treating active CD (MacLellan et al., 2017). Section 1.2.3.1 will provide more information on the use of exclusive enteral nutrition (EEN) for inducing remission in pediatric CD.

Several other environmental factors, such as vitamin D exposure, oral contraceptives, cigarette smoking, and non-steroidal anti-inflammatory drugs have been studied for their role in CD development; however, due to the heterogenous design of environmental studies, results are inconclusive (Ananthakrishnan et al., 2012; Barclay et al., 2009; Boyko, Theis, Vaughan, & Nicol-Blades, 1994; Cholapranee & Ananthakrishnan, 2016; Lewis & Abreu, 2017; Meyer, Ramzan, Heigh, & Leighton, 2006; Parkes, Whelan, & Lindsay, 2014). In general, risk for CD is increased in individuals who smoke cigarettes, use oral contraceptive or non-steroidal anti-inflammatory drugs, and lack vitamin D.

1.2.2.2: Immunological Factors

CD pathogenesis is a result of an immune imbalance that leads to persistent intestinal mucosal damage and loss of tolerance to commensal bacteria (Figure 2). At the innate level,

neutrophil recruitment to sites of inflammation is increased, macrophage bacterial clearance is defective, and dendritic cell receptor expression and pro-inflammatory cytokine production are increased (de Souza & Fiocchi, 2016; Hart et al., 2005; Segal, 2018). Additionally, pro-inflammatory signaling and adaptive immune cell activation is increased in natural killer (NK) cells, NKT cells, and innate lymphoid cells (ILCs) (de Souza & Fiocchi, 2016; Takayama et al., 2010). At the adaptive level, T helper cells (T_H cells), specifically T_H1 and T_H17, are increased and T regulatory (T_{reg}) cells are decreased (Yamada et al., 2016). Furthermore, CD patients develop antibodies in their serum against commensal intestinal bacteria, such as *Escherichia coli* (*E. coli*), and microbial antigens, such as glycans (Dotan et al., 2006; Mow et al., 2004). To date, studies have shown that the severity of CD is associated with higher titres and variety of antibacterial antibodies (Mow et al., 2004).

Both immune and non-immune cells actively secrete a large number of soluble mediators that play a role in initiating and perpetuating inflammation. The first chemokine described for CD was CXC-motif chemokine ligand (CXCL) 8, but since then several others have been implicated, including CC-motif chemokine ligand (CCL)5, CCL3, CCL4, CCL2, CCL7, CXCL10, and CXCL5 (de Souza & Fiocchi, 2016). Within the gut, CXCL8 is secreted by a variety of cell types and functions to induce neutrophil chemotaxis and degranulation.

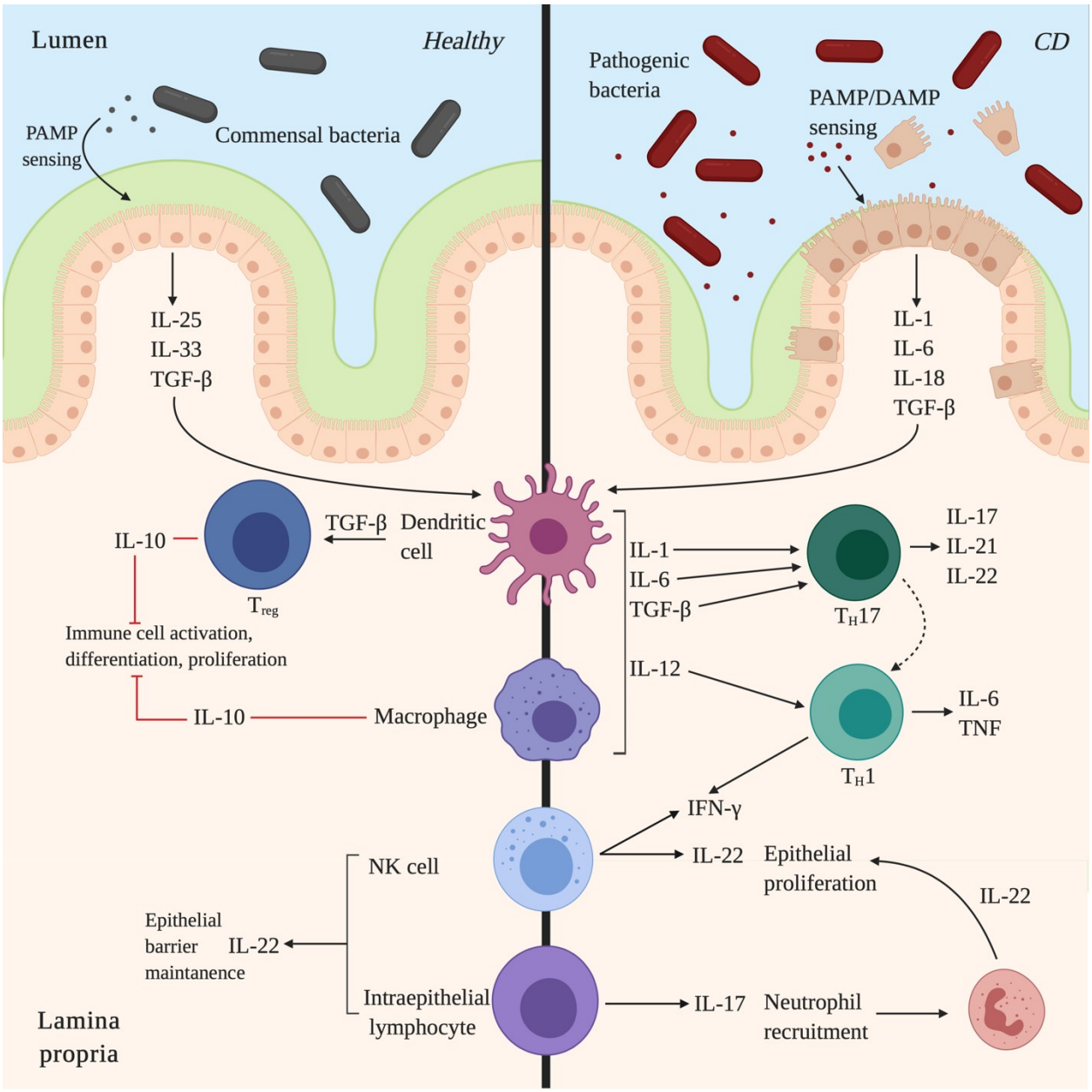


Figure 2. Schematic of immunological differences between healthy and CD mucosa. In healthy individuals, PAMPS are sensed through PRRs on IECs resulting in the production of mediators that activate innate immune cells. Activation of dendritic cells and macrophages reduce further immune cell activity and activation of NK cells and IELs helps to maintain the epithelial barrier. In CD individuals, PRRs respond to PAMPS and DAMPS resulting in the production of different inflammatory mediators by IECs. Innate immune cells activate both T_H17 and T_H1 cells as well as recruit neutrophils to sites of epithelial disturbances. Figure created with BioRender.

1.2.2.3: Microbial Factors

CD patients have different intestinal bacteria than healthy individuals (Sun et al., 2019). Evidence suggests that CD is not caused by a single bacterium but instead by significant changes to overall microbial community (i.e., dysbiosis) (Alhagamhmad et al., 2016). CD-associated dysbiosis shows an overall decrease in bacterial diversity with a general reduction in the phylum Firmicutes and expansion in the phylum Proteobacteria (Joossens et al., 2011). Accompanying the loss of beneficial bacteria is the relative increase in commensal bacteria with immune-stimulating properties (Frank et al., 2011). For example, *Bacteroidetes*, *Enterobacteriaceae*, *Ruminococcus,* and *Pseudomonas* are more frequently observed in CD patients (Alhagamhmad et al., 2016). Moreover, CD patients are at an increased risk of developing *Clostridium difficile* infections due to decreased intestinal microbial diversity and altered immune activity (D'Aoust, Battat, & Bessissow, 2017). Changes to bacterial composition result in changes to bacterial metabolite production. Metabolites, such as short-chain fatty acids (SCFAs), play an important role influencing the host immune system. For example, individuals who consume a western diet are low in SCFA-producing bacteria and this loss in bacteria is accompanied by a reduction in colonic T_{reg} development (Dalal & Chang, 2014).

1.2.2.4: Genetic Factors

A genetic predisposition to CD was first proposed in the 1930s following epidemiological observation that many patients had a family history of CD (Liu & Anderson, 2014). Approximately fifty years later in 1988, the first twin survey was published describing concordance rates of 58.3% and 3.9% for monozygotic and dizygotic twins, respectively (Tysk, Lindberg, Jarnerot, & Floderus-Myrhed, 1988). Subsequent twin studies have continued to support the genetic link to CD, albeit to a lesser degree (Brant, 2011). Alongside twin studies,

family studies have observed that 2-14% of patients have a family history of CD and that first-degree relatives, especially siblings, have a 5% greater risk of developing CD (Freeman, 2002; Halme, 2006). The results from twin and family studies motivated efforts to begin identifying genetic regions responsible for CD.

Linkage studies, a family-based method, identify regions of the genome that underlie disease susceptibility by demonstrating co-segregation of a disease with genetic markers of known chromosomal location (Liu & Anderson, 2014). In 1996, chromosome 16 was the first CD-susceptibility locus identified (named IBD1) and in 2001, *nucleotide-binding oligomerization domain-containing 2* (*NOD2*) was the first relevant CD-susceptibility gene identified (Hugot et al., 2001, 1996). The three most common *NOD2* variants, all loss of function, are R702W, G908R, and L1007fs; however, other *NOD2* variants do occur (Huang et al., 2015; Lesage et al., 2002; Ng et al., 2012). The risk for CD development increases 15-40-fold in *NOD2* homozygotes or compound heterozygotes and 2-4-fold in *NOD2* heterozygotes (Economou, Trikalinos, Loizou, Tsianos, & Ioannidis, 2004; Siegmund & Zeitz, 2011). Among healthy white Europeans, approximately 11%–14% are heterozygous and 0.4%–0.9% homozygous or compound-heterozygous for *NOD2* variants, which suggests that a *NOD2* variant alone is not enough for the initiation of CD (Kennedy et al., 2018). Additional linkage studies were conducted following the identification of *NOD2,* but the results were largely disappointing due to poor replicability.

Genome-wide association studies (GWAS), a hypothesis-free method**,** identify single nucleotide polymorphisms (SNPs) with allele frequencies that differ significantly between healthy and diseased individuals. GWAS have identified over 200 CD risk loci involved in a variety of cellular processes including, but not limited to, autophagy (*NOD2, ATG16L1, IRGM*),

innate mucosal defense (*ITLN1*), epithelial barrier (*NOD2* and *MUC19)*, immune cell recruitment (*CCL11/CCL2/CCL7/CCL8*), and immune tolerance (*IL27, IL23, IL23R*) (Khor, Gardet, & Xavier, 2011; Liu et al., 2015). Besides *IL-10* mutations that are associated with early-onset CD, an increasing number of genetic variants driving very early onset IBD are being identified, with some being amenable to bone marrow transplant to treat the underlying immune defect (Uhlig et al., 2014). Identified variants only account for approximately 14% of the variance explained for CD (J.-S. Chen et al., 2018).

Current assumptions of complex diseases impose restrictions that limit the effectiveness of models used to study CD genetics. Future models of heritability should consider including new assessments that incorporate the other factors—environmental, microbial, and immunological—that contribute to CD development.

1.2.3: Treatment

CD does not have a cure, thus treatments function to first induce and then maintain remission. The three primary goals of CD treatment are controlling intestinal inflammation (to allow for mucosal healing), optimizing quality of life, and minimizing treatment-related toxicities. The three main CD treatments include surgery, medication, and enteral nutrition. As EEN is a first line therapy for pediatric CD, it will be reviewed in greater detail in this section.

1.2.3.1: Exclusive Enteral Nutrition

EEN is a liquid formula diet administered orally or through a nasogastric tube for a period of 8-12 weeks (Ruemmele et al., 2014). There are two types of formulas: polymeric, such as Modulen (Nestlé), and elemental, such as Elemental O28 (Nutricia). Polymeric and elemental refer to whether the protein source is delivered as intact proteins or individual amino acids, respectively. As both types of formulas are effective, polymeric is used more often than

elemental as it is relatively palatable and cheap (Whitten, Rogers, Ooi, & Day, 2012). In comparison to other treatment options, EEN is effective in inducing remission in 60-80% of cases while also addressing CD-associated nutritional deficits, allowing for mucosal healing, and bettering growth outcomes (Borrelli et al., 2006; Dunn et al., 2016; F.M. Ruemmele et al., 2014). Despite recommendation as a first line therapy, EEN use in North America is significantly lower than in Europe, in part because the mechanism of action is poorly understood (Forbes et al., 2017; Lawley et al., 2018; Van Limbergen et al., 2015).

Recent studies suggest that EEN may induce remission by changing the gut microbiota; however, documented changes vary considerably between patients and study cohorts (Gatti et al., 2017; MacLellan et al., 2017). Comparing different studies is complicated by the variability in methodology (e.g., sample site and type of EEN formula) and personal heterogeneity. Nevertheless, patterns do arise. At disease-onset and before EEN, CD patients generally have different compositions of intestinal bacteria than healthy individuals (Gevers et al., 2014). Importantly, the bacterial composition at disease-onset may serve to indicate which patients will respond well to EEN. A study by Dunn et al. (2016) observed that patients who sustained remission began EEN with *Akkermansia muciniphila, B. fragilis*, *Bacteroides ovatus, Lachnospiraceae,* and *Ruminococceae* (Dunn et al., 2016). Once EEN has begun, bacterial diversity tends to decrease while abundance tends to increase (Gerasimidis et al., 2014). Following the end of EEN, bacterial diversity increases (Gerasimidis et al., 2014; Kaakoush, Day, Leach, Lemberg, & Mitchell, 2016). These studies suggest that EEN may induce remission by creating a dysbiotic state that disrupts CD-associated bacteria and allows for recolonization.

EEN also changes bacterial metabolite production. Notably, genes involved in transporting spermidine/putrescine (involved in cell growth) are increased during EEN, which

may indicate epithelial cell growth and healing (Quince et al., 2015). Furthermore, following two-weeks of EEN, CD patients have a significant decrease in toxic microbial metabolites as well as other products such as SCFAs (Walton et al., 2016).

1.3: NOD2: An intracellular bacterial sensor

As stated in section 1.1.1.1, PRRs alert the immune system to foreign invaders by recognizing PAMPs and DAMPs (Amarante-Mendes et al., 2018). As NOD2, a type of NLR, is one of the best studied susceptibility genes for CD, this section will provide an overview of NOD2 structure and function.

1.3.1: NOD2 structural elements

NLR-family members are generally comprised of three domains: an amino (N)-terminal domain that is composed of protein–protein interaction cassettes, such as caspase-recruitment domains (CARDs) or pyrin domains; a central NOD (otherwise known as a NACHT) domain, which facilitates self-oligomerization and has adenosine triphosphate (ATP)ase activity; and a carboxy (C)-terminal leucine rich repeat (LRR) domain that is involved in ligand sensing (J. P.-Y. Ting & Davis, 2005). Variations in the N-terminus are used to categorize NLR into subfamilies, such as NLRs containing CARD (NLRC) and NLRs containing Pyrin (Biswas & Kobayashi, 2013). NOD2, encoded by *CARD15*, has an N-terminal domain with two CARDs (therefore NOD2 is part of the NLRC subfamily), a central NOD, and a C-terminal LRR (Strober, Murray, Kitani, & Watanabe, 2006). NOD2 is a 110 kDa protein with 1040 amino acids (Sidiq, Yoshihama, Downs, & Kobayashi, 2016). The three loss of function *NOD2* variants mentioned in section 1.2.2.4—R702W (missense), G908R (missense), and L1007fs (frameshift)—all occur within the LRR (Sidiq et al., 2016). For the rest of the thesis, R702W, G908R, and L1007fs are referred to as V8, V12, and V13, respectively.

1.3.2: NOD2 expression

Within the intestine, NOD2 is expressed intracellularly by a variety of cells including both hematopoietic (T cells, B cells, macrophages, dendritic cells, and mast cells) and non-hematopoietic cells (Paneth cells, stem cells, goblet cells, and enterocytes) (Al Nabhani, Dietrich, Hugot, & Barreau, 2017; Ferrand et al., 2019). NOD2 expression by enteroendocrine cells, tuft cells, and M-cells remains unconfirmed (Ferrand et al., 2019). NOD2 localizes on the basolateral side of the epithelial cell, but may re-localize to the apical membrane (Barnich, Aguirre, Reinecker, Xavier, & Podolsky, 2005; Kabi & McDonald, 2012).

1.3.3: Ligand recognition

NOD2 binds directly to muramyl dipeptide (MDP), a component of the Gram-negative and -positive peptidoglycan layer, through the LRR domain (Grimes, Ariyananda, Melnyk, & O'Shea, 2012; Mo et al., 2012). NOD2 does not respond to the inactive chiral isomer of MDP (MDP_{DD} or MDP_{LL}) (Traub, von Aulock, Hartung, & Hermann, 2006). The mechanism by which MDP enters the cell remains unclear; however, five general mechanisms are being explored: phagosomes, outer membrane vesicles, transmembrane channels, endocytosis, and junctions (between neighbouring cells) (Caruso, Warner, Inohara, & Núñez, 2014; Kaparakis-Liaskos & Ferrero, 2015; Marina-Garcia et al., 2009; Nakamura et al., 2014; Vavricka et al., 2004). CD-associated NOD2 variants have disrupted MDP sensing (Mirkov, Verstockt, & Cleynen, 2017).

1.3.4: NOD2 signaling

Basally, NOD2 is inactive as the LRRs fold back onto the NOD domain hindering self-oligomerization (Tigno-Aranjuez & Abbott, 2012). Upon exposure to MDP, NOD2 oligomerizes, recruiting and activating receptor-interacting serine/threonine-protein kinase 2 (RIPK2, also called RIP2 or RICK), which selectively promotes signaling through the nuclear

factor kappa-light-chain-enhancer of activated B cells (NF-κB) or mitogen-activated protein kinase (MAPK) pathways (Figure 3). The NF-κB pathway involves the recruitment of IKKγ, a component of the inhibitor of κB kinase (IKK) complex, and the phosphorylation of IKKβ. Phosphorylated IKKβ then phosphorylates IκB, which leads to proteasomal degradation of IκB and release of the p50 and p65 subunits of NF-κB. The heterodimer p50-p65 initiates the transcription of NF-κB-dependent pro-inflammatory genes, including cytokines and chemokines (Feerick & McKernan, 2017). Alternatively, the MAPK pathway involves phosphorylation of c-Jun N-terminal kinases (JNK) and p38, and subsequent induction of activator protein-1-dependent (AP-1) gene transcription. Signaling through NF-κB and MAPK is influenced by RIPK2 ubiquitination and binding partners (Boyle, Parkhouse, & Monie, 2014; Clark, Marinis, Cobb, & Abbott, 2008).

NOD2 activation can also induce autophagy, a highly conserved catabolic process that delivers cytoplasmic cargo (e.g., misfolded proteins, damaged organelles and invading microbes) to lysosomes for controlled degradation (Yu, Chen, & Tooze, 2018). NOD2 recruits and co-localizes with autophagy-related 16 like 1 (ATG16L1) at sites of bacterial entry (Travassos, Carneiro, Girardin, & Philpott, 2010). Human epithelial, macrophage, and dendritic cells that express *NOD2* variants fail to recruit ATG16L1 and to initiate autophagosome formation at the site of bacterial entry, which significantly impairs bacterial killing (Homer, Richmond, Rebert, Achkar, & McDonald, 2010; Travassos et al., 2010).

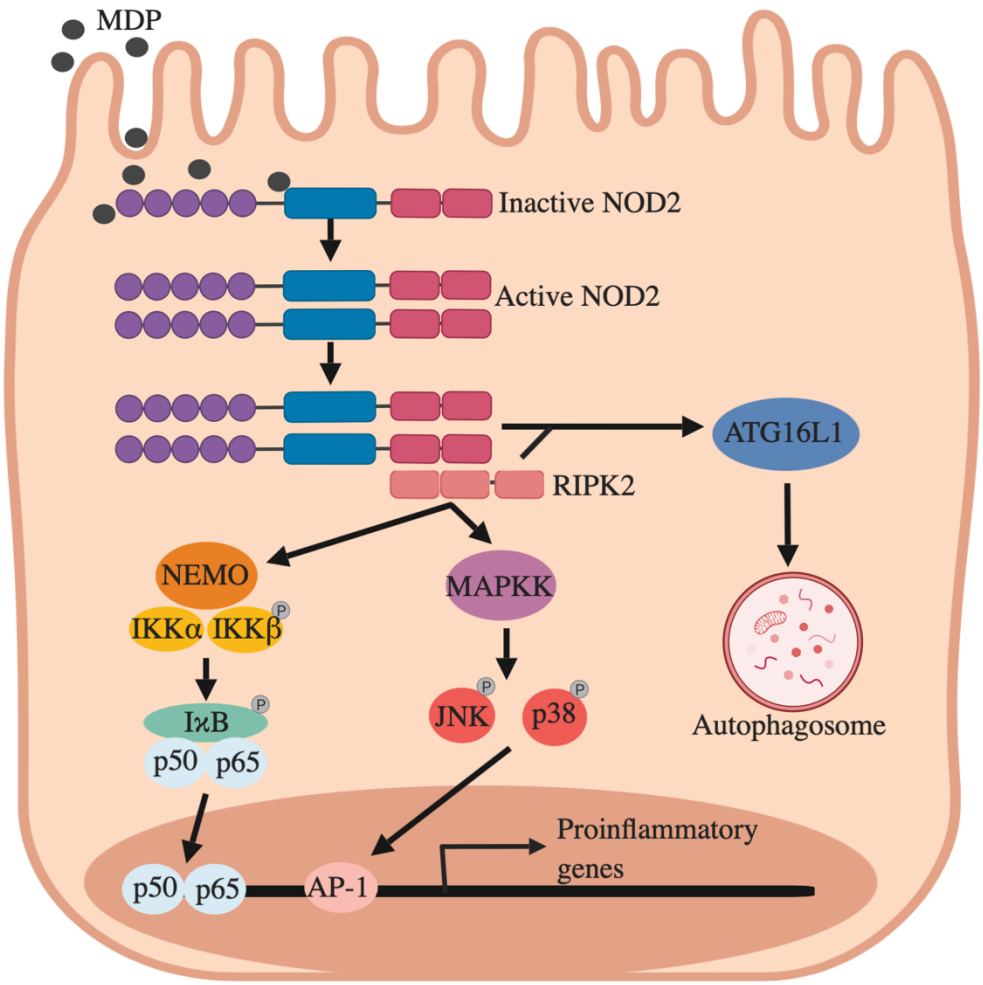


Figure 3. Cellular pathways downstream of NOD2 activation. Upon activation by MDP, NOD2 oligomerizes, recruiting and activating RIPK2. RIPK2 activation can lead to the activation of either NF-κB, MAPK, or autophagy. Signaling through NF-κB and MAPK results in pro-inflammatory gene expression. Autophagy can also be activated in a RIPK2-independent manner. Figure created with BioRender.

1.3.5: NOD2 regulation

Regulation of NOD2 is important for maintaining cellular homeostasis and preventing inappropriate immune responses to stimuli. NOD2 signaling is tightly regulated by a collection of protein interactions, post-transcriptional modifications, and post-translational modifications. Protein interactions pertinent to this thesis are described in section 1.4.3.1. MicroRNAs (miRNAs) are small non-coding RNAs that regulate gene expression post-transcriptionally (O'Brien, Hayder, Zayed, & Peng, 2018; Wu et al., 2010). Several miRNAs are thought to play important roles in NOD2 regulation (Negroni, Pierdomenico, Cucchiara, & Stronati, 2018). For example, miRNA-320, which can bind to three separate sites (called a, b, and c) in the 3′-untranslated region of NOD2, is inversely correlated with NOD2 expression during inflammation (Pierdomenico et al., 2016). As a result of recent studies describing miRNA activity and CD, new research is exploring the therapeutic potential of miRNA inhibitors that specifically target CD-activated miRNAs (Soroosh, Koutsioumpa, Pothoulakis, & Iliopoulos, 2018). Post-translationally, NOD2 signaling is regulated by ubiquitination, leading to proteasomal degradation following MDP stimulation, and *O*-GlcNAcylation, leading to a prolonged half-life (Baker et al., 2017). Additionally, NOD2 is subject to epigenetic modification through methylation of CpG sites in the NOD2 promoter (Acevedo et al., 2015; Karatzas, Gazouli, Safioleas, & Mantzaris, 2014; Nimmo et al., 2012). The addition of methyl groups leads to gene inactivation.

1.3.6: Role in Innate and Adaptive Immunity

NOD2-activated innate and adaptive immunity creates a tolerogenic gut environment that can also function to launch attacks against invading pathogens. Commensal bacteria and NOD2 work within a feedback loop, whereby commensal bacteria promote NOD2 expression, which in

turn prevents bacterial overexpansion (Feerick & McKernan, 2017). One way NOD2 controls bacterial density is by inducing the production of antimicrobial peptides, such as defensins, by Paneth cells (de Bruyn & Vermeire, 2017; Petnicki-Ocwieja et al., 2009; Tan, Zeng, & Zhi, 2015). CD patients with *NOD2* variants have reduced defensin production from Paneth cells, as well as abnormal mucin secretion from goblet cells and increased bacterial internalization by enterocytes (Petnicki-Ocwieja et al., 2009; Ramanan, Tang, Bowcutt, Loke, & Cadwell, 2014; Saxena, Lopes, Poon, & McKay, 2017; Wehkamp et al., 2005). The NOD2-activated innate response is further characterized by expression of various pro-inflammatory immune factors, including interleukin (IL)-1β, tumor necrosis factor (TNF), IL-6, CCL2, CXCL8, and CXCL2 (Philpott, Sorbara, Robertson, Croitoru, & Girardin, 2014). Expression of chemokines recruits and primes innate immune cells, such as neutrophils and monocytes, to sites of infection (Ajendra et al., 2016; Kim et al., 2011). The NOD2-adaptive response is characterized by a T_H2-type polarization profile, although co-stimulation with TLR agonists primes T_H1 and T_H17 responses (Magalhaes et al., 2008).

NOD2 signaling and synergistic activity with TLRs protects against infections caused by various bacteria, such as *Yersinia pseudotuberculosis*, *Listeria monocytogenes, Citrobacter rodentium, Bacillus anthracis, Salmonella enterica, Streptococcus pneumoniae,* and adherent-invasive *E. coli* (Al Nabhani et al., 2017; Caruso et al., 2014; Kim et al., 2008). Loss of *NOD2* increases susceptibility to infection; however, as PAMPS are recognized by multiple PRR, *NOD2* deficiency only modestly affects bacterial clearance.

1.4: Heat Shock Protein (HSPs)

HSPs are evolutionarily conserved and are essential mediators of cellular homeostasis. As molecular chaperones, HSPs help fold newly synthesized polypeptides, assemble multiprotein

complexes, prevent abnormal protein structure, maintain protein conformation through refolding, mediate intracellular protein trafficking, and degrade damaged proteins via the ubiquitin-proteasome pathway (Binder, 2014; Broere, van der Zee, & van Eden, 2011; Li, Menoret, & Srivastava, 2002). In addition to intracellular functions, HSPs can act extra-cellularly to activate immune cells through antigen presentation (Pockley & Henderson, 2018).

1.4.1: HSP90 Family

HSPs are categorized into six families based on molecular weight: small HSPs, HSP40, HSP60, HSP70, HSP90, and large HSPs. Representative proteins from each HSP family can be found in both prokaryotes and eukaryotes. The HSP90 family interacts with numerous functionally and structurally diverse client proteins and thus remains a popular area of research.

Members of the HSP90 family consist of three conserved domains: a N-terminal ATP-binding site, a middle domain for activating N-terminus ATP hydrolysis and substrate binding, and a C-terminal domain for dimerization (Pearl & Prodromou, 2006). Amongst eukaryotes, there are four homologs found in different cellular compartments: cytosolic heat-shock inducible HSP90α, cytosolic constitutively-expressed HSP90β, mitochondrial TNF receptor associated protein 1 (Trap1), and endoplasmic glucose-related protein 94 (grp94) (Johnson, 2012). These four homologs are thought to have arisen from duplication events of high temperature protein G (HtpG), a bacterial HSP (Chen, Zhong, & Monteiro, 2006).

Three HtpG lineages evolved (groups A, B, and C) following two gene duplication events. Most bacteria express at least one copy of *HtpG*; however, copy number is variable (Table 1). The number of copies of *HtpG*, which are never part of the same group, does not strongly correlate with membership in a particular bacterial phylum. Each HtpG group has a signature sequence, however, the protein is so divergent that there is not one common signature

sequence for all three groups (Chen et al., 2006). The archaeaon, *Methanosarcina mazei* (*M. mazei*), has an HSP90-like gene that is used to root gene trees comparing different HtpG amino acid sequences. Unlike HSP90, deleting *HtpG* is not lethal to bacterial cells; however, growth and biofilm formation is impaired at high temperatures as is biosynthesis and secretion of certain enzymes as observed in *E. coli* (Grudniak, Pawlak, Bartosik, & Wolska, 2013).

Table 1. Number of *HtpG* copies found in common gut bacterial genera. Copy number can range from 0-3, with many genera containing at least one copy.

Bacterial Genus	**Number of *HtpG* copies**
Bacteroides	1-2
Parabacteroides	1
Enterobacter	1
Clostridium	0-3
Ruminococcus	1
Escherichia	1-2

HtpG has both protective and pathogenic activity. For some bacteria, such as *Salmonella*, *Leptospira*, *Edwardsiella*, *Porphyromonas* and *Francisella*, HtpG functions as a virulence factor to aid in pathogenesis and persistent infection. For example, healthy individuals have higher serum antibody levels against *Porphyromonas gingivalis* (*P. gingivalis*) HtpG than patients with either chronic or aggressive periodontitis (Shelburne et al., 2008). HtpG activity, however, is not always associated with disease. In a metagenomic analysis of fecal samples collected from healthy and CD children, HtpG abundance was highest in the healthy population (Dunn et al., 2016). Furthermore, amongst the children with CD, those able to sustain remission following EEN had a greater abundance of HtpG than those children who did not sustain remission (Dunn et al., 2016). The results from this study suggest that HtpG may have some role in maintaining intestinal homeostasis and regulating inflammation.

1.4.2: Factors that Influence Endogenous HSP Abundance

In addition to constitutive expression under physiological conditions, endogenous HSPs can be induced by four types of stimuli: (1) physical, such as heat shock and radiation; (2) chemical; (3) microbial; and (4) dietary. This section will focus on microbial and dietary factors that influence host HSP expression.

HSP expression within the intestines is influenced by interactions with molecules derived from intestinal bacteria. For example, lipopolysaccharide (LPS) from *E. coli* induces HSP25 and a sporulating factor from *Bacillus subtilis* induces HSP27 (Kojima et al., 2004; Okamoto et al., 2012). As the colon has greater bacterial diversity and density than the small intestine, the colon tends to have higher expression levels of HSPs. Accordingly, when bacteria are absent in germ-free mice, colonic HSP expression is either reduced or completely lost (Hu et al., 2010). It is not surprising that antibiotics, which often accompany CD treatments and are known to change microbial composition, also change HSP expression. Mice treated with broad-spectrum antibiotics such as metronidazole exhibit reduced intestinal expression of HSP25 and HSP72 (Rakoff-Nahoum, Paglino, Eslami-Varzaneh, Edberg, & Medzhitov, 2004). This finding suggests that antibiotic therapy may disrupt HSP expression within the intestines, therefore increasing susceptibility to infection and inflammation.

Nutrients (or lack of) also alter HSP expression. For example, in fasting piglets, HSP27 and HSP90 expression is increased, while HSP70 expression does not change (Lallès & David, 2011). That is not to say that HSP70 does not play a role in dietary-mediated cellular protection. Glutamine, the most abundant free amino acid in the body, induces HSP70 expression during times of cellular stress (David et al., 2014). Several other amino acids also impact HSP expression, such as arginine, which restores physiological levels of HSP70 in intestinal Caco-2

cells; and threonine, which induces HSP25 expression in rat small intestinal IEC-18 cells (Baird, Niederlechner, Beck, Kallweit, & Wischmeyer, 2013; Lenaerts et al., 2007). In addition to free amino acids, proteins can increase HSP expression (de Moura, Lollo, Morato, Carneiro, & Amaya-Farfan, 2013). Remember that the first line therapy in pediatric CD is EEN, which is a liquid formula composed of either individual amino acids or intact proteins. As previously described, EEN is thought to induce remission by modifying the intestinal bacterial microbiota. Seeing that gut bacteria and their products are important inducers of HSPs during stress, perhaps EEN is not only affecting the gut microbiota but also the gut expression of HSPs (Finlayson-Trick, Connors, Stadnyk, & Van Limbergen, 2018).

1.4.3: HSPs and CD

As HSPs play an important role in mediating stress and inflammation, it is not unexpected that HSP polymorphisms or altered HSP abundances lead to disease. In general, HSP abundance appears to be increased in CD patients. Furthermore, autoantibodies against HSPs are found in CD patients. There are two theories as to why these antibodies appear: firstly, the body may develop antibodies against bacterial HSPs, which then cross-react with endogenous HSPs leading to inflammation; and secondly, antibodies may form against endogenous HSPs released into the extracellular space following injury to the epithelium as DAMPS (Maunder, 2000).

Specific HSP polymorphisms, such as those identified for HSP70, are associated with CD risk (Nam et al., 2007). HSP70 polymorphisms disrupt the epithelial barrier and enable bacterial infection (Nam et al., 2007; Takahashi et al., 2017). Besides HSP70, the role of other HSP polymorphisms in CD development is poorly understood. Furthermore, HSPs often work in conjunction with one another, such as HSP70 and HSP90, therefore future studies need to investigate how polymorphisms impact the combined activities of HSPs.

1.4.3.1: HSP Interactions with NOD2

HSP90 is constitutively associated with NOD2 until MDP binds to NOD2 and the complex dissociates (Lee, Biswas, Liu, & Kobayashi, 2012). HSP90 requires both NOD2 CARD motifs, as the study by Lee et al. (2012) showed that HSP90 fails to interact with NOD2 if one or more of the motifs is deleted. Whether HSP90 can bind to NOD2 variants remains unclear. NOD2 activity is further regulated by interacting with the substrate binding domain of HSP70. Grimes et al. (2014) demonstrated that the half-life of NOD2 could be increased or decreased depending on HSP70 overexpression or down-regulation, respectively. In addition to wild-type (WT) NOD2, HSP70 stabilizes the three most common NOD2 variants and restores proper responsiveness to MDP (Mohanan & Grimes, 2014).

1.5: Thesis Overview

The Dunn et al. (2016) study informed my decision to examine the role of HtpG on innate immunity. In the Dunn et al. (2016) study, metagenomic sequencing was conducted on fecal samples collected at various time points from pediatric CD patients on EEN and their healthy siblings. HUMAaN was used to process the sequences and to predict Kyoto Encyclopedia of Genes and Genomes (KEGG) orthologs, pathways, and modules. Between the healthy siblings and the CD patients, there were eight pathways with community-level differences and low false positive rates. These pathways were organized into three groups: (1) pathways with no prior connection and/or mechanistic connection to IBD, (2) pathways connected to IBD, and (3) pathways with roles in innate immunity. Pathways from group 1 included nitrotoluene degradation, polycyclic aromatic hydrocarbon degradation, glyocylate and dicarboxylate metabolism, and fatty acid metabolism. Group 2 comprised of ascorbate, aldarate, and sphingolipid metabolism pathways. Finally, group 3 consisted of the butanoate metabolism

pathway and the NOD-like receptor signaling (NLSR) pathway. Genes from the group 3 pathways were the most intriguing for further investigation as they already had a plausible mechanistic connection to IBD. CD patients with non-sustained remission had a higher mean relative abundance of the butanoate metabolism pathway and a lower mean relative abundance of the NLSR pathway than healthy controls. I decided to examine the NLSR pathway in greater detail because while NLSR is not a bacterial pathway, bacteria encode homologs to NLSR genes, such as HtpG.

The research in this thesis consequently focuses on the characterization of the bacterial heat shock protein, HtpG, in the context of pediatric CD. Firstly, I sought to describe how HtpG changes in response to EEN, specifically observing HtpG abundance, bacterial source, and group identity. My second objective was to describe the extracellular (anti-)inflammatory activity of HtpG in an intestinal cell culture model. Finally, as HtpG and HSP90 are homologs and as HSP90 interacts with NOD2 intra-cellularly, I examined whether HtpG and NOD2 WT/V8/V12/V13 interacted *in vitro.* I propose that bacterial HtpG abundance changes with the pediatric CD patient's disease and is capable of acting directly on intestinal epithelial cells in an anti-inflammatory capacity.

CHAPTER 2 METHODS

2.1: HtpG Dry-lab Analysis

Dry-lab analysis used relative abundance from shotgun metagenomic sequencing data previously collected and analyzed for the Dunn et al. (2016) study. Fourteen patients (CD1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, and 15) were analyzed at timepoints ranging from baseline to 96-weeks. As described in the study, disease activity was assessed using wPCDAI, with a score >12.5 indicating non-sustained remission and a score <12.5 indicating remission. All dry-lab work was accomplished with assistance from Dr. Katherine Dunn (Department of Biology, Dalhousie University).

2.1.1: HtpG Gene Tree

Fragments of *HtpG* sequences were identified from the metagenomic dataset for each patient at each timepoint. Fecal, skin, and saliva samples collected from healthy individuals were identified from the Human Microbiome Project online database (https://hmpdacc.org/hmp/). *HtpG* query fragments were translated in six reading frames, producing six-protein sequences that were compared to a protein sequence database using Basic Local Alignment Search Tool (blastx, translated nucleotide to protein). Protein targets with e-values <0.001 provided confidence in protein identity and were selected for downstream analysis. Prior to alignment, target sequences were culled such that bacterial species were only represented once in the dataset (i.e., generated gene trees do not reflect density of sequences). Protein target sequences were then aligned using Multiple Alignment using Fast Fourier Transform (MAFFT) online. Exported "Phylip (full name, padded & interleaved)" files were uploaded for Randomized Axelerated Maximum Likelihood (RAxML) analysis, which was initially run with both CAT and GAMMA models using the LG exchangeability matrix and frequencies estimated from the data. Both

models generated similar results, therefore subsequent analysis was completed with just the Gamma model. Bootstrap support was estimated from 100 bootstrap replicates and nodes with support greater than 70% were indicated as an asterisk on the tree. Trees were visualized using FigTree program.

2.2: Growth of Bacterial Cultures

Bacterial strains used in this study are listed in Table 2. *E. coli* was routinely grown at 37°C overnight either shaking in liquid, or on solid lysogeny broth (LB) medium (10 mg/mL tryptone, 5 mg/mL yeast extract, 10 mg/mL sodium chloride [NaCl], with/without 20 mg/mL agar). The following antibiotics (Sigma) were added when appropriate: 100 µg/mL ampicillin (Amp) and 50 µg/mL kanamycin (Kan). Frozen stocks (0.5 mL overnight culture, 0.25mL 80% glycerol) were kept at -80°C.

Table 2. Bacterial strains used in this study

Strains	Description	Genotype	Source
5-alpha Competent *E. coli* (High Efficiency)	*E. coli* strain used for cloning	*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	New England Biolabs, CAT# C2987H
BL21(DE3) Competent *E. coli*	*E. coli* strain used for protein overexpression	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdSλ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	New England Biolabs, CAT# C2530H

2.3: Cloning *HtpG* from *B. fragilis* Genomic Deoxyribonucleic Acid (DNA)

B. fragilis was selected based on the observation in the Dunn et al. (2016) study that *Bacteroides* sp. produce the most HtpG. *B. fragilis* (ATCC 25285) was provided by Dr. Jason Leblanc (Halifax, Nova Scotia). To generate an anaerobic environment, thioglycolate broth [acumedia, CAT#7160A] was placed into evacuation jars with a palladium catalyst and air was

removed using the automated Anoxymat system [Mart Microbiology, Drachten, Netherlands]. Genomic deoxyribonucleic acid **(**DNA) was isolated using the Wizard® Genomic DNA Purification Kit [Promega, CAT# A1120] as described by the manufacturer.

Group B *B. fragilis HtpG* was amplified from genomic DNA using the HotStarTaq® *Plus* Master Mix Kit [Qiagen, CAT# 203643]. The manufacturer's instructions were followed with minor modification. Instead of maintaining the annealing temperature for all 30 cycles, a touchdown polymerase chain reaction (PCR) was conducted such that the annealing temperature started 10°C above the final temperature and decreased 10°C over the first ten cycles (1°C per cycle). Primers used for amplification were ordered from Integrated DNA Technologies (IDT) and are outlined in Table 3.

Table 3. **Primers used to amplify group B *B. fragilis* HtpG from genomic DNA.** Restriction enzyme cleavage site are underlined.

Name	**Primer Sequence (5′→ 3′)**	**Restriction Site**
HtpG-Forward	CAAGGATCCATGCAAAAAGGTAA	BamHI
HtpG-Reverse	GGCAAGCTTTTAAATCAGCTCAATG	HindIII

Endonuclease restriction digestions were prepared in a 50 µL reaction volume with the following reagents: 1 µL BamHI, 1 µL HindIII, 1 µg of DNA, 5 µL 10X NEBuffer, and H_2O. Digests were incubated at 37°C for 1 hour. Digested DNA was visualized using FroggaBio Novel Juice [CAT #LD001-1000] on a 1% agarose gel (50 mL 1x ethylenediaminetetraacetic acid [EDTA], 0.5 g agarose). NEB 1 kb DNA ladder [CAT# N3232L] was used as a DNA ladder. DNA bands located at the correct molecular weights were excised from the gel using a razor blade. Excised DNA was purified using the Wizard® SV Gel and PCR Clean-Up System [CAT #A9281] as per the manufacturer's instructions. All ligations used T4 DNA ligase [NEB,

CAT #M0202S] and were prepared as outlined by the manufacturer with a 3:1 ratio of insert to vector (pET28a). Ligated plasmids were used to transform *E. coli* DH5α via heat-shock. Approximately 100 ng of plasmid DNA was added to thawed competent cells (100 µL) and incubated on ice for 30 minutes. The cells were then incubated for 1 minute at 42°C, before returning to the ice for 1-2 minutes. Next, LB medium (500 µL) was added to the cells and incubated at 37°C shaking for 60 minutes. Cells (50 µL) were plated on selective LB solid medium and incubated at 37°C overnight. The construct, pET28a-HtpG, was confirmed via sequencing.

2.4: Cloning *B. Fragilis HtpG* using Synthetic gBlocks

Group B *B. fragilis* HtpG nucleotide sequence (KEGG database BF2409) and group C *B. fragilis* HtpG nucleotide sequence (KEGG database BF0237) were synthesized (IDT) to include a BamHI cut site at the N-terminus and a 6x-His tag and XhoI cut site at the C-terminus (Appendix B). The digestion, ligation, and transformation were conducted as previously described in section 2.3, except for the change in vector (pCR3.1) and endonucleases. Cloning was confirmed via sequencing.

2.5: Cobalt Bead Affinity Purification of Group B *B. fragilis* HtpG from *E. coli*

pCR3.1-HtpG (B) was used to transform BL21 *E. coli* as previously described. Following the overnight incubation, a single colony was used to inoculate 5 mL of LB medium. Cultures were incubated at 37°C shaking for approximately 7 hours, and then diluted 1:100 into 500 mL of LB medium. Isopropyl β-D-1-thiogalactopyranoside (IPTG, 1 mM final concentration) and Amp (100 µg/mL) were added to the cultures, which were then returned to the 30°C shaking incubator to grow overnight. Next, the cultures were pelleted at 4000 xg for 25 minutes at 4°C. Pellets were resuspended in 4 mL wash buffer (20 mM disodium phosphate [Na_2PO_4], 500 mM

NaCl, pH 8.0) and sonicated 6x 30 seconds on ice using a probe sonicator. Cells were then centrifuged at 8000 xg for 30 minutes during which point the Poly-Prep® Chromatography Columns [BIORAD, CAT# 7311550] were prepared. To the column, 300 µL of HisPur™ Cobalt Resin [ThermoFisher Scientific, CAT #89964] was added and washed with 900 µL of ddH_2O followed by 900 µL of wash buffer. Once the centrifugation was complete, supernatant was added to the column and collected as the "flow through" fraction. The pellet was resuspended in 4 mL of wash buffer and collected as the "pellet" fraction. The column was then washed three times with 10 mL of wash buffer per round. Sample "wash" fractions were collected for all three washes. Protein was eluted from the column via three washes with 250 µL of elution buffer per wash (20 mM Na_2PO_4, 500 mM NaCl, 150 mM Imidazole, pH 8.0). Protein purification was visualized via SDS-PAGE gel electrophoresis and Coomassie Blue stain (described in section 2.6).

2.6: SDS-PAGE & Immunoblotting

Sodium dodecyl sulfate (SDS)-polyacrylamide gels (10%) were made and cast using the BIORAD TGX™ FastCast Acrylamide Kit [CAT# 1610172] as per the manufacturer's instruction. Protein samples were quantified using the BIORAD Quick Start™ Bradford Protein Assay [CAT# 5000201]. In preparation for loading, protein (approximately 10-20 µg/lane) was combined with BIORAD 4x Laemmli sample buffer [CAT# 161-0737] already mixed with 2-mercaptoethanol and incubated at 95°C for 5 minutes. Froggabio BLUeye Prestained Protein ladder [CAT# PM007-055] was used as a protein ladder. Gels were run in 1x running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3) for approximately 30 minutes at 250 V or as long as it took the loading front to reach the end of the gel.

To analyze protein loading, gels were first washed three times with ddH_2O for 5 minutes and then covered with Coomassie Brilliant Blue [Biomatik, CAT# A2329-5G]. Gels were microwaved for 10 seconds and stained with agitation at room temperature for 20-30 minutes. Gels were then de-stained using ddH_2O for 2-3 hours at room temperature. Gels were imaged using a BIORAD ChemiDoc Touch Imaging System [CAT# 17001401].

Gels intended for immunoblot analysis were not stained with Coomassie Brilliant Blue. Instead, proteins were transferred from gels to polyvinylidene difluoride (PVDF) membranes via a BIORAD Trans-Blot® Turbo™ Transfer System [CAT# 1704150] with Bio-Rad Trans-Blot® Turbo™ Transfer Buffer [CAT# 100026938]. Proteins were transferred at a constant voltage of 25 V for 7 minutes. Membranes were then blocked with 5% bovine serum albumin (BSA) in Tris-buffered saline with Tween-20 (TBST, 145 mM NaCl, 5 mM Tris-hydrochloric acid (HCl), pH 7.5, 0.1% Tween-20) for 1 hour at room temperature on a shaking incubator. Primary antibody was then diluted to the concentrations described in Table 4 in the same 5% BSA blocking buffer. Membranes were incubated in primary antibody overnight at 4°C on a shaking platform.

Table 4. Antibody sources and dilutions.

Antibody	Dilution	Source
NOD2 (2D9) HRP	1:1000	Santa Cruz Biotechnology, CAT# sc-56168
His Tag Antibody (AD1.1.10)[HRP]	1:2000	Novusbio, CAT# NB100-63173
DYKDDDDK Epitope Tag Antibody (29E4.G7)[HRP]	1:5000	Novusbio, CAT# NBP1-97393

Next, membranes were washed 6x 5 minutes with TBST following which the diluted secondary antibody was applied to the membrane and incubated (shaking) for 1 hour at room

temperature. Membranes were then washed 6x 5 minutes with TBST, developed using the BIORAD Clarity Max™ Western ECL Blotting Substrate [CAT# 170-5060], and imaged using a BIORAD ChemiDoc Touch Imaging System.

2.7: Cell Culture & Maintenance

HEK293T and HT29 cells were maintained in a 5% CO_2 incubator at 37°C. Cells were passaged every 3-5 days (80-90% confluency) by removing the culture medium and washing with 5 mL 1X phosphate buffered saline (PBS). Trypsin (0.05%) [ThermoFisher, CAT#90057] with 0.5 M EDTA [Invitrogen CAT#25300-054] was used to lift adherent cells with a contact time of 5 minutes at 37°C. Cells were then diluted in 10 mL of fresh media (approximately 1:10-1:50 dilution). Cell were then incubated as previously described. For HEK293T cells, Dulbecco's Modified Eagle Medium (DMEM) was supplemented with 10% Fetal Bovine Serum (FBS), and 50 U/mL penicillin-streptomycin. For HT29 cells, Minimum Essential Media (MEM) was supplemented with 10% FBS, 2 mM L-glutamine, 1 nM sodium pyruvate, and 50 U/mL penicillin-streptomycin.

2.7.1: FuGENE Transfection

The following mammalian-expression plasmids were used to transfect HEK293T cells: His-tagged HtpG (pCR3.1-HtpG [B]), Flag-tagged NOD2WT (pCMV-NOD2WT—constructed by Dr. Johan Van Limbergen), Flag-tagged NOD2 V8 (pCMV-NOD2V8—constructed by Dr. Johan Van Limbergen), Flag-tagged NOD2 V12 (pCMV-NOD2V12—constructed by Dr. Johan Van Limbergen), Flag-tagged NOD2V13 (pCMV-NOD2V13—constructed by Dr. Johan Van Limbergen), and a green fluorescent protein (GFP) control (pcDNA3.1-eGFP—provided by Dr. John Rohde).

Cells were seeded at a density of 1.8 x 10^6 cells per 10 cm culture plastic dish. After 24 hours, Promega FuGENE® HD Transfection Reagent [CAT# E2311] was used to transfect cells as per the manufacturer's instructions with minor modification. For a 10 cm dish, 6 µg of DNA was used with 18 µl of FuGENE® Reagent. For a 15 cm dish, 18 µg of DNA was used with 54 µl of FuGENE® Reagent. Transfection efficiency was verified with a GFP control 24 hours post-transfection.

2.7.2: Cell Lysis using RIPA Buffer

Twenty-four hours post-transfection, media was removed from the dish, cells were washed twice with cold 1X PBS, and collected using a cell scraper into chilled 1.5 mL microcentrifuge tubes. Cells were subsequently centrifuged at 500 xg for 5 minutes at 4°C. Supernatant was removed and the pellet was resuspended in 1 mL of cold 1X PBS. Cells were centrifuged at 2500 xg for 5 minutes at 4°C. The supernatant was removed and the pellet was resuspended in 40 µL of lysis buffer (radioimmunoprecipitation assay [RIPA] buffer [50mM Tris-HCl, 150mM NaCl, 50mM Na_2HPO_4, 5mM EDTA, 5mM ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid (EGTA), 0.25% w/v sodium deoxycholate, 0.1% Nonidet p-40, and pH adjusted to 7.5 with 1N HCl]; 10x Protease Inhibitor [Sigma, CAT# P2714]; and 1 mM phenylmethylsulfonyl fluoride). Cells were incubated on ice for 15 minutes. Cells were then centrifuged at 14000 xg for 10 minutes at 4°C. Supernatant was collected and stored at -20°C. Cell lysate was used downstream for affinity purifications using cobalt beads.

2.7.3: Cell Lysis using TRITON X-100 Lysis Buffer

Twenty-four hours post-transfection, media was removed from the dish and cells were washed twice with cold 1X Tris-buffered saline (TBS, 50 mM Tris HCl, with 150 mM NaCl, pH 7.4). Care was taken not to dislodge the cells. To each plate, 1 mL of lysis buffer (50 mM Tris HCl, pH 7.4, with 150 mM NaCl, 1 mM EDTA, and 1% TRITON X-100) was added. The plates

were then incubated at 4°C on a shaker for 30 minutes. Next, cells were scraped and collected in chilled microcentrifuge tubes. Cells were then centrifuged at 4°C for 10 minutes at 12,000 xg. Supernatant was collected in a chilled microcentrifuge tube and stored at -20°C. Cell lysate was used downstream for co-immunoprecipitation using anti-FLAG antibodies.

2.8: Co-Precipitation of HtpG and NOD2 (WT, V8, V12, V13)

HEK293T cells were transfected with His-tagged HtpG and either Flag-tagged NOD2 WT, V8, V12, or V13. Cobalt bead affinity purification was used to detect proteins that interacted with His-tagged HtpG. Similarly, ANTI-FLAG® M2 Affinity Gel (Sigma, CAT#A2220) purification was used to detect proteins that interacted with Flag-tagged NOD2 WT/variants.

2.8.1: 6x-His Fusion Protein Co-Affinity Precipitation

Two 10 cm culture dishes were seeded with 1.8 x 10^6 HEK293T cells. One plate was transfected with pCR3.1-His-HtpG (B) and pCMV-Flag-NOD2WT/V8/V12/V13 and the other plate was transfected with pCMV-Flag-NOD2WT/V8/V12/V13 alone. After 24 hours, cells were lysed as described in section 2.7.2. His-tagged HtpG was affinity purified as described in section 2.5 with modifications. Instead of columns, 1.5 mL microcentrifuge tubes were used. The collected supernatant was incubated with 300 µL of HisPur Cobalt Resin for 1 hour with end-over-end rotation at 4°C. Beads were then pelleted for 1.5 minutes at 700 xg and 4°C. Unbound sample was collected using a 30G needle and stored as the flow-through sample. The beads were washed 3x 15 minutes with end-over-end rotation at 4°C. Wash buffer was supplemented with 5 mM imidazole. Following each wash, buffer was removed by pelleting the beads as described previously and stored as wash fractions. The beads were then incubated with elution buffer for 30 minutes with end-over-end rotation at 4°C. The bound fraction was collected by pelleting the

beads as described previously. Affinity purifications were repeated on separate occasions three times.

2.8.2: Flag Fusion Protein Co-Immunoprecipitation

Two 15 cm culture dishes were seeded with 5 x 10^6 HEK293T cells. One plate was transfected with pCR3.1-His-HtpG (B) and pCMV-Flag-NOD2WT and the other plate was transfected with pCMV-Flag-NOD2WT alone. After 24 hours, media was removed and cells were lysed as described in section 2.7.3. The affinity purification occurred as described in the Sigma ANTI-FLAG® M2 Affinity Gel immunoprecipitation protocol with minor modification. Instead of 40 µl of ANTI-FLAG M2 affinity gel per 1 mL of cell lysate, 100 µl of affinity gel was used. All centrifugation steps occurred at 18,200 xg. After washing the beads twice with 1X TBS, the beads were washed once with 0.5 mL of 0.1 M glycine HCl, pH 3.5. Cell lysate was incubated with the affinity gel on an end-over-end spinner for at least 2 hours in 4°C. Beads were washed 3 x15 minutes with 1 mL 1X TBS. Protein was eluted using 200 µl of 0.1 M glycine HCl, pH 3.5 as described in the protocol. Co-immunoprecipitations were attempted on three separate occasions.

2.9: Mass Spectrometry of Purified HtpG

HtpG was purified as described in section 2.5 and visualized by SDS-PAGE stained with Coomassie Brilliant Blue as described in section 2.6. One gel slice (purified HtpG in elution lane) of approximately 1 mm x 6.5 mm was excised using a gel cutting pipette tip and incubated in dH_2O for 2 hours. The slice was then cut into smaller 1 mm x 1 mm slices and washed twice with dH_2O. Slices were reduced with 10 mM dithiothreitol at 56°C, then alkylated with 55 mM iodoacetamide for 30 minutes in the dark at room temperature. Then, slices were dried with 200 µl of acetonitrile and immersed in 20 µg/mL of trypsin for 2 hours. Samples were incubated at

37°C overnight following the addition of 50 mM ammonium bicarbonate. Peptides were extracted using a mix of 50% acetonitrile in 5% formic acid (100 µl). Extract was dried using a vacuum centrifuge and pellets were resuspended in 20 µl 3% acetonitrile and 0.5% formic acid.

The processed gel slice was analyzed via electrospray liquid chromatography-tandem mass spectrometry (LC-MS/MS) using a VelosPRO orbitrap mass spectrometer with an UltiMate 3000 Nano-LC. A PicoFRIT C18 self-packed 75 mm x 60 cm capillary column [New Objective, CAT# PF360-50-##-CE-5] was used for initial separation at 300 mL/min. MS data was collected at a resolution of 30,000 with 10 successive MS/MS spectra in higher-energy collisional dissociation and collision-induced dissociation mode.

Sample preparation and raw data analysis was conducted by staff of the Dalhousie Centralized Operation of Research Equipment proteomics facility using Proteome Discoverer 2.2. Spectra were also searched manually against *B. fragilis* group B HtpG as well as the cRAP database of common MS contaminants.

2.10: Enzyme-linked Immunosorbent Assay (ELISA)

HT29 cells were seeded at a density of 10,000 cells per well in a 96-well plate. After approximately 24 hours, cells were stimulated with either TNF (10 ng/mL) [Sigma, CAT#T6674], recombinant *B. fragilis* group B (r) HtpG [CUSABIO, CAT# CSB-EP711569BAAB] (two-fold serial dilutions starting at 50 µg/mL to 0.19 µg/mL), or TNF (10 ng/mL) and rHtpG (two-fold serial dilutions starting at 50 µg/mL to 0.19 µg/mL) and incubated for approximately 18 hours. rHtpG expresses the first 326 amino acids in the N-terminus and contains a 10x-His-sumo-tag in the N-terminus and a Myc-tag in the C-terminus. rHtpG concentration was based on concentrations used by Shelburne et al. (2007). Following the incubation, supernatant was collected and stored at -80°C. Supernatant was used in the

Invitrogen IL-8 Human Matched Antibody Pair kit [CAT# BMS204-3MST] as described by the manufacturer. Samples were run in triplicate wells and two independent experiments were conducted.

2.10.1: Cell Viability

Cell viability was assessed using the alamarBlue Cell Viability Reagent [ThermoFisher Scientific, CAT#DAL1025]. After supernatant was removed, alamarBlue Reagent was added to the cells and incubated at 37°C for 2 hours. The resulting absorbance was read on a spectrophotometer at 570 nm and plotted accordingly.

2.11: Inflammatory Cytokine Antibody Profiling

HT29 cells were seeded at a density of 0.3 x 10^6 cells per well in a 6-well plate. After approximately 24 hours, cells were stimulated with either TNF (10 ng/mL), rHtpG (50 μg/mL), or TNF (10 ng/mL) and rHtpG (50 μg/mL) and incubated for approximately 18 hours. Following the incubation, supernatant was collected and stored at -80°C. All reagents were components of the abcam Human Inflammation Antibody Array Membrane [CAT# ab134003] and were used as described by the manufacturer. As a note, the membranes were incubated with supernatant overnight at 4°C shaking. As such, the membranes were washed following the Large Volume Wash protocol. Membranes were developed using chemiluminescence. Cell viability was assessed using the alamarBlue Cell Viability reagent described in section 2.10.1.

2.12: Statistical Analysis

GraphPad Prism 4 (GraphPad) software was used to test for statistical significance. To compare multiple groups, one-way analysis of variance (ANOVA) test followed by multiple comparison of means Bonferroni test was applied. Data are shown as mean ± standard error of the mean (SEM). *P* values ≤ 0.05 were considered statistically significant.

CHAPTER 3 RESULTS

3.1: Dry-lab Analysis of HtpG

Our lab previously conducted a metagenomic study identifying HtpG as a candidate for distinguishing CD patients able to sustain remission from those who relapse (Dunn et al., 2016). In comparison to patients who relapsed, patients able to sustain remission had a higher relative abundance of HtpG. Despite knowing this pattern, little was known about HtpG in the context of pediatric CD.

3.1.1: HtpG Relative Abundance is Markedly Increased in the Gut

To put HtpG in the context of CD, I examined HtpG relative abundance at several different body sites from healthy individuals. Metagenomic data were downloaded from the Human Microbiome Project online database for five major body sites including nose, mouth, skin, gut (represented by stool samples), and urogenital tract. Throughout the different body sites, *DnaK*, another bacterial HSP, maintained a consistent relative abundance (Figure 4). *HtpG*, in comparison, had a lower relative abundance than *DnaK* at every body site except the gut. These data suggest that at least amongst healthy individuals, the gut is a unique reservoir for *HtpG* read counts.

3.1.2: HtpG Relative Abundance and Bacterial Contributors Change During and After EEN

To examine how HtpG changes in response to EEN, specifically observing *HtpG* relative abundance and bacterial source, I analyzed metagenomic data collected for the Dunn et al. (2016) study from fourteen CD patients during a 96-week period. *HtpG* relative abundance fluctuated to varying degrees for each patient with no obvious pattern associated with those patients who sustained remission (CD2, CD3, CD9, CD10) and those patients who relapsed (Figure 5). One possible confounder is fluctuations in other organisms lacking *HtpG* in the

presence of steady *HtpG* gene levels. As *HtpG* relative abundance changed so too did the bacterial phyla contributing *HtpG* read counts. Figure 6 highlights two patients—CD1 (non-sustained remission) and CD2 (sustained remission)—who were not on any medications at baseline. Both patients began EEN with *HtpG* derived from Bacteroidetes and Firmicutes, although to varying abundance levels. At the 12-week EEN endpoint, the main bacteria contributing *HtpG* in patient CD1 were Actinobacteria and Firmicutes, whereas the main bacterium in patient CD2 was Bacteroidetes. Figure 7 examines the bacterial phyla producing HtpG in two patients who began EEN on antibiotics. Patient CD3, who sustained remission, maintained a Firmicutes signature throughout EEN and gained an additional signal from Verrucomicrobia. Patient CD5, who relapsed, began and ended EEN with *HtpG* read counts derived from Bacteroidetes and Firmicutes. Patient CD5 provided an additional fecal sample at 24-weeks, which showed a decrease in Bacteroidetes-derived *HtpG* read counts and an increase in Firmicutes-derived *HtpG* read counts. The data did not clearly describe a pattern of *HtpG* relative abundance following the end of EEN, thus I decided to examine HtpG group identity.

3.1.3: HtpG Groups A and B Found Consistently Throughout and Following EEN

Gene trees describe the evolutionary history of a gene of interest providing evidence for gene duplication and speciation events. The program RAxML, which generates gene trees using aligned amino acid sequences, was used to determine HtpG group identity. Among the 14 CD patients, 6 random patients were examined (CD4, CD6, CD8, CD10, CD13, and CD14) at three separate timepoints—baseline, 12-weeks, and 24-weeks. The three corresponding gene trees had similar branch lengths, thus could be compared to one another. In the patient samples, only groups A and B HtpG, never group C, were identified at the three timepoints (Figure 8). Group A was supported by a 81% bootstrap value at baseline, a 69% bootstrap value at 12-weeks, and a

80% bootstrap value at 24-weeks. The lower bootstrap value at 12-weeks potentially indicates that the initial alignment could be improved. Group B was supported at all timepoints by 100% bootstrap values. Group A+B were supported by a 93% bootstrap value at baseline, a 89% bootstrap value at 12-weeks, and a 96% bootstrap value at 24-weeks. HtpG group identity was also examined using fecal, saliva, and skin metagenomic samples from healthy individuals. HtpG groups A and B were found in the fecal and saliva samples (Figure 9 and 10, respectively). In Figure 9, group A was supported by a 67% bootstrap value and group B was supported by a 100% bootstrap value. Groups A+B were supported by a 70% bootstrap value. In Figure 10, groups A+B were supported by a 77% bootstrap value and group B was supported by a 71% bootstrap value. In Figure 11, HtpG groups B and C were found in skin samples. Group B was supported by a 100% bootstrap value and group C was supported by a 74% bootstrap value.

All of this goes to show that HtpG groups may be body site specific. As the gut appears to only contain groups A and B HtpG, I decided to examine the effects of group B HtpG from the commensal gut bacterium, *B. fragilis*, on innate immunity.

3.2: Extracellular HtpG Activity

3.2.1: HtpG Dampens TNF-Induced CXCL8 Production

Peptides and proteins can be detected and quantified using ELISA, a plate-based assay that involves the immobilization of an antigen and subsequent detection using enzyme-conjugated antibodies. As previous work by Shelburne et al. (2007) observed concentration-dependent CXCL8 secretion following treatment with *P. gingivalis,* I conducted a CXCL8 ELISA to measure secretion following treatment with group B *B. fragilis* HtpG. HT29 cells were stimulated with two-fold serial dilutions of *B. fragilis* group B rHtpG starting at 50 μg/mL. HtpG activity was concentration-dependent with concentrations between 6.25-12.5 μg/mL increasing

CXCL8 secretion (Figure 12A). HtpG concentrations below 6.25 μg/mL or above 12.5 μg/mL did not significantly increase CXCL8 secretion above that observed in the untreated control. Next, to examine HtpG inflammatory activity in an inflamed environment, I co-treated cells with the same serial-dilution of rHtpG and a constant concentration of TNF. As demonstrated in Figure 13, TNF-treatment alone caused HT29 cells to secrete a high concentration of CXCL8. When cells were co-treated with TNF and rHtpG, only the highest concentration of HtpG (50 μg/mL) significantly reduced CXCL8 secretion in comparison to TNF-treated cells. To confirm HtpG activity when co-treated with TNF, HT29 cells were stimulated with either TNF (10 ng/mL), rHtpG (50 μg/mL), or TNF (10 ng/mL) and rHtpG (50 μg/mL). Once again, a high concentration of rHtpG in combination with TNF significantly reduced CXCL8 secretion (Figure 14). Cell viability was not significantly impacted by treatment conditions (Figure 12-14B). In order to get a better understanding of the TNF-dampening property of HtpG, I examined the response of forty inflammatory mediators to HtpG and/or TNF stimulation.

3.2.2: HtpG Activates Inflammatory Mediators at High Concentrations

Cytokine arrays (like ELISAs but membrane-based) allow for simultaneous analysis of multiple cytokines in the same sample. To get an overview of the inflammatory mediators produced following rHtpG, TNF, or rHtpG and TNF treatment, I applied *B. fragilis* (group B) rHtpG (50 μg/mL), TNF (10 ng/mL), or rHtpG (50 μg/mL) and TNF (10 ng/mL) to HT29 cells for 18 hours and then applied the supernatants to four separate membranes (one for untreated cells). The assay qualitatively surveyed the inflammatory profile induced by the test conditions; the larger the dot, the greater the concentration of the inflammatory mediator. Untreated HT29 cells had some inflammatory signals, which is not unexpected for a transformed cancer cell line (Figure 15). Notably, untreated HT29 cells produced IL-8 and tissue inhibitors of

metalloproteinases (TIMP)-2. Additional inflammatory markers such as IL-3, IL-16, lymphotoxin, soluble TNF RII, PDGF-BB, CCL24, M-CSF, CCL15, soluble TNF RI, CXCL10, IL-6 soluble receptor, CCL4, CCL5, and IL-12 p40 were also produced by untreated HT29 cells, but to a lesser degree than CXCL8 and TIMP-2. TNF-treated HT29 cells induced IL-3, IL-16, lymphotoxin, CXCL8, CCL11, CXCL10, CCL2, CCL24, PDGF-BB, ICAM-1, IFNγ, CCL4, and IL-12 p40. rHtpG treatment induced an inflammatory profile with similarities to TNF, but with some notable differences: IL-13, IL-15, TNF, CCL3, CXCL9, CCL8, CXCL10, IL-17, and IL-7 expression was increased; and soluble TNF RI, soluble TNF RII, TIMP-2, CCL4, and CCL15 expression was decreased. Finally, the combination of rHtpG and TNF treatment produced an inflammatory profile that more closely resembled rHtpG treatment than TNF treatment.

HtpG potentially induces the observed inflammatory mediator profile by interacting with extra- and/or intra-cellular PRRs. To gain insight into whether HtpG can act intra-cellularly, I turned to affinity purification.

3.3: Intracellular HtpG Activity

3.3.1: HtpG Protein Expression

To examine intracellular interaction between HtpG and NOD2 (WT and variants), I began by cloning group B *B. fragilis* HtpG. My initial attempts to clone HtpG (B) from *B. fragilis* gDNA into *E. coli* (using pET28a-HtpG) resulted in elongated *E. coli*, potentially indicating activation of the SOS response (Figure 16B). As such, I synthesized group B and C *B. fragilis* HtpG. Group B HtpG was successfully cloned into pCR3.1 as determined by sequencing (Figure 16C), but group C remains un-cloned due to technical difficulties. The pCR3.1 vector allows for both mammalian and bacterial expression. As shown in Figure 16D, HtpG was cobalt

bead affinity purified from HEK293T cells transfected with pCR3.1-HtpG (B). HtpG purification methods were confirmed using LC-MS/MS.

3.3.2: Co-Affinity Purification of HtpG and NOD2 (WT, V8, V12, V13)

Protein-protein interaction can be assessed using common approaches such as co-immunoprecipitation and co-affinity purification, whereby a protein of interest expressing an affinity or epitope tag is purified from cell lysate and the presence of interacting proteins are determined. In my experiments, I used 6x-His-tagged HtpG and FLAG-tagged NOD2 WT/V8/V12/V13 to transfect HEK293T cells that have no endogenous NOD2.

HEK293T cell lysates co-expressing His-HtpG or FLAG-NOD2 (either WT, V8, V12, or V13) were subject to cobalt bead His-affinity purification, competitive elution with imidazole, and analysis by SDS-PAGE and Western blotting. The presence of NOD2 (~110 kDa) was detected using anti-FLAG antibodies and the presence of HtpG (~72 kDa) was detected using anti-His antibodies. HEK293T cells transfected with only NOD2 (WT/V8/V12/V13) were used as negative controls. Expression of His-HtpG (Figure 17A, lane 8) did not aid in the recovery of NOD2 WT as measured by anti-FLAG immunoblot (Figure 17B, lane 8). However, expression of His-HtpG did aid in the recovery of NOD2 V8 (Figure 18B), V12 (Figure 19B), and V13 (Figure 20B) as measured by anti-FLAG immunoblot. Additional species were observed in all Western blots representing, most likely, breakdown products of expressed proteins or cross-reacting proteins.

To clarify the interaction, anti-FLAG beads were used instead of cobalt beads for purification. FLAG-NOD2 proteins were eluted and analyzed via SDS-PAGE and Western blotting. Despite multiple attempts to elute FLAG-NOD2, no protein was observed in the eluate (data not shown). Furthermore, FLAG-NOD2 was not observed following bead boiling at 95°C

for 5 minutes (data not shown). Due to a lack of time, FLAG purification protocol could not be optimized.

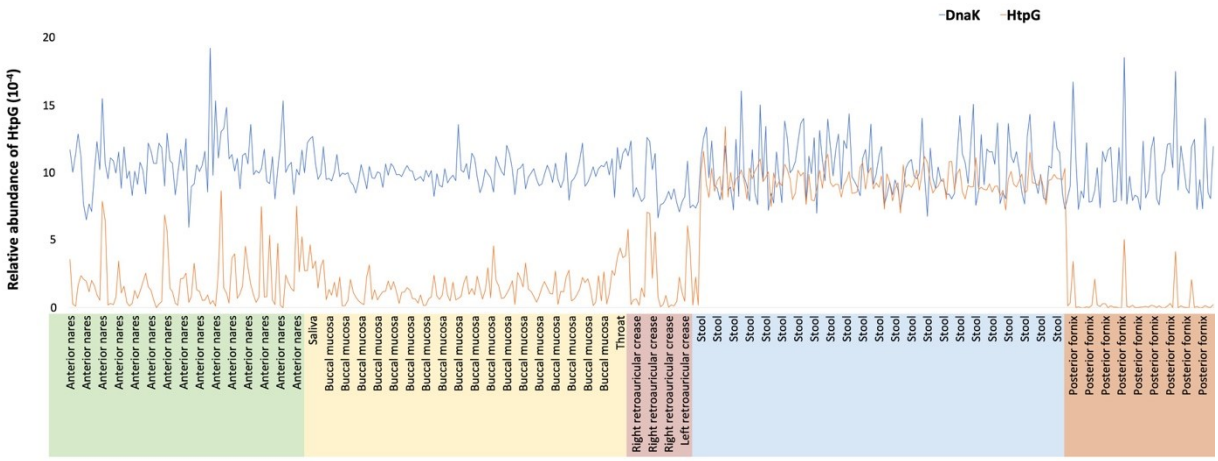


Figure 4. HtpG relative abundance is consistently higher in stool samples than in samples collected at nasal, oral, skin, and urogenital tract sites from healthy controls (Human Microbiome Project database). The relative abundance of HtpG was plotted for each body site sampled: nose is green, mouth is yellow, skin is red, gut is blue, and urogenital tract is orange. Each body site is displayed with a variable number of samples that reflect the availability on the Human Microbiome Project database. HtpG is represented by an orange line and DnaK, another bacterial HSP, is represented by a blue line.

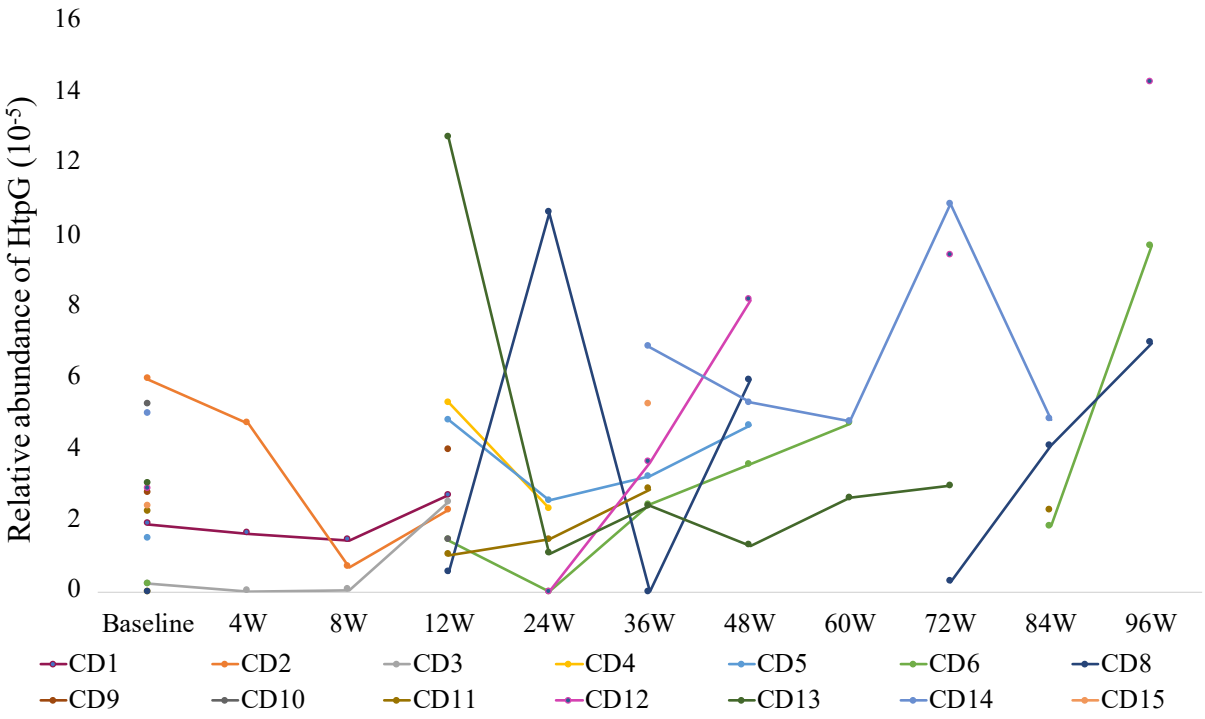


Figure 5. *HtpG* relative abundance fluctuates throughout and following EEN for 14 CD patients enrolled in the Dunn et al. (2016) study. Fecal samples were collected from each CD patient at various timepoints (not necessarily every timepoint). Metagenomic sequencing was performed on DNA collected from fecal samples. The relative abundance of HtpG was plotted for each time point sampled. Patients with a relative abundance of zero were still plotted. Each patient is represented by a different color.

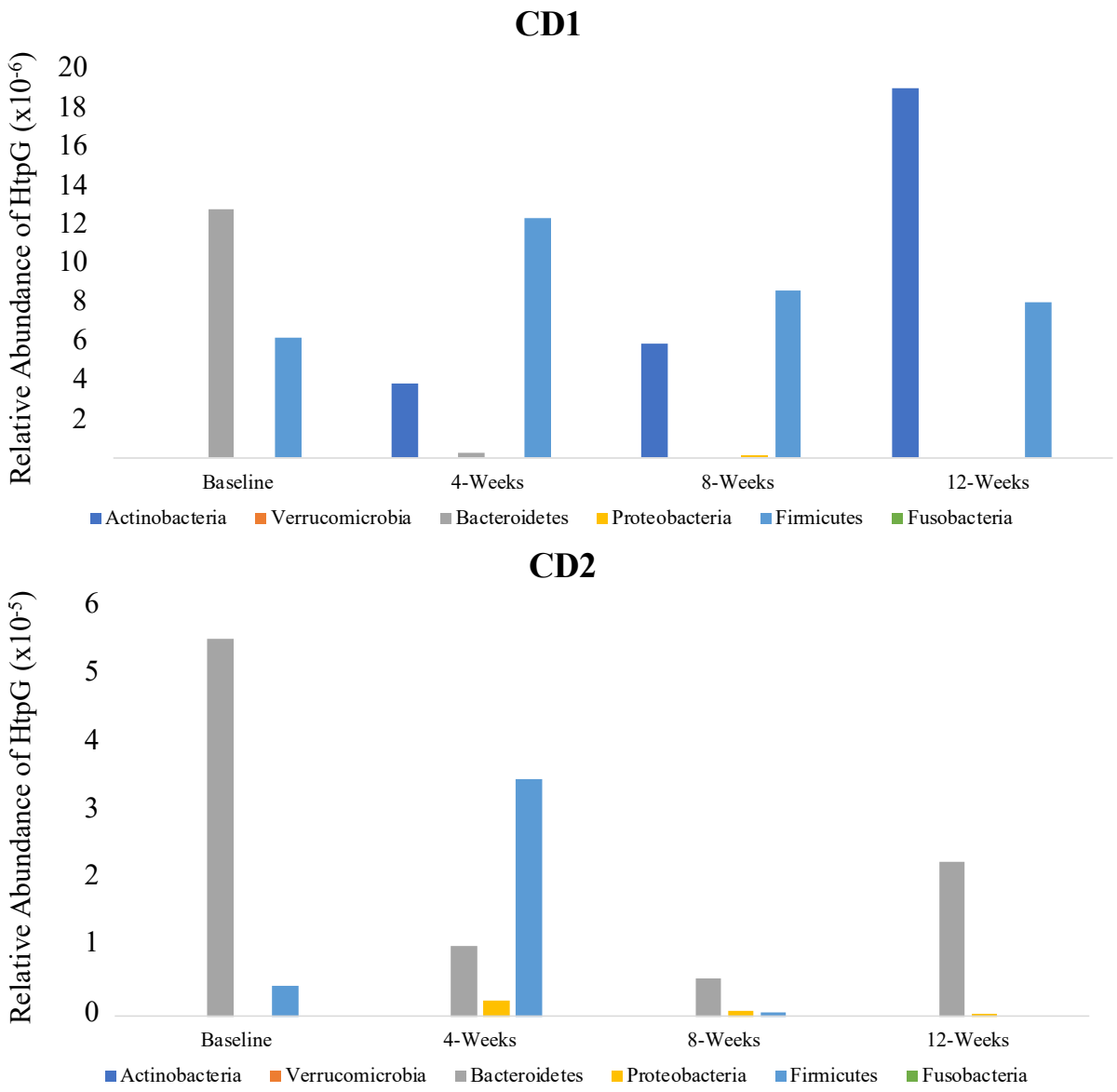


Figure 6. Bacterial source of *HtpG* changed throughout EEN in two patients (CD1 and CD2) who received no other medications during EEN therapy. Fecal samples were collected from each CD patient at baseline, 4-weeks, 8-weeks, and 12-weeks. Metagenomic sequencing was performed on DNA collected from fecal samples. The relative abundance of *HtpG* was plotted for each time point. At 12-weeks, patient CD1 was classified as non-sustained remission and patient CD2 was classified as sustained remission based on wPCDAI scores. The six most prominent phyla contributing *HtpG* are represented by different colours.

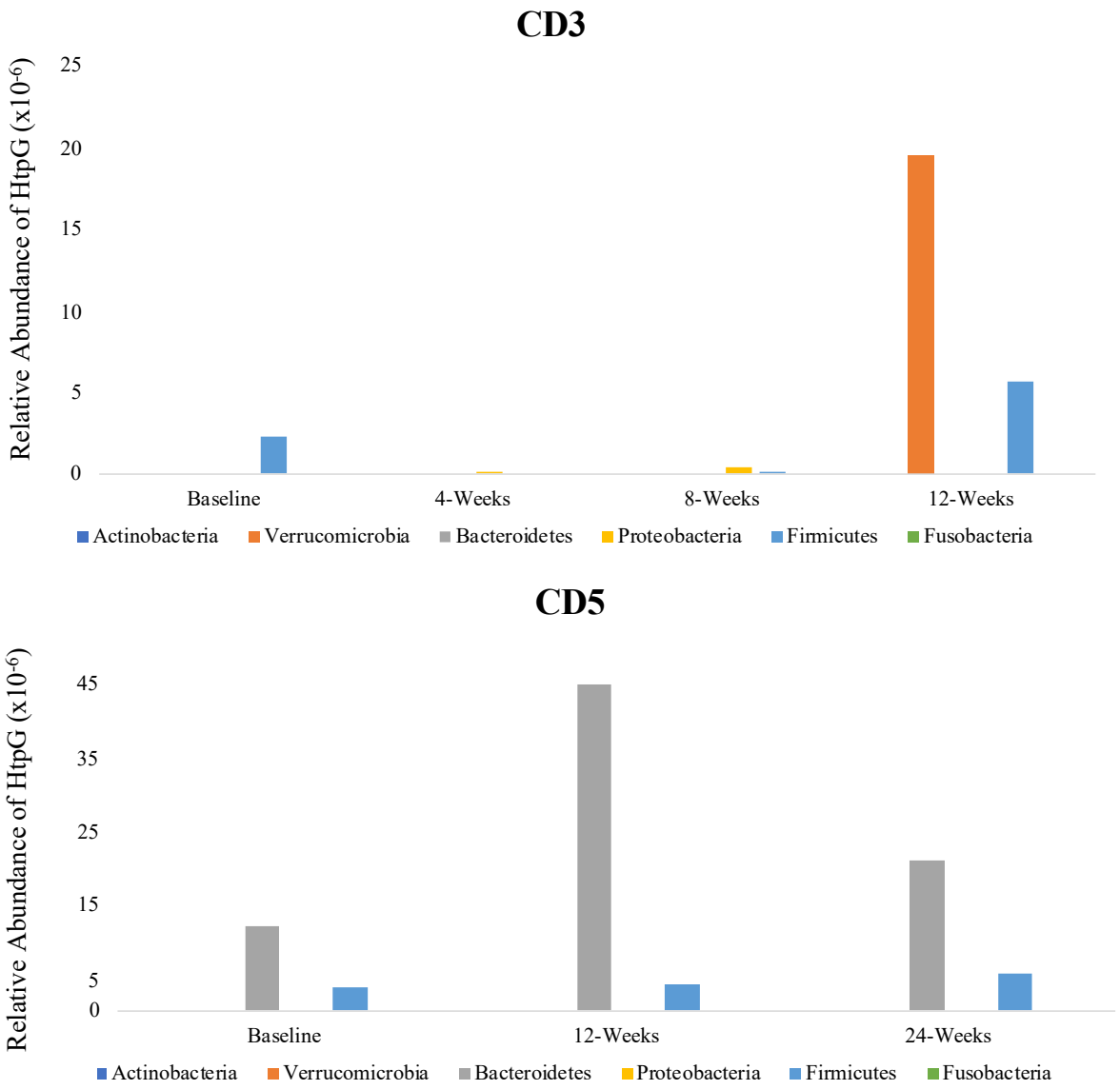


Figure 7. Bacterial source of *HtpG* changed throughout EEN in two patients (CD3 and CD5) who received other medications during EEN therapy. Patient CD3 provided fecal samples at baseline, 4-weeks, 8-weeks, and 12-weeks. Patient CD5 provided fecal samples at baseline, 12-weeks, and 24-weeks. Metagenomic sequencing was performed on DNA collected from fecal samples. The relative abundance of *HtpG* was plotted for each time point sampled. At 12-weeks, patient CD3 was classified as sustained remission and patient CD5 was classified as non-sustained remission based on wPCDAI scores. The six most prominent phyla contributing *HtpG* are represented by different colours.

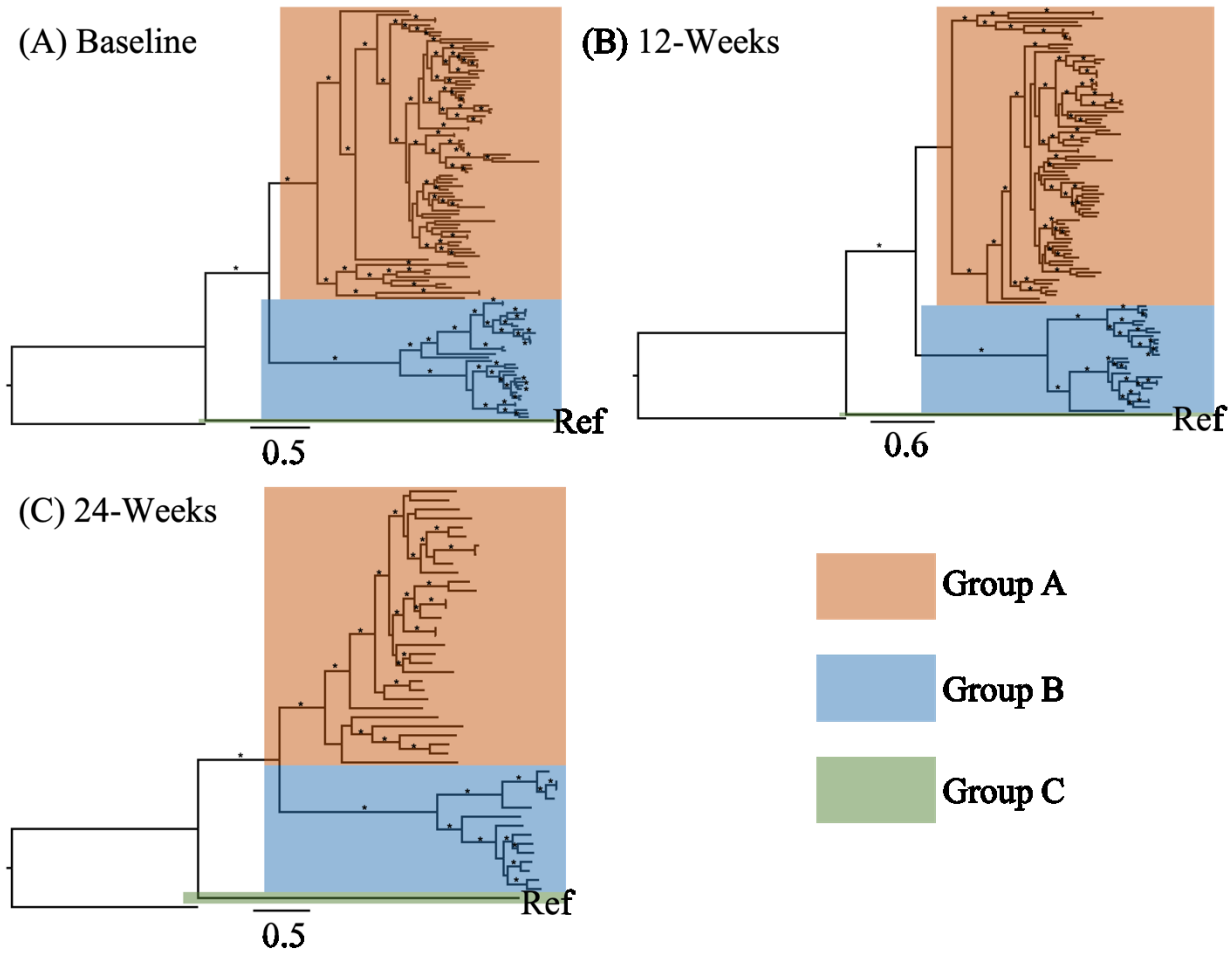


Figure 8. Group C HtpG not identified in metagenomic sequencing of six CD patient fecal samples at three separate time points (baseline, 12-weeks, and 24-weeks). From the metagenomic dataset collected for the Dunn et al. (2016) study, *HtpG* nucleotide fragments were identified from six CD patients (CD4, CD6, CD8, CD10, CD13, and CD14) at (A) baseline, (B) 12-weeks, and (C) 24-weeks. *HtpG* query fragments were compared to a protein sequence database using blastx. Protein targets with e-values <0.001 were selected for downstream alignment. MAFFT was used to align amino acid sequences and RAxML was used to generate gene trees. RAxML was run with GAMMA models using the LG exchangeability matrix and frequencies estimated from the data. Trees do not show density as bacterial species are only represented once. Bootstrap support was estimated from 100 bootstrap replicates and nodes with support greater than 70% are indicated as an asterisk on the tree. Trees are rooted with HtpG from *Methanosarcina mazei.* REF refers to group C *B. fragilis* HtpG amino acid sequence used to demonstrate group C, but not found in patient data. The scale for the branch lengths gives the mean number of substitutions per amino acid.

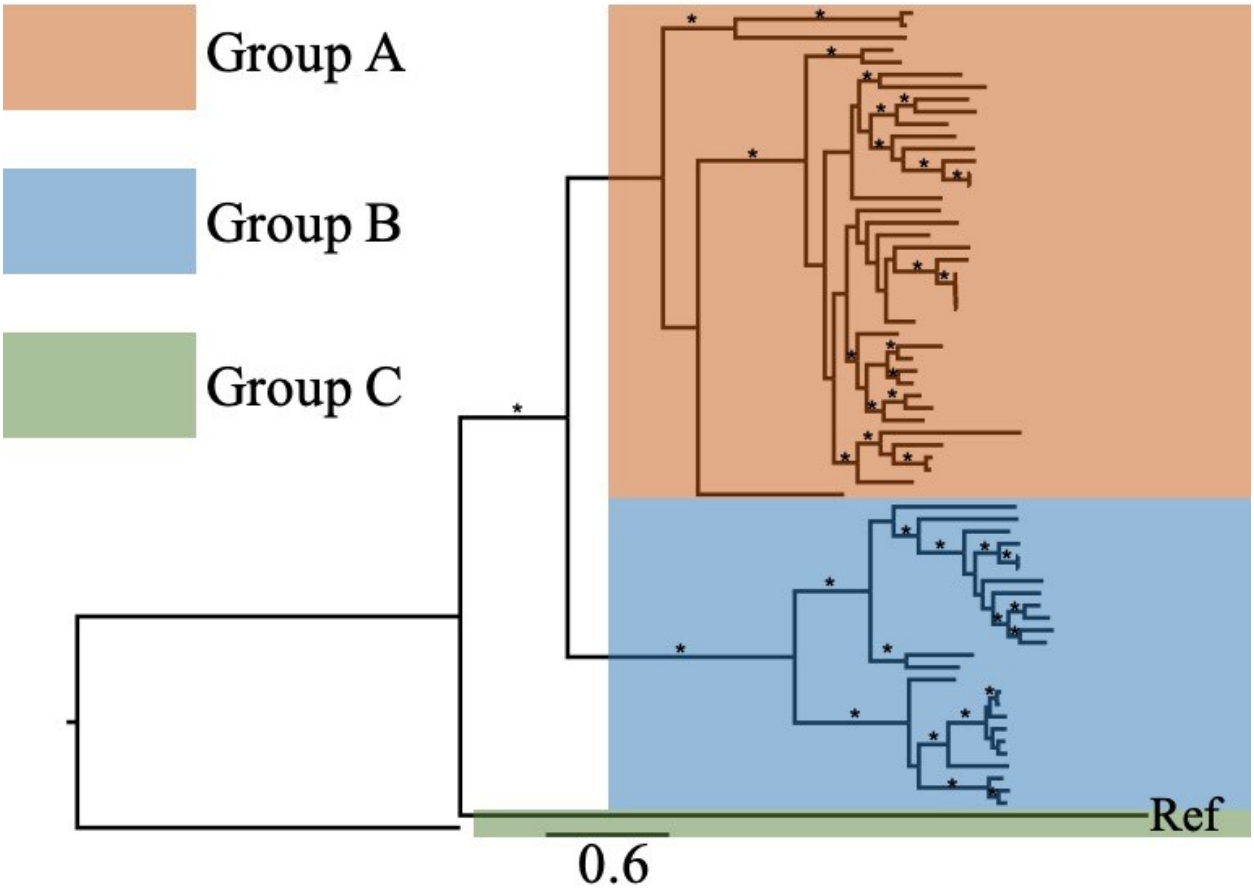


Figure 9. Group C HtpG not identified in metagenomic sequencing of healthy control fecal sample (sibling of patient CD10). From the metagenomic dataset collected for the Dunn et al. (2016) study, *HtpG* nucleotide fragments were identified from the healthy control. *HtpG* query fragments were compared to a protein sequence database using blastx. Protein targets with e-values <0.001 were selected for downstream alignment. MAFFT was used to align amino acid sequences and RAxML was used to generate gene trees. RAxML was run with GAMMA models using the LG exchangeability matrix and frequencies estimated from the data. Trees do not show density as bacterial species are only represented once. Bootstrap support was estimated from 100 bootstrap replicates and nodes with support greater than 70% are indicated as an asterisk on the tree. Trees are rooted with HtpG from *Methanosarcina mazei.* REF refers to group C *B. fragilis* HtpG amino acid sequence used to demonstrate group C, but not found in patient data. The scale for the branch lengths gives the mean number of substitutions per amino acid.

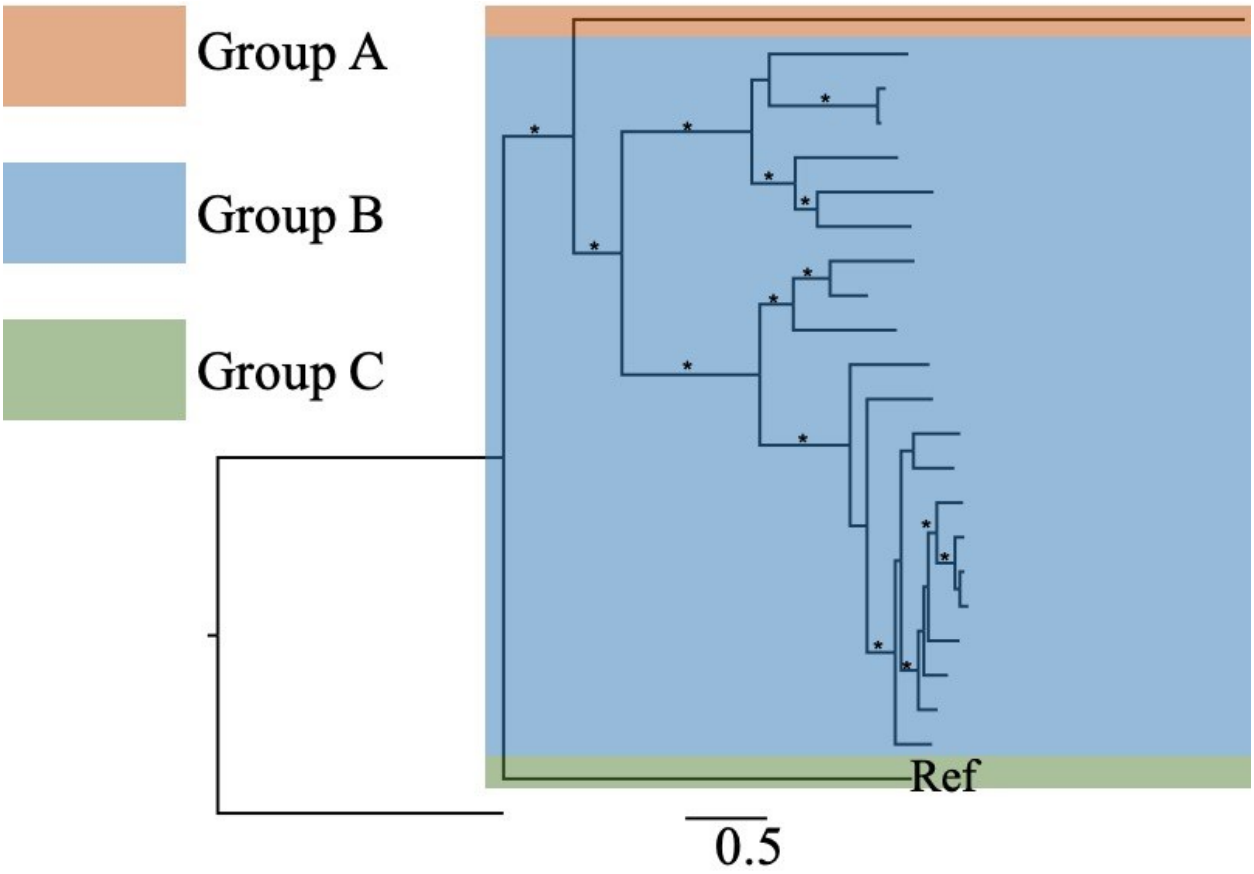


Figure 10. Group C HtpG not identified in metagneomic sequencing of healthy control saliva sample (Human Microbiome Project database). *HtpG* nucleotide sequences were identified from saliva samples donated by five healthy individuals on the Human Microbiome Project online database. *HtpG* query fragments were compared to a protein sequence database using blastx. Protein targets with e-values <0.001 were selected for downstream alignment. MAFFT was used to align amino acid sequences and RAxML was used to generate gene trees. RAxML was run with GAMMA models using the LG exchangeability matrix and frequencies estimated from the data. Trees do not show density as bacterial species are only represented once. Bootstrap support was estimated from 100 bootstrap replicates and nodes with support greater than 70% are indicated as an asterisk on the tree. Trees are rooted with HtpG from *Methanosarcina mazei.* REF refers to group C *B. fragilis* HtpG amino acid sequence used to demonstrate group C, but not found in patient data. The scale for the branch lengths gives the mean number of substitutions per amino acid.

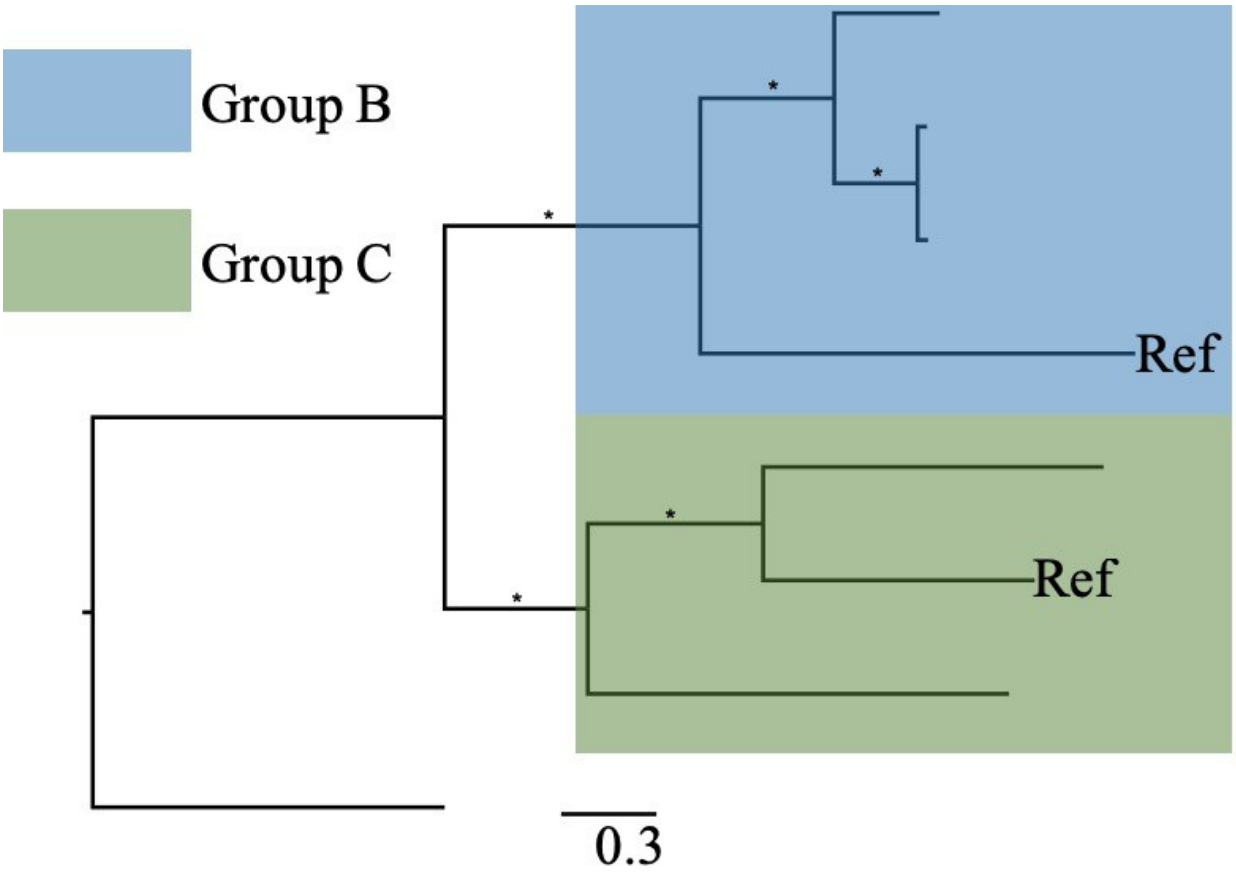
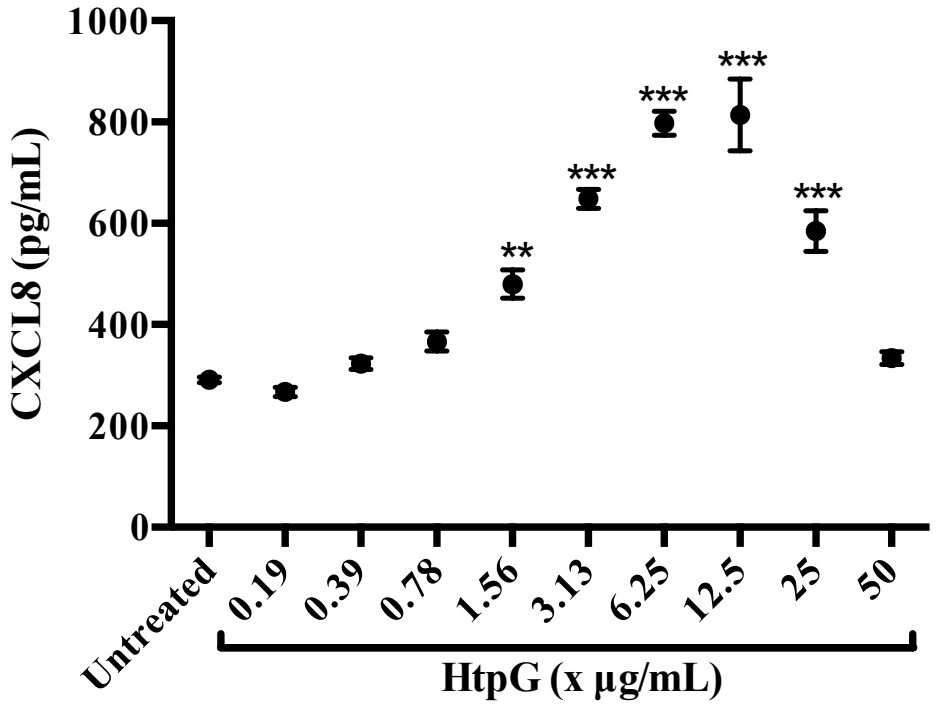


Figure 11. Group C HtpG identified in metagneomic sequencing of healthy control skin sample (Human Microbiome Project database). *HtpG* nucleotide sequences were identified from skin samples donated by four healthy individuals on the Human Microbiome Project online database. *HtpG* query fragments were compared to a protein sequence database using blastx. Protein targets with e-values <0.001 were selected for downstream alignment. MAFFT was used to align amino acid sequences and RAxML was used to generate gene trees. RAxML was run with GAMMA models using the LG exchangeability matrix and frequencies estimated from the data. Trees do not show density as bacterial species are only represented once. Bootstrap support was estimated from 100 bootstrap replicates and nodes with support greater than 70% are indicated as an asterisk on the tree. Trees are rooted with HtpG from *Methanosarcina mazei.* REF refers to group B and C *B. fragilis* HtpG amino acid sequence used to demonstrate group B and C, but not found in patient data. The scale for the branch lengths gives the mean number of substitutions per amino acid.

A)


1000
800
600
400
200
0
CXCL8 (pg/mL)
**
***
***
***
***
Untreated
0.19
0.39
0.78
1.56
3.13
6.25
12.5
25
50
HtpG (x µg/mL)

B)

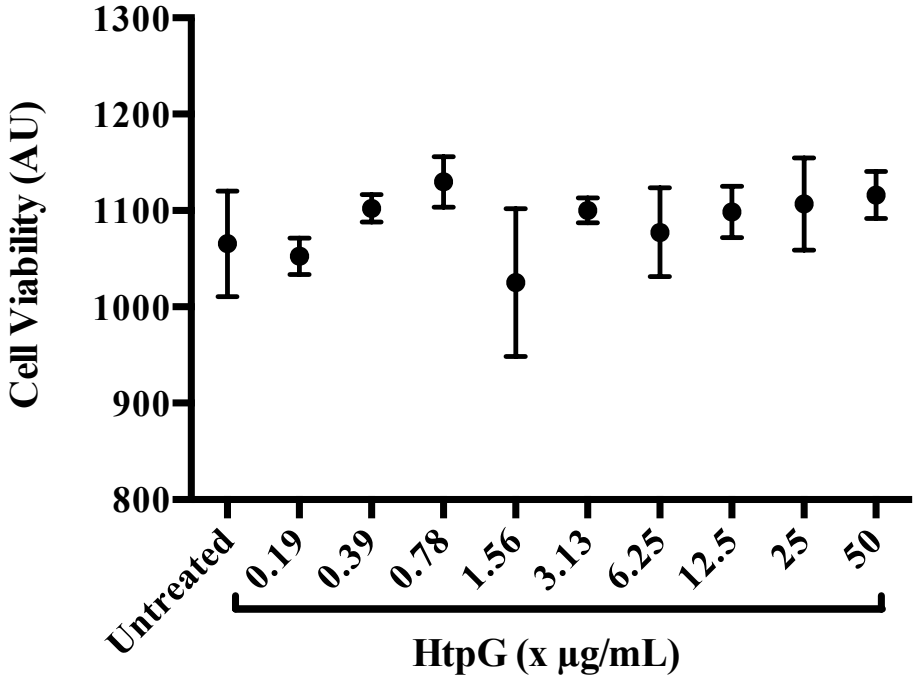
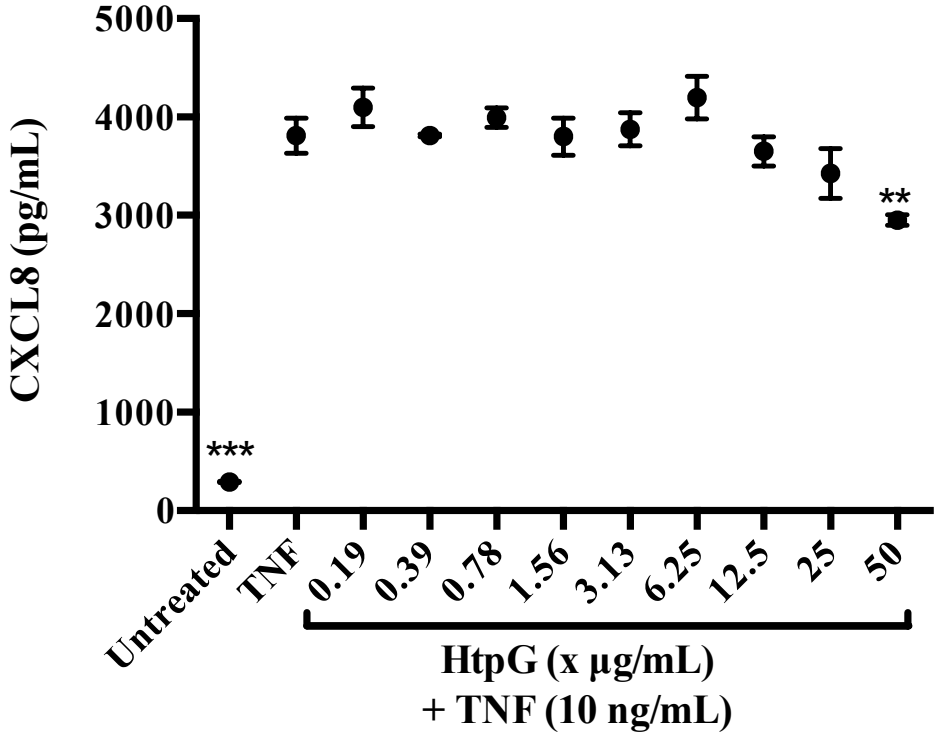

1300
1200
1100
1000
900
800
Cell Viability (AU)
Untreated
0.19
0.39
0.78
1.56
3.13
6.25
12.5
25
50
HtpG (x µg/mL)

Figure 12. Group B *B. fragilis* HtpG acts in a concentration-dependent manner to stimulate CXCL8 secretion in HT29 cells. HT29 cells were stimulated for 18 hours with HtpG alone (starting at 50 μg/mL followed by two-fold serial dilution). (A) CXCL8 levels were measured in supernatant by ELISA colorimetric detection. (B) Cell viability was measured using alamarBlue Cell Viability Reagent. Data are presented as mean ± SEM of triplicate wells and are representative of two independent experiments. The mean of each condition was compared to the mean of the untreated. $^{**}P \leq 0.01$, $^{***}P \leq 0.001$ using a one-way ANOVA test with Bonferonii multiple-comparison test analysis.

A)


5000
4000
3000
2000
1000
0
CXCL8 (pg/mL)
***
**
Untreated
TNF
0.19
0.39
0.78
1.56
3.13
6.25
12.5
25
50
HtpG (x µg/mL)
+ TNF (10 ng/mL)

B)

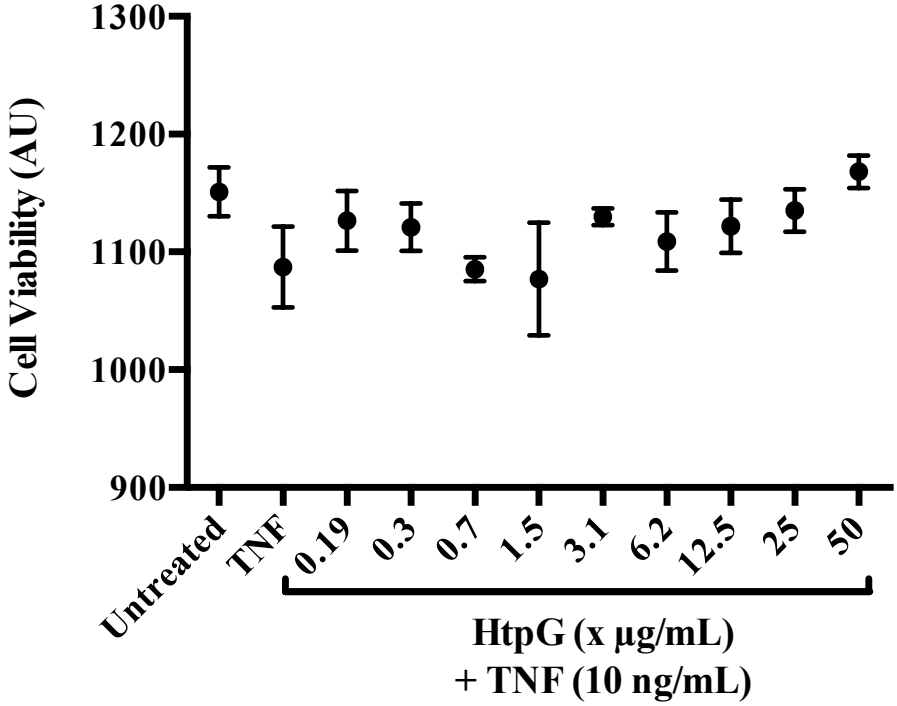
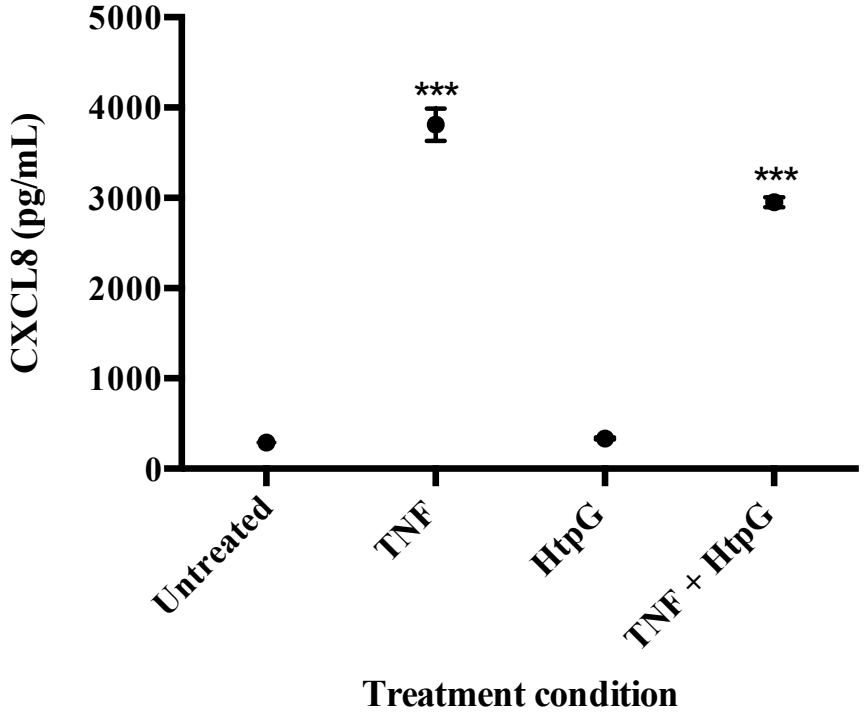

1300
1200
1100
1000
900
Cell Viability (AU)
Untreated
TNF
0.19
0.3
0.7
1.5
3.1
6.2
12.5
25
50
HtpG (x µg/mL)
+ TNF (10 ng/mL)

Figure 13. Group B *B. fragilis* HtpG, in combination with TNF, does not act in a concentration-dependent manner to stimulate CXCL8 secretion in HT29 cells. HT29 cells were stimulated for 18 hours with TNF alone (10 ng/mL) or with TNF (10 ng/mL) and HtpG (starting at 50 μg/mL followed by two-fold serial dilution). Supernatants were subsequently collected and incubated with the IL-8 Human Matched Antibody Pair kit [CAT# BMS204-3MST]. Colorimetric detection was used to measure CXCL8 expression. (B) Cell viability was measured using alamarBlue Cell Viability Reagent. Data are presented as mean ± SEM of triplicate wells. The mean of each condition (untreated and HtpG + TNF) was compared to the mean of the TNF treatment.$^{**}P \leq 0.01$, $^{***}P \leq 0.001$ using a one-way ANOVA test with Bonferonii multiple-comparison test analysis.

A)


5000
4000
3000
2000
1000
0
CXCL8 (pg/mL)
***
***
Untreated
TNF
HtpG
TNF + HtpG
Treatment condition

B)

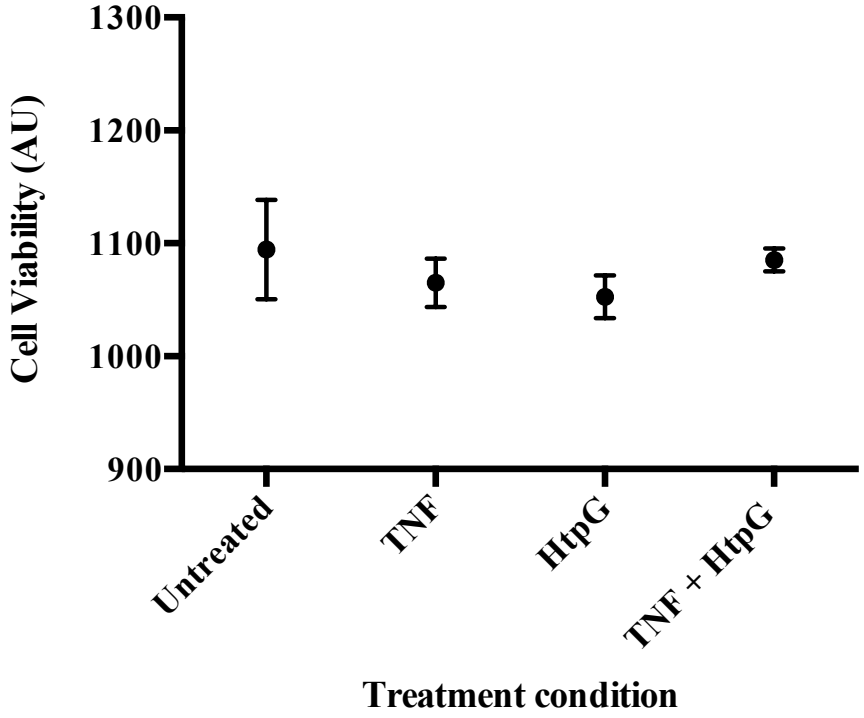
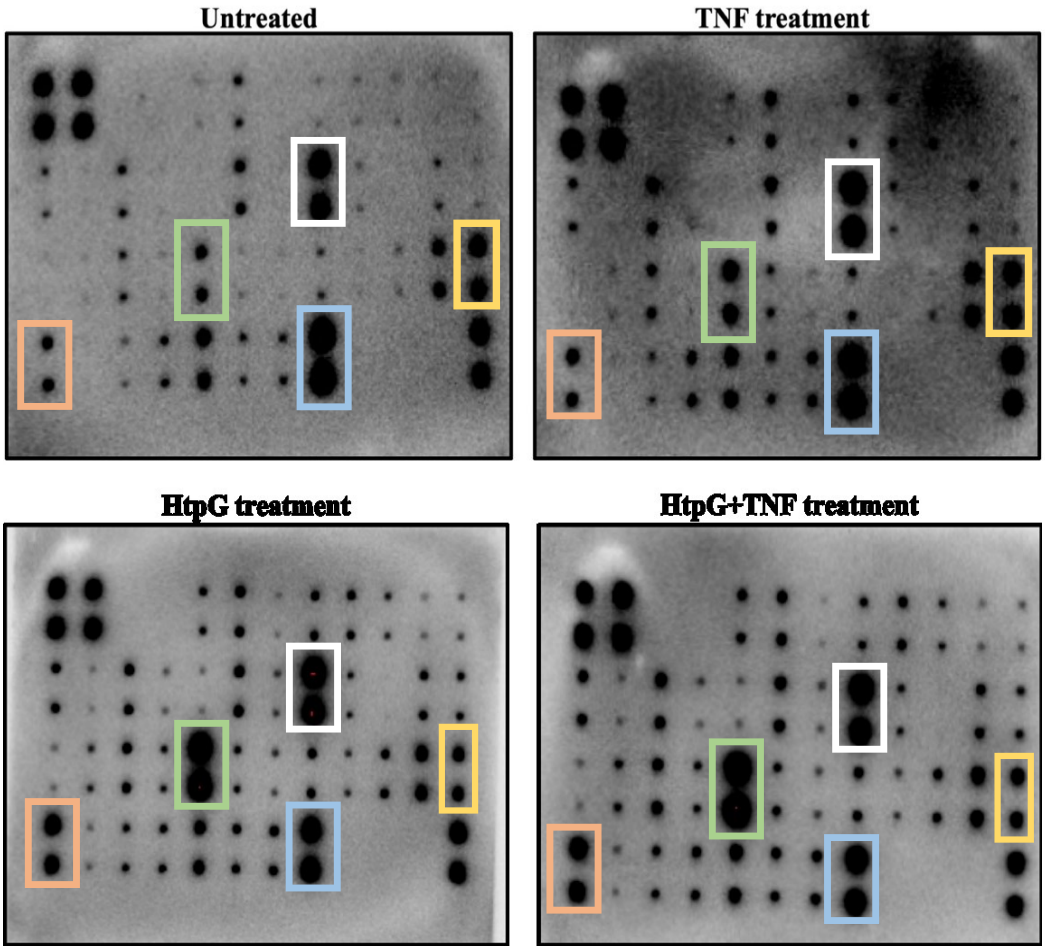

1300
1200
1100
1000
900
Cell Viability (AU)
Untreated
TNF
HtpG
TNF + HtpG
Treatment condition

Figure 14. TNF-induced CXCL8 secretion in HT29 cells is reduced by high concentration of group B *B. fragilis* HtpG. HT29 cells were stimulated for 18 hours with either TNF alone (10 ng/ml), HtpG alone (50 μg/mL), or TNF (10 ng/ml) and HtpG (50 μg/mL). Supernatants were subsequently collected and incubated with the IL-8 Human Matched Antibody Pair kit [CAT# BMS204-3MST]. Colorimetric detection was used to measure CXCL8 expression. (B) Cell viability was measured using alamarBlue Cell Viability Reagent. Data are presented as mean ± SEM of triplicate wells and are representative of two independent experiments. The mean of each condition was compared to the mean of the untreated.$***P \leq 0.001$ using a one-way ANOVA test with Bonferonii multiple-comparison test analysis

(A)



(B)

1	**9**	**17**	**25**	**33**	**41**	**49**	**57**	**65**	**73**	**81**	**89**
2	10	18	26	34	42	50	58	66	74	82	90
3	11	19	27	35	43	51	59	67	75	83	91
4	12	20	28	36	44	52	60	68	76	84	92
5	13	21	29	37	45	53	61	69	77	85	93
6	14	22	30	38	46	54	62	70	78	86	94
7	15	23	31	39	47	55	63	71	79	87	95
8	16	24	32	40	48	56	64	72	80	88	96

Positive Controls	1, 2, 9, 10, 95, 96	**TNF**	23, 24	**S TNF RII**	47, 48	**CXCL9**	69, 70
Negative Controls	17, 18, 25, 26, 87, 88	**IL-4**	27, 28	**GCSF**	49, 50	**IFN-γ**	73, 74
Blank	71, 72, 79, 80	**IL-17**	29, 30	**IL-7**	51, 52	**IL-11**	75, 76
IL-1β	3, 4	**Lymphotoxin**	31, 32	**CCL8**	53, 54	**CCL3**	77, 78
IL-13	5, 6	**CCL11**	33, 34	**PDGF-BB**	55, 56	**I-309**	81,82
CCL5	7, 8	**IL-6**	35, 36	**GM-CSF**	57, 58	**IL-12 p40**	83, 84
IL-2	11, 12	**CXCL10**	37, 38	**IL-8**	59, 60	**CCL4**	85, 86
IL-15	13, 14	**S TNF RI**	39, 40	**M-CSF**	61, 62	**IL-1α**	89, 90
TGF-β1	15, 16	**CCL24**	41, 42	**TIMP-2**	63, 64	**IL-12 p70**	91, 92
IL-3	19, 20	**IL-6sR**	43, 44	**ICAM-I**	65, 66	**CCL15**	93, 94
IL-16	21, 22	**CCL2**	45, 46	**IL-10**	67, 68		

Figure 15. High concentration of group B *B. fragilis* HtpG in combination with TNF shifts the inflammatory cytokine profile to look more like HtpG treatment alone. (A) HT29 cells were stimulated for 18 hours with either TNF alone (10 ng/ml), HtpG alone (50 μg/mL), or TNF (10 ng/ml) and HtpG (50 μg/mL). Supernatants were subsequently collected and incubated with Human Inflammation Antibody Array Membranes. Membranes were developed using chemiluminescence. Notable mediators are indicated by coloured boxes: white, IL-8; light blue, TIMP2; light green, CXCL10; orange, CCL5; and yellow, CCL15. (B) Identity index describing the inflammatory mediator represented by each dot.

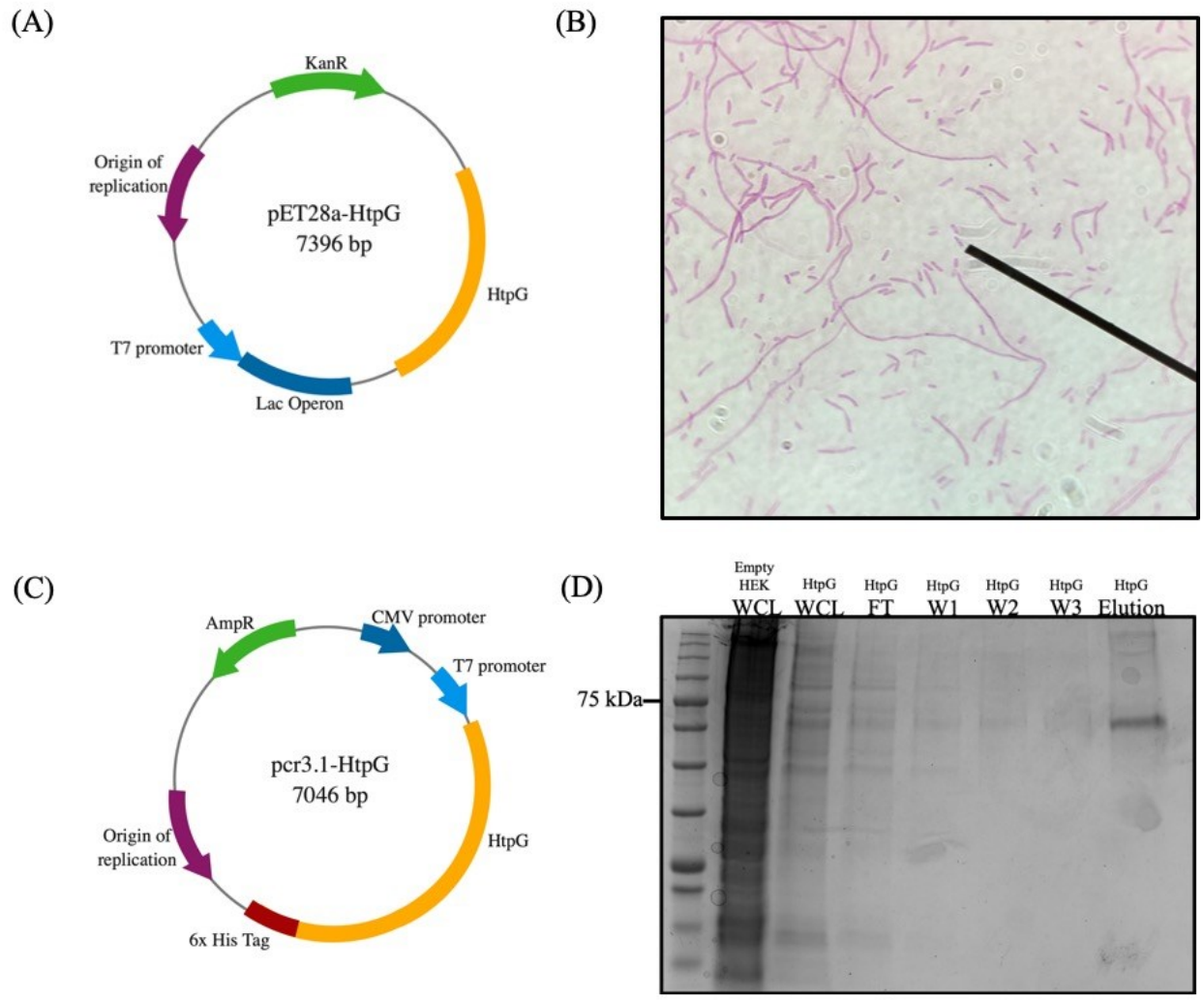


Figure 16. Cloning and purification of group B *B. fragilis* HtpG. (A) Plasmid encoding group B *B. fragilis HtpG* amplified from gDNA. (B) *E. coli* adopted an elongated phenotype following transformation with pET28a-HtpG plasmid. (C) Group B *B. fragilis HtpG* gene block was cloned into the pcr3.1 vector that contains a CMV promoter for mammalian expression and a T7 promoter for bacterial expression. Cloning was confirmed via sequencing. (D) Plasmid was transfected into HEK293T cells. Lysates (WCL) were collected from both un-transfected and transfected HEK293T cells and cobalt bead affinity purified. Flow-through (FT), washes (W1/2/3), and elution were collected. The product was analyzed via SDS-PAGE stained with Coomassie.

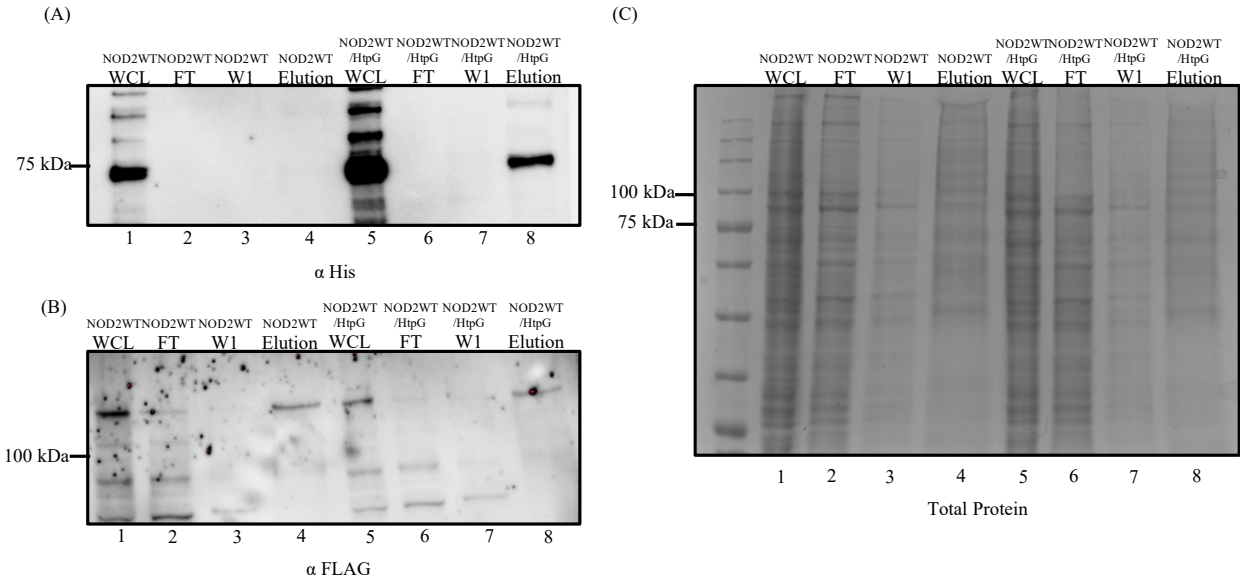


Figure 17. HtpG does not interact with NOD2 WT *in vitro*. HEK293T cells were co-transfected with HtpG and NOD2 WT. Lysates were collected and cobalt bead affinity purified. The product was analyzed via SDS-PAGE probed with (A) α-His antibody, (B) α-FLAG antibody, or (C) total protein. Representative blots of several independent experiments. Mobility of molecular weight standards identified on the left side of each blot in kDa.

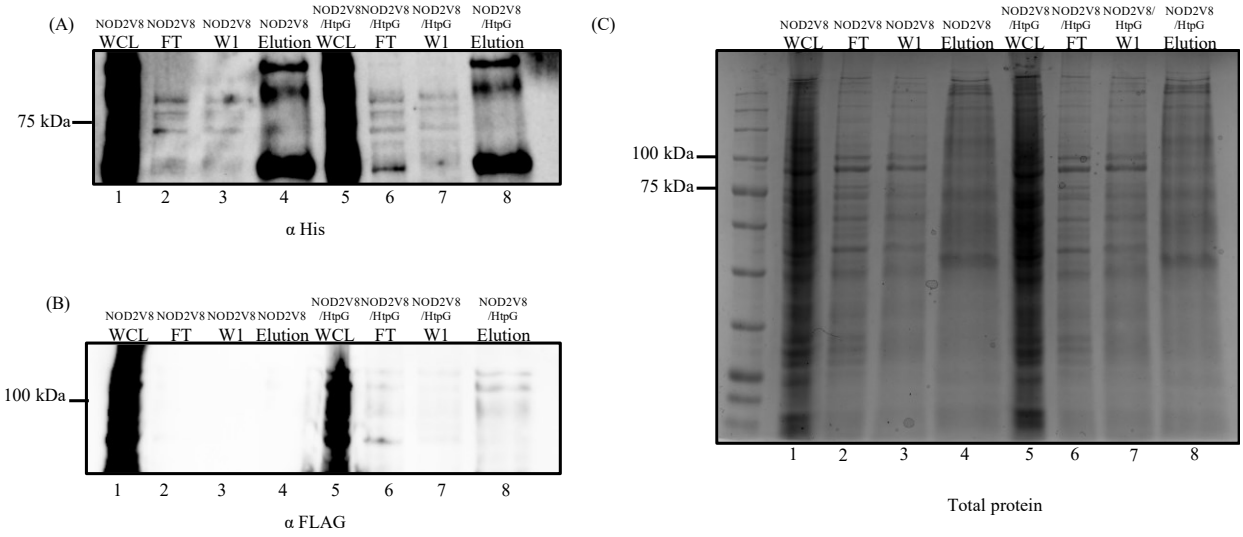


Figure 18. HtpG interacts with NOD2 V8 *in vitro*. HEK293T cells were co-transfected with HtpG and NOD2 V8. Lysates were collected and cobalt bead affinity purified. The product was analyzed via SDS-PAGE probed with (A) α-His antibody, (B) α-FLAG antibody, or (C) total protein using Coomassie stain. Representative blots of several independent experiments. Mobility of molecular weight standards identified on the left side of each blot in kDa.

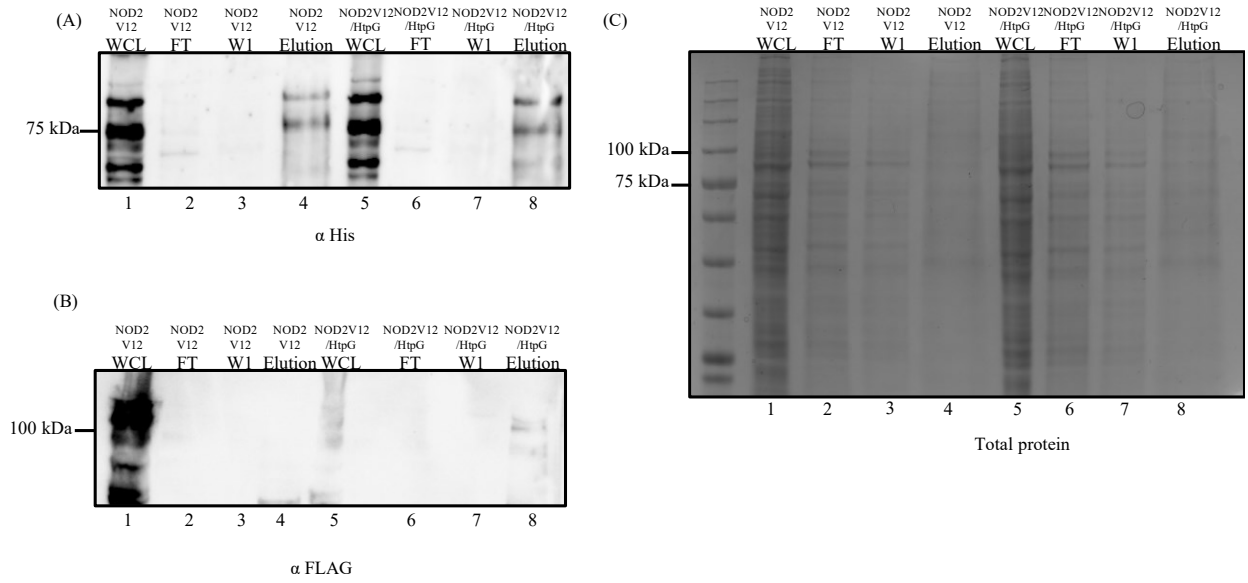


Figure 19. HtpG interacts with NOD2 V12 *in vitro*. HEK293T cells were co-transfected with HtpG and NOD2 V12. Lysates were collected and cobalt bead affinity purified. The product was analyzed via SDS-PAGE probed with (A) α-His antibody, (B) α-FLAG antibody, or (C) total protein using Coomassie stain. Representative blots of several independent experiments. Mobility of molecular weight standards identified on the left side of each blot in kDa.

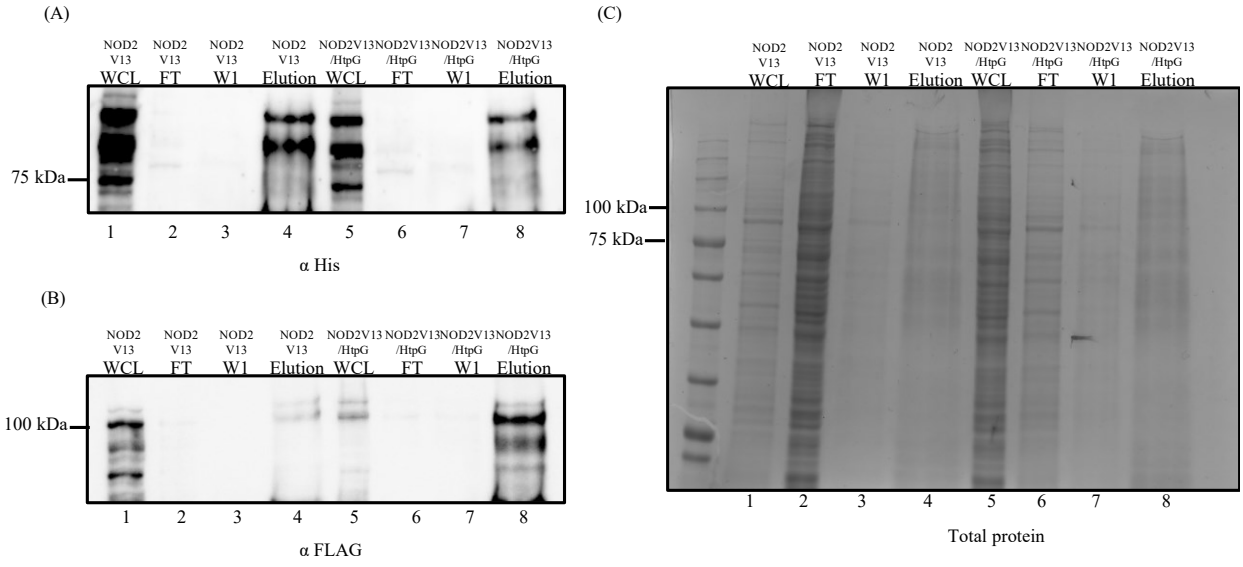


Figure 20. HtpG interacts with NOD2 V13 *in vitro*. HEK293T cells were co-transfected with HtpG and NOD2 V13. Lysates were collected and cobalt bead affinity purified. The product was analyzed via SDS-PAGE probed with (A) α-His antibody, (B) α-FLAG antibody, or (C) total protein using Coomassie stain. Representative blots of several independent experiments. Mobility of molecular weight standards identified on the left side of each blot in kDa.

CHAPTER 4 DISCUSSION

4.1: Summary of Major Findings

This thesis offers important insight into the activity of HtpG in the context of CD. Herein I have described how *HtpG* relative abundance changes with the pediatric CD patient's disease status and I have provided support that HtpG acts directly on intestinal epithelial cells in an anti-inflammatory capacity. Finally, through *in vitro* analysis, I have observed HtpG to interact with three variants of the intracellular PRR, NOD2. To my understanding, this is the first bacterial protein shown to interact with NOD2. Collectively, these results constitute an advancement in our understanding of how and why a patient's HtpG profile may influence their disease state.

4.2: Implications and Relevance of Major Findings

4.2.1: The Intestines Harbour a Unique Relative Abundance of HtpG

Advances in next-generation sequencing have enabled a comprehensive examination of the gut microbiome by providing details on bacterial classification and predicted gene function. The Human Microbiome Project is one of the best known microbiome studies having characterized microbial communities through 16S rRNA gene and metagenomic sequencing from 300 healthy individuals at several body sites. Using this database, I observed that the gut, in comparison to the mouth, nose, skin, and urogenital tract, has a marked increase in the relative abundance of *HtpG* (Figure 4). As a HSP, HtpG is important for regulating the balance between immune activation and suppression. The greater relative abundance of *HtpG* in the gut of healthy individuals may correspond to an equally high concentration of HtpG. If so, HtpG could represent a continuous source of microbial stimuli that enables the maintenance of immune homeostasis and tolerance. Strikingly, this pattern of increased relative abundance in the gut is not observed for another bacterial HSP, *DnaK*. While it might be tempting to assume that the

increase in *HtpG* read counts reflect the unique bacterial composition and density of the gut, it is important to remember that *HtpG* can be found in the genomes of most bacterial genera (Notably, *Neisseria*, *Staphylococcus*, *Lactococcus*, *Streptococcus*, *Lactobacillus*, and *Bifidobacterium* do not contain *HtpG*). Potentially, the increase in *HtpG* relative abundance in the gut is a result of the gut environment itself (e.g., stressors such as pH, oxygen, nutrient deprivation). The relative abundance of *DnaK* may not experience a similar increase in the gut as not all HSPs respond the same way to environmental stressors. Regardless of why, the unique increase of *HtpG* relative abundance in the gut provides impetus to explore the relationship between HtpG and gastrointestinal inflammatory disorders such as CD.

The Dunn et al. (2016) study, using shotgun metagenomic sequencing, observed that healthy individuals have a higher relative abundance of *HtpG* than CD patients, who amongst themselves have a higher relative abundance of *HtpG* if they sustain remission following EEN (Dunn et al., 2016). Upon dissecting the metagenomic dataset further, I observed that a patient's relative abundance of *HtpG* rarely increases or decreases steadily over the course of EEN. Instead, the relative abundance of *HtpG* fluctuates between samples, changing the most once EEN is complete and the individual's regular diet is reintroduced. Diet and dietary restrictions, such as EEN, directly impact microbial composition and metabolite production. As such, the fluctuations in bacterial phyla contributing *HtpG* read counts are not surprising. Just as EEN places metabolic pressure on SCFA-producing bacteria to produce less SCFAs, EEN may also pressure bacteria to either produce higher or lower concentration of HtpG.

Fluctuations in *HtpG* relative abundance following EEN may suggest a role for HtpG in colonization resistance. During and following EEN, commensal and pathogenic bacteria fight to recolonize the intestines through use of various mechanisms including virulence factors, biofilm

formation, and nutrient deprivation. While never directly examined in the context of colonization, several different groups have described HtpG as a virulence factor and have noted that knocking out *HtpG* results in decreased biofilm formation (Garcie et al., 2016; Grudniak, Klecha, & Wolska, 2018; Yamanaka et al., 2009). Additionally, by educating the immune system, HtpG enables bacteria to maintain their colonization status long-term. *Bacteroides,* one of the most abundant commensals within the healthy gut and the main contributor of *HtpG* read counts in the Dunn et al. (2016) dataset, uses capsules and biofilms for persistent colonization (Reis, Silva, Laranjeira, Pinheiro, & Carvalho, 2014; Sproule-Willoughby et al., 2010).

Of the four CD patients I examined in greater detail, patients CD2 and CD3 who sustained remission had lower relative abundances of *HtpG* than patients CD1 and CD5 who did not sustain remission, which seemingly conflicts with the observation made by the Dunn et al. (2016) study. Furthermore, the dominant bacteria contributing the *HtpG* read counts—Bacteroidetes and Firmicutes—were found in both patients who relapsed or sustained remission. These results highlight that CD has a complex etiology and microbial products are just one contributing factor to disease development and status.

4.2.2: HtpG Group Identity Appears Site Specific

Unique to this thesis is the examination of HtpG group identity in the context of health and disease. Over a decade has passed since Chen et al. (2006) outlined three HtpG groups and yet there have been few advances in characterizing HtpG group expression and functionality. By analyzing HtpG amino acid sequences, I observed groups A and B HtpG in CD-patient fecal samples as well as healthy individual fecal and saliva samples. I also observed groups B and C HtpG in healthy individual skin samples. HtpG groups, therefore, appear to be body site specific, which may indicate that there are functional differences between the groups. In general, group A

HtpG is the larger of the three groups. According to the Chen et al. (2006) article, group A HtpG contains other HSP90 family members including HSP90α, HSP90β, and Trap1. For this reason, HtpG from group A may primarily function as chaperones. Several studies describe the protein-stabilizing activity of group A HtpG. For example, *E. coli* group A HtpG interacts and stabilizes DnaA, a protein involved in the initiation of bacterial chromosome DNA replication (Grudniak, Markowska, & Wolska, 2015). In addition, their study demonstrates that the concentration of HtpG is important for interaction to occur—too low a concentration and HtpG cannot interact with DnaA. If this result can be generalized and HtpG concentration is important most interactions, then perhaps in CD patients there is not enough HtpG to interact with the NOD2 variants. Other examples of group A HtpG chaperone activity include interactions with Cas3 in *E. coli*, LexA in *Pseudomonas plecoglossicida*, and TilS in *Shewanella oneidensis* (Honoré, Méjean, & Genest, 2017; Huang et al., 2019; Yosef, Goren, Kiro, Edgar, & Qimron, 2011). In comparison, HtpG from group B and potentially group C may contribute to pathogenesis and colonization. While the number of studies examining group B and C HtpG are limited, group B *P. gingivalis* HtpG is known to induce inflammation in human monocytes (Shelburne, Coopamah, Sweier, An, & Lopatin, 2007). I recognize, however, that functional similarities can occur in the absence of sequence similarities and therefore there may be crossover between the functions of different groups. Moreover, HtpG activity may not be generalizable to the group level and may in fact be species specific.

As a final note, all previous studies on HtpG have failed to report on group identity (in the examples provided in the last paragraph, I determined HtpG group identity), which makes looking for group trends challenging. With the field of HtpG ripe for exploration, the need for

common nomenclature and accurate reporting is required. I propose the use of the system outlined by Chen et al. (2006) – groups A, B, and C—as it is simple and efficient.

4.2.3: HtpG as a Mediator of Intestinal Inflammation

As HtpG can induce inflammatory signaling after application to cell cultures, HtpG released or escaped from bacteria most likely acts through PRRs on the surface of intestinal epithelial cells. Previously, Shelburne et al. (2007) examined the inflammatory properties of HtpG-derived from *P. gingivalis* on human monocytic and microvascular vein endothelial cells. Using CXCL8 production as a measurement of HtpG signaling, with more CXCL8 in this case corresponding to more HtpG, CD91, also known as low-density lipoprotein receptor-related protein 1, and TLR4 were identified as HtpG receptors (Shelburne et al., 2007). Previous work describing HSP signaling via TLR4 is speculated to be confounded by endotoxin contamination; however, Shelburne et al. note that there was no detectable levels of LPS in their final preparation of HtpG (Bausinger et al., 2002). In addition to monocytic and microvascular vein endothelial cells, TLR4 and CD91 are both located on intestinal epithelial cells (Dheer et al., 2016). TLR4 is found on both the apical and basolateral sides of the small and large intestinal epithelium, whereas CD91 localization appears limited to the colon, although there are only a few studies describing cellular placement (Price et al., 2018). While I did not directly examine TLR signaling, I did observe in Figure 15 an increased secretion of IL-1, IL-2, IL-6, IL-12 p40 and p70, CCL2, CCL5, and ICAM-I following rHtpG treatment of HT29 cells. Each of these mediators are downstream of NF-κB activation, a focal point of all TLR signaling pathways (Furrie et al., 2005). As PRRs can recognize multiple bacterial antigens, future work will be important in identifying other receptors used by HtpG. For example, cells lacking MyD88 will lack signalling through most TLRs implicating other receptors important for HtpG.

Within the gut, HtpG presents a paradox whereby a low relative abundance is associated with CD (inflamed state) and a high relative abundance is associated with health. It remains unclear, however, how *HtpG* relative abundance relates to HtpG concentration. *In vitro* analysis shows that rHtpG acts in a concentration dependent manner to influence inflammation (Figure 12A). At a high concentration (50 μg/mL), rHtpG induces HT29 cells to produce inflammatory mediators that recruit macrophages, T-cells, B-cells, NK cells, and dendritic cells (Figure 15). Most notably, rHtpG treatment increases the production of CCL5 and CXCL10, and decreases the production of TIMP2 and CCL15. Several studies have observed increased levels of CCL5, CXCL10, TIMP2, and CCL15 in IBD patients in comparison to healthy controls. While CCL5, CXCL10, and CCL15 contribute to immune cell recruitment, TIMP2, as an endogenous regulator of metalloproteinases, blocks extracellular matrix catabolism. Therefore, the inflammatory profile generated by rHtpG suggests that a high concentration of HtpG not only aids in immune cell recruitment but also encourages the breakdown of the cellular matrix. While high concentrations of HtpG appear pro-inflammatory, perhaps in the context of the gut this level of inflammatory signaling and cellular renewal is part of maintaining tolerance to commensal bacteria (van Eden et al., 2017).

Unlike HtpG, the proinflammatory cytokine TNF presents no paradox with regards to immune activation in CD. Patients with active CD, in comparison to healthy controls, have increased levels of TNF in their serum and stool (Adegbola, Sahnan, Warusavitarne, Hart, & Tozer, 2018). Moreover, within the intestinal mucosa and submucosa of CD patients, TNF secretion from mononuclear cells is increased (Murch, Braegger, Walker-Smith, & MacDonald, 1993). *In vitro* studies have implicated TNF in several pathologic processes including neutrophil accumulation (by activating coagulation responses), granuloma formation (by recruiting T cells

and macrophages), and increased epithelial permeability (by activating metalloproteinases) (Adegbola et al., 2018). Furthermore, blocking TNF by means of anti-TNF monoclonal antibodies is associated with increases in the risk of opportunistic infections. I treated HT29 cells with TNF, either alone or with rHtpG, to observe not only how HT29 cells respond to TNF, but also how HtpG, which appears pro-inflammatory by itself, alters the inflammatory effects of TNF. Strikingly, the rHtpG concentration-dependent CXCL8 curve observed in Figure 12 disappeared when HT29 cells were co-treated with TNF and rHtpG. Only the highest concentration of rHtpG (50 μg/mL) significantly reduced CXCL8 secretion below the level induced by TNF alone (Figure 13). As such, HtpG appears to dampen TNF-induced CXCL8 secretion. The ability of HtpG to dampen TNF-signaling in general is observed in Figure 15 where the mediator panel induced by rHtpG and TNF looks more similar to the panel induced by rHtpG alone than TNF alone. Previous studies have demonstrated a role for HSPs in dampening TNF activity. For example, TNF-induced death was reduced in mice that were heat shocked 12-24 hours before TNF challenge (heat shocking increased HSP70 production in various organs) (Van Molle et al., 2002). The protective effects of heat shock by HSP70 were confirmed by repeating the experiment with HSP70-deficient mice (Van Molle et al., 2002). An exception may be in those patients that raise antibodies to HtpG where TNF-induced inflammation may remain high despite apparent high concentrations of HtpG. Despite the similarities between the inflammatory mediator profiles generated by co-treatment and rHtpG-treatment alone, co-treated cells experienced a slight increase in CCL11 (also known as eotaxin-1) and IL-16 secretion. Eotaxins are small peptides secreted to attract eosinophils and other cell types. Several studies have established a relationship between increased numbers of intestinal eosinophils and IBD pathology (Powell, Walker, & Talley, 2010). Similar to CCL11, IL-16, secreted by several cell

types including eosinophils, is increased in IBD (Seegert et al., 2001). The role of gastrointestinal eosinophils in IBD has yet to be fully elucidated, but studies suggest they may contribute to regulatory, inflammatory, and/or tissue repair functions (Filippone, Sahakian, Apostolopoulos, & Nurgali, 2019). The increased signaling for immune cells such as eosinophils most likely reflects the presence of TNF (the concentration of CCL11 is similar between co-treated cells and TNF-treated cells), which functions to activate granulocytes.

Finally, as a chaperone, HtpG may bind to TNF impacting the ability of TNF to exert its inflammatory effects. To my understanding, no studies have demonstrated such an interaction between HtpG (or any HSP) and TNF; however, interaction between Trap1, a member of the HSP90 family, and the intracellular domain of the type 1 receptor for TNF has been reported (Song, Dunbar, Zhang, Guo, & Donner, 1995). As the use of anti-TNF therapy continues to increase amongst pediatric CD patients, the nature of TNF and HtpG in mediating inflammation should be examined in greater depth (Kaplan et al., 2018).

4.2.4: HtpG Interactions with NOD2

E. coli HtpG shares 50% sequence identity with HSP90 (B. Chen et al., 2006). HtpG, therefore, has the potential to interact with host proteins and receptors that recognize HSP90. Mimicry describes the evolutionary process used by organisms large and small to imitate the appearance and/or function of other organisms or their surrounding environment. Bacteria, primarily pathogenic species, use molecular mimicry to manipulate host cell physiology. Molecular mimics may be bacterial homologs of host proteins or new bacterial effector proteins that arose through convergent evolution. The array of proteins injected into host cells through bacterial secretion systems provide several examples of molecular mimicry. Alternatively, the *Bacteroides* protein BfUbb, with 63% identity to human ubiquitin, is exported through OMVs

(Stewart, D M Edgar, Blakely, & Patrick, 2018). If HtpG is indeed a molecular mimic and interacts with proteins/receptors associated with HSP90, the question remains how HtpG accesses the cytoplasm within intestinal epithelial cells. There are at least three potential mechanisms to explore: bacterial phagocytosis and pinocytosis, outer membrane vesicles, and secretion systems. Within the intestines and despite the mucus barrier, *Bacteroides* have been located not only in association with the colonic mucosa, but also intracellularly (Swidsinski et al., 2002). Furthermore, *B. fragilis* are known to transport HtpG through OMVs (Zakharzhevskaya et al., 2017). A schematic model of HtpG transport through *B. fragilis* OMVs is summarized in Figure 21.

NOD2-mediated inflammatory signaling is tightly regulated through interaction with HSP90. I hypothesized, due to shared sequence identity, that HtpG may also interact with NOD2 following the dissociation of the NOD2-HSP90 complex. Through affinity purification, I examined four separate interactions between HtpG and either NOD2 WT/V8/V12/V13. I observed that 6x-His-HtpG aided in the recovery of Flag-NOD2 variants but not NOD2 WT, as determined by the appearance of a 110 kDa species via FLAG immunoblot. While it remains unclear where HtpG binds to NOD2, it cannot be forgotten that the only differences between NOD2 WT and the variants are mutations in the LRR domain. The observed interactions between HtpG and NOD2 variants are therefore unexpected considering HSP90 interacts with NOD2 CARD domains, not the LRR domain. The LRR domain plays an important role in inhibiting NOD2 self-oligomerization and therefore NOD2 activation. As such, it remains unclear if HtpG binds to the LRR domain when LRRs are folded back inhibiting NOD2 activation, or if HtpG binds once NOD2 is activated. Potentially, the nucleotide substitutions

observed in NOD2 variants, which results in proteins with different charges and polarities than NOD2 WT, make the environment more amenable for HtpG to bind.

Interaction between HtpG and NOD2 variants provides a fascinating new hypothesis for the role of microbial and genetic factors in the development of CD. Figure 22 is a schematic of three possible scenarios linking HtpG and NOD2 WT/variants to intestinal epithelial cell homeostasis: (1) the individual is healthy with NOD2 WT, (2) the individual is healthy with NOD2 variant, and (3) the individual has CD with NOD2 variant. In these scenarios, I am assuming that the relative abundance of *HtpG* as observed in the Dunn et al. (2016) study corresponds to an equal concentration of HtpG. In the first scenario, the individual has functioning NOD2 that is known to play an important role in maintaining tolerance. Furthermore, the individual has a high concentration of HtpG in the lumen that may act through PRR to educate the immune system. In the second scenario, the NOD2 variant means that the individual has a loss of tolerance due to dampened MDP sensing, but the high concentration of HtpG may be able to maintain some level of immune education. In the final scenario, the individual not only has reduced MDP sensing and loss of tolerance due to NOD2 variant, but also they have a reduced concentration of HtpG. As NOD2 variants are hypo-responsive to MDP and as HtpG appears to interact with the LRR domain that senses ligands, HtpG binding may further reduce MDP sensing. The reduction of MDP sensing in combination with the reduction in HtpG concentration, as well as other factors not examined in this thesis, may lead to the chronic inflammation observed in patients with CD.

4.3: Limitations of Experimental Systems

4.3.1: Thesis Foundation Based on Metagenomic Sequencing Data from One Cohort of Patients

The foundation of this thesis—that HtpG occurs in different concentrations depending on health/CD status—is based on the observation made from metagenomic sequencing data collected during the Dunn et al. (2016) study. While metagenomic sequencing is a major advance from 16S rRNA gene sequencing, it does not provide information on cell-specific gene expression features or the proteins inside a cell. As such, the *HtpG* relative abundance pattern observed in the Dunn et al. (2016) study may not necessarily translate into a similar protein concentration pattern when examined using proteomics. Furthermore, I recognize that by selecting samples from the Dunn et al. (2016) study I risk having an underpowered study and may not come to same conclusions as the Dunn et al. study. I realize I may also risk a lack of reproducibility. In addition, I do not know the genotypes of the Dunn et al. study participants. Despite the shortcomings in my sample of the Dunn et al. study, preliminary exploration of the Human Microbiome Project 2 online database showed detectable amounts of HtpG within the gut. Continued analysis of the Human Microbiome Project 2 online database will help to better characterize HtpG in IBD individuals. Additionally, analysis of other pediatric CD cohorts where children receive EEN will determine whether changes to HtpG occur consistently and whether these changes are driven by the same key bacteria.

4.3.2: Profile of Inflammatory Mediators Generated by HT29 Cells Following rHtpG, TNF, or Combined Treatment

The inflammatory mediator array provided great insight into the types of mediators induced following treatment with either rHtpG, TNF, or rHtpG and TNF. Without the array, I would not have observed the ability of HtpG to influence the inflammatory profile when co-administered with TNF, a result that was also observed using the CXCL8 ELISA. Nevertheless,

the array had three main limitations. Firstly, the array was qualitative not quantitative. As such, the observed changes to protein secretion, both large and small, should be quantified using a more sensitive analysis, such as an ELISA.

Secondly, HT29 cells are a transformed cell line and as such may have characteristics not found in primary colonic epithelial cells. For example, TIMP2 secretion is known to be increased in transformed cells and I observed a large concentration of TIMP2 in my untreated cells (in comparison to the positive controls). That being said, studies have shown that HT29 cells and freshly isolated intestinal epithelial cells similarly secrete other inflammatory mediators, such as CXCL8, CCL2, GM-CSF, and TNF (Jung et al., 1995).

Finally, while many of the inflammatory mediators examined are known to be secreted by intestinal epithelial cells, some have not previously been described. For example, IL-12 p70 (heterodimer composed of a p35 and p40 subunit) is naturally produced by immune cells such as dendritic cells and macrophages. While research describing IL-12 p70 and p40 secretion from intestinal epithelial cells remains unclear, a study did observe regulated expression of IL-12 p35 in HT29 cells (Maaser, Egan, Birkenbach, Eckmann, & Kagnoff, 2004). Moreover, a study using oviduct cells (also a mucosal epithelium) observed transcription of both IL-12 p35 and p40 subunits as well as secretion of IL-12 p70 into culture supernatant (R. M. Johnson, 2004). Therefore, it is feasible that HT29 cells may also be able to secrete IL-12 p70 and p40 as observed in Figure 15. There were three other inflammatory mediators that were surprising, IFNγ, IL-2, and IL-16, as all previous reports suggest that they are produced only by immune cells. Detection of these mediators may be a result of poor antibody specificity or transformed cancer cell signaling.

4.3.3: Assumption that HtpG Can Get Into Intestinal Epithelial Cells

To examine a possible interaction between HtpG and NOD2, I transfected HEK293T cells with my genes of interest, and in by doing so, I created an artificial environment. One of the major assumptions of this environment is that HtpG can get into intestinal epithelial cells. While it was necessary to reduce the system to begin investigations on potential intracellular HtpG interactions, I am most likely missing important interactions between bacteria and host cells that may represent routes of HtpG exposure. Furthermore, I recognize that the observed interaction between HtpG and NOD2 variants may be due to the "sticky-ness" of this artificial environment. In my Western blots there were several additional protein species, which suggests that HtpG, as a chaperone, may bind to lots of cellular proteins. Future purifications may look to increase imidazole concentration or add some detergent, such as SDS, to reduce non-specific binding.

4.4: Proposed Future Directions

4.4.1: HtpG-Induced Inflammation in Intestinal Epithelial Cells, Macrophages, and Organoid Models

In this study, I examined the inflammatory effects of HtpG on the colonic colorectal cancer cell line HT29 using a membrane array and an ELISA. Due to the variability in cell lines, the results generated from these experiments need to be confirmed in at least two other intestinal cells lines, such as the other colorectal cancer cell lines T84 and Caco-2. In addition, HtpG inflammatory activity should also be characterized using macrophage and dendritic cell lines.

Intestinal epithelial cells are suitable for initial experiments, however, they do not accurately represent the complexity of the intestinal epithelium (i.e., multiple cell subtypes and cell polarization). Intestinal organoids (also called "enteroids") are 3D primary epithelial cell constructs that incorporate many physiologically relevant features of the intestinal epithelium (N. de Souza, 2018). As such, it would be interesting to use enteroids to examine the impact of HtpG

on cell subtypes found *in vivo*. The enteroids could initially come from mice, but experiments could be validated using enteroids derived from human biopsies. Based on the results from the inflammatory mediator array, I propose future experiments begin by analyzing CCL5, CXCL10, TIMP2, and CCL15. In addition to analysis of inflammatory mediators through ELISA assays, changes in genes expression should be assessed through transcriptomic profiling in both the various cell lines and enteroids. Analysis of gene expression in the enteroids will provide fascinating insight into how specific intestinal cell subtypes (sorted using flow cytometry) respond to HtpG.

In the extra-cellular experiments conducted for this thesis, I used a purchased recombinant *B. fragilis* group B HtpG, not my homemade HtpG. In a preliminary endotoxin screen of my homemade *B. fragilis* group B HtpG by Mushfiqur Rahman, endotoxin levels were greater than 100 pg *E. coli* LPS per µg of recombinant protein. Endotoxin levels in the future will need to be decreased by affinity chromatography on an agarose gel containing Polymyxin B, an inhibitor of biological activities of LPS. The results generated previously should be confirmed using homemade HtpG as full length HtpG may behave differently than truncated HtpG.

4.4.2: HtpG-NOD2 Interaction Confirmed Using Alternative Methods

The interaction between HtpG and NOD2 variants provides an interesting avenue to explore in terms of HtpG intracellular activity; however, the interaction must first be confirmed using alternative methods. Other techniques I considered, but ultimately did not have time to pursue included co-immunoprecipitation using anti-FLAG antibodies, mass spectrometry, and co-localization using fluorescently-labelled HtpG and NOD2. If HtpG interacts with NOD2 variants, then the mechanism of action and resulting cellular effects will need to be elucidated. If the interactions between HtpG and NOD2 variants are disproven then a broader approach

should be taken to examine other potential HtpG-protein interactions. A yeast-2-hybrid could be used to conduct the initial screen with HtpG acting as bait.

4.4.3: HtpG and Colorectal Cancer (CRC)

IBD patients having a 2-10 fold increase in developing CRC (Bhatt, Redinbo, & Bultman, 2017). Many of the same factors that contribute to CD development, such as chronic inflammation, *NOD2* mutations, and intestinal dysbiosis, also contribute to CRC development (Irrazábal, Belcheva, Girardin, Martin, & Philpott, 2014; Lasry, Zinger, & Ben-Neriah, 2016). Recent studies suggest that the instability/poor signaling of NOD2 mutants could be corrected by enhancing the activity of chaperone proteins that stabilize NOD2, including HSP90. However, in cancer cells, Hsp90 plays a critical role in stabilizing oncoproteins and inhibiting apoptotic activity. As the bacterial homolog of HSP90, the role of HtpG in CRC remains unknown. Elevated levels of reactive oxygen species are a feature of almost all cancers. A previous report showed that loss of HtpG in the bacterium *Edwardsiella tarda* made the cells more sensitive to stress induced by reactive oxygen species (Dang, Hu, & Sun, 2011). Future experiments could focus on exploring the relationship between CRC and HtpG by better describing the effects of HtpG on reactive oxygen species. Controlling HtpG abundance and antagonizing Hsp90 cancer-promoting activity may provide a personalized strategy to reduce the risk of CRC in IBD patients.

4.5: Concluding Remarks

Our understanding of commensal bacteria in influencing host immunity and physiology has evolved rapidly over the past several years. This work illustrates how translational metagenome analysis of smaller cohorts with matching of time points, phenotypes, and treatment exposures, could inform modification of bacterial pathways (e.g., increasing abundance of the

bacterial protein HtpG) to reduce mucosal inflammation. As the prevalence of CD, especially among the pediatric population, continues to rise globally, a better understanding of HtpG could be used to inform new or improve current CD treatments.

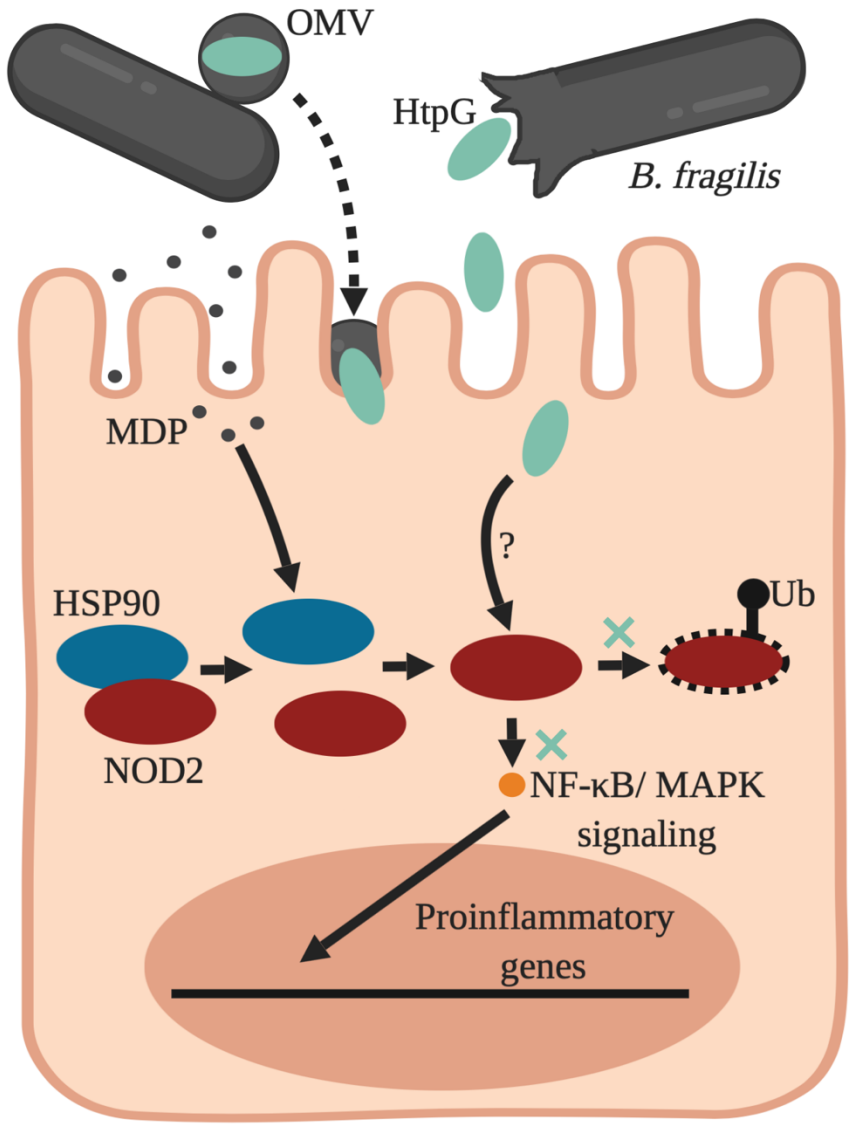


Figure 21. Schematic for *B. fragilis* group B HtpG intracellular activity. In intestinal epithelial cells, NOD2 is constitutively associated with HSP90 unless bacterial MDP binds. NOD2 then signals through NF-kB and MAPK pathways to activate proinflammatory genes, following which NOD2 is degraded. I propose that *B. fragilis* HtpG enters intestinal epithelial cells through potential mechanisms, such as outer membrane vesicles (OMVs) or phagocytosis. Then, based on my results, HtpG interacts with NOD2 variants. This interaction may impact downstream signaling processes and perhaps even the fate of NOD2. Figure created with BioRender.

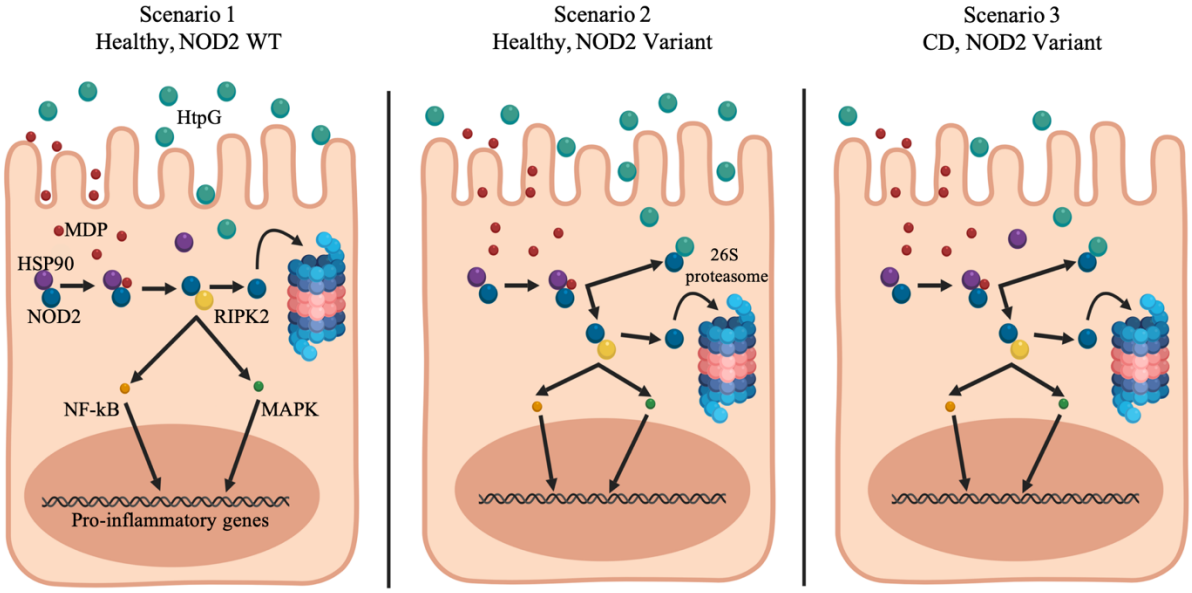


Figure 22. Schematic showing the involvement of HtpG in intestinal epithelial cell homeostasis in healthy and CD individuals. In scenario 1, a higher concentration of HtpG in the lumen and intact NOD2 signaling gives rise to tolerance. In scenario 2, despite the reduced ability to sense MDP and therefore a loss of tolerance, the high concentration of HtpG in the lumen enables some immune cell education through interaction with PRR. In scenario 3, the reduced ability of NOD2 signaling and the low concentration of HtpG results in inflammation. Figure created with BioRender.

References

Acevedo, N., Reinius, L. E., Vitezic, M., Fortino, V., Söderhäll, C., Honkanen, H., … Kere, J. (2015). Age-associated DNA methylation changes in immune genes, histone modifiers and chromatin remodeling factors within 5 years after birth in human blood leukocytes. *Clinical Epigenetics*, *7*(1), 34. https://doi.org/10.1186/s13148-015-0064-6

Adegbola, S. O., Sahnan, K., Warusavitarne, J., Hart, A., & Tozer, P. (2018). Anti-TNF Therapy in Crohn's Disease. *International Journal of Molecular Sciences*, *19*(8). https://doi.org/10.3390/ijms19082244

Agace, W. W., & McCoy, K. D. (2017). Regionalized Development and Maintenance of the Intestinal Adaptive Immune Landscape. *Immunity*, *46*(4), 532–548. https://doi.org/10.1016/j.immuni.2017.04.004

Ahern, P. P., Faith, J. J., & Gordon, J. I. (2014). Mining the human gut microbiota for effector strains that shape the immune system. *Immunity*, *40*(6), 815–823. https://doi.org/10.1016/j.immuni.2014.05.012

Ajendra, J., Specht, S., Ziewer, S., Schiefer, A., Pfarr, K., Parčina, M., … Hübner, M. P. (2016). NOD2 dependent neutrophil recruitment is required for early protective immune responses against infectious Litomosoides sigmodontis L3 larvae. *Scientific Reports*, *6*, 39648. https://doi.org/10.1038/srep39648

Al Nabhani, Z., Dietrich, G., Hugot, J.-P., & Barreau, F. (2017). Nod2: The intestinal gate keeper. *PLOS Pathogens*, *13*(3), e1006177. https://doi.org/10.1371/journal.ppat.1006177

Alhagamhmad, M. H., Day, A. S., Lemberg, D. A., & Leach, S. T. (2016). An overview of the bacterial contribution to Crohn disease pathogenesis. *Journal of Medical Microbiology*, *65*(10), 1049–1059. https://doi.org/10.1099/jmm.0.000331

Allaire, J. M., Crowley, S. M., Law, H. T., Chang, S.-Y., Ko, H.-J., & Vallance, B. A. (2018). The Intestinal Epithelium: Central Coordinator of Mucosal Immunity. *Trends in Immunology*, *39*(9), 677–696. https://doi.org/10.1016/j.it.2018.04.002

Amarante-Mendes, G. P., Adjemian, S., Branco, L. M., Zanetti, L. C., Weinlich, R., & Bortoluci, K. R. (2018). Pattern Recognition Receptors and the Host Cell Death Molecular Machinery. *Frontiers in Immunology*, *9*, 2379. https://doi.org/10.3389/fimmu.2018.02379

Ananthakrishnan, A. N., Khalili, H., Higuchi, L. M., Bao, Y., Korzenik, J. R., Giovannucci, E. L., … Chan, A. T. (2012). Higher predicted vitamin D status is associated with reduced risk of Crohn's disease. *Gastroenterology*, *142*(3), 482–489. https://doi.org/10.1053/j.gastro.2011.11.040

Ananthakrishnan, A. N., Khalili, H., Konijeti, G. G., Higuchi, L. M., de Silva, P., Korzenik, J. R., … Chan, A. T. (2013). A prospective study of long-term intake of dietary fiber and risk of Crohn's disease and ulcerative colitis. *Gastroenterology*, *145*(5), 970–977. https://doi.org/10.1053/j.gastro.2013.07.050

Ashida, H., Ogawa, M., Kim, M., Mimuro, H., & Sasakawa, C. (2011). Bacteria and host interactions in the gut epithelial barrier. *Nature Chemical Biology*, *8*(1), 36–45. https://doi.org/10.1038/nchembio.741

Baird, C. H., Niederlechner, S., Beck, R., Kallweit, A. R., & Wischmeyer, P. E. (2013). L-Threonine induces heat shock protein expression and decreases apoptosis in heat-stressed intestinal epithelial cells. *Nutrition (Burbank, Los Angeles County, Calif.)*, *29*(11–12), 1404–1411. https://doi.org/10.1016/j.nut.2013.05.017

Baker, P. J., De Nardo, D., Moghaddas, F., Tran, L. S., Bachem, A., Nguyen, T., … Masters, S. L. (2017). Posttranslational Modification as a Critical Determinant of Cytoplasmic Innate Immune Recognition. *Physiological Reviews*, *97*(3), 1165–1209. https://doi.org/10.1152/physrev.00026.2016

Barclay, A. R., Russell, R. K., Wilson, M. L., Gilmour, W. H., Satsangi, J., & Wilson, D. C. (2009). Systematic Review: The Role of Breastfeeding in the Development of Pediatric Inflammatory Bowel Disease. *The Journal of Pediatrics*, *155*(3), 421–426. https://doi.org/10.1016/j.jpeds.2009.03.017

Barnich, N., Aguirre, J. E., Reinecker, H.-C., Xavier, R., & Podolsky, D. K. (2005). Membrane recruitment of NOD2 in intestinal epithelial cells is essential for nuclear factor-{kappa}B activation in muramyl dipeptide recognition. *The Journal of Cell Biology*, *170*(1), 21–26. https://doi.org/10.1083/jcb.200502153

Bausinger, H., Lipsker, D., Ziylan, U., Manié, S., Briand, J.-P., Cazenave, J.-P., … Hanau, D. (2002). Endotoxin-free heat-shock protein 70 fails to induce APC activation. *European Journal of Immunology*, *32*(12), 3708–3713. https://doi.org/10.1002/1521-4141(200212)32:12<3708::AID-IMMU3708>3.0.CO;2-C

Benchimol, E. I., Bernstein, C. N., Bitton, A., Carroll, M. W., Singh, H., Otley, A. R., … Kaplan, G. G. (2017). Trends in Epidemiology of Pediatric Inflammatory Bowel Disease in Canada: Distributed Network Analysis of Multiple Population-Based Provincial Health Administrative Databases: *American Journal of Gastroenterology*, *112*(7), 1120–1134. https://doi.org/10.1038/ajg.2017.97

Benchimol, E. I., Kaplan, G. G., Otley, A. R., Nguyen, G. C., Underwood, F. E., Guttmann, A., … Bernstein, C. N. (2017). Rural and Urban Residence During Early Life is Associated with Risk of Inflammatory Bowel Disease: A Population-Based Inception and Birth Cohort Study: *American Journal of Gastroenterology*, *112*(9), 1412–1422. https://doi.org/10.1038/ajg.2017.208

Benchimol, E. I., Mack, D. R., Guttmann, A., Nguyen, G. C., To, T., Mojaverian, N., … Manuel, D. G. (2015). Inflammatory Bowel Disease in Immigrants to Canada And Their Children: A Population-Based Cohort Study: *American Journal of Gastroenterology*, *110*(4), 553–563. https://doi.org/10.1038/ajg.2015.52

Bevins, C. L., & Salzman, N. H. (2011). Paneth cells, antimicrobial peptides and maintenance of intestinal homeostasis. *Nature Reviews. Microbiology*, *9*(5), 356–368. https://doi.org/10.1038/nrmicro2546

Bhatt, A. P., Redinbo, M. R., & Bultman, S. J. (2017). The role of the microbiome in cancer development and therapy. *CA: A Cancer Journal for Clinicians*, *67*(4), 326–344. https://doi.org/10.3322/caac.21398

Binder, R. J. (2014). Functions of heat shock proteins in pathways of the innate and adaptive immune system. *Journal of Immunology (Baltimore, Md.: 1950)*, *193*(12), 5765–5771. https://doi.org/10.4049/jimmunol.1401417

Biswas, A., & Kobayashi, K. S. (2013). Regulation of intestinal microbiota by the NLR protein family. *International Immunology*, *25*(4), 207–214. https://doi.org/10.1093/intimm/dxs116

Borrelli, O., Cordischi, L., Cirulli, M., Paganelli, M., Labalestra, V., Uccini, S., … Cucchiara, S. (2006). Polymeric Diet Alone Versus Corticosteroids in the Treatment of Active Pediatric Crohn's Disease: A Randomized Controlled Open-Label Trial. *Clinical Gastroenterology and Hepatology*, *4*(6), 744–753. https://doi.org/10.1016/j.cgh.2006.03.010

Boyko, E. J., Theis, M. K., Vaughan, T. L., & Nicol-Blades, B. (1994). Increased Risk of Inflammatory Bowel Disease Associated with Oral Contraceptive Use. *American Journal of Epidemiology*, *140*(3), 268–278. https://doi.org/10.1093/oxfordjournals.aje.a117246

Boyle, J. P., Parkhouse, R., & Monie, T. P. (2014). Insights into the molecular basis of the NOD2 signalling pathway. *Open Biology*, *4*(12). https://doi.org/10.1098/rsob.140178

Brant, S. R. (2011). Update on the heritability of inflammatory bowel disease: The importance of twin studies: *Inflammatory Bowel Diseases*, *17*(1), 1–5. https://doi.org/10.1002/ibd.21385

Broere, F., van der Zee, R., & van Eden, W. (2011). Heat shock proteins are no DAMPs, rather "DAMPERs." *Nature Reviews. Immunology*, *11*(8), 565; author reply 565. https://doi.org/10.1038/nri2873-c1

Carroll, M. W., Hamilton, Z., Gill, H., Simkin, J., Smyth, M., Espinosa, V., … Jacobson, K. (2016). Pediatric Inflammatory Bowel Disease Among South Asians Living in British Columbia, Canada: A Distinct Clinical Phenotype. *Inflammatory Bowel Diseases*, *22*(2), 387–396. https://doi.org/10.1097/MIB.0000000000000651

Caruso, R., Warner, N., Inohara, N., & Núñez, G. (2014). NOD1 and NOD2: Signaling, host defense, and inflammatory disease. *Immunity*, *41*(6), 898–908. https://doi.org/10.1016/j.immuni.2014.12.010

Chassaing, B., Kumar, M., Baker, M. T., Singh, V., & Vijay-Kumar, M. (2014). Mammalian gut immunity. *Biomedical Journal*, *37*(5), 246–258. https://doi.org/10.4103/2319-4170.130922

Chen, B., Zhong, D., & Monteiro, A. (2006). Comparative genomics and evolution of the HSP90 family of genes across all kingdoms of organisms. *BMC Genomics*, *7*, 156. https://doi.org/10.1186/1471-2164-7-156

Chen, J.-S., Hu, F., Kugathasan, S., Jorde, L. B., Nix, D., Rutherford, A., … Guthery, S. L. (2018). Targeted Gene Sequencing in Children with Crohn's Disease and Their Parents: Implications for Missing Heritability. *G3: Genes|Genomes|Genetics*, *8*(9), 2881–2888. https://doi.org/10.1534/g3.118.200404

Cholapranee, A., & Ananthakrishnan, A. N. (2016). Environmental Hygiene and Risk of Inflammatory Bowel Diseases: A Systematic Review and Meta-analysis. *Inflammatory Bowel Diseases*, *22*(9), 2191–2199. https://doi.org/10.1097/MIB.0000000000000852

Clark, N. M., Marinis, J. M., Cobb, B. A., & Abbott, D. W. (2008). MEKK4 sequesters RIP2 to dictate NOD2 signal specificity. *Current Biology: CB*, *18*(18), 1402–1408. https://doi.org/10.1016/j.cub.2008.07.084

Crosnier, C., Stamataki, D., & Lewis, J. (2006). Organizing cell renewal in the intestine: Stem cells, signals and combinatorial control. *Nature Reviews. Genetics*, *7*(5), 349–359. https://doi.org/10.1038/nrg1840

Dalal, S. R., & Chang, E. B. (2014). The microbial basis of inflammatory bowel diseases. *The Journal of Clinical Investigation*, *124*(10), 4190–4196. https://doi.org/10.1172/JCI72330

Dang, W., Hu, Y., & Sun, L. (2011). HtpG is involved in the pathogenesis of Edwardsiella tarda. *Veterinary Microbiology*, *152*(3–4), 394–400. https://doi.org/10.1016/j.vetmic.2011.05.030

D'Aoust, J., Battat, R., & Bessissow, T. (2017). Management of inflammatory bowel disease with Clostridium difficile infection. *World Journal of Gastroenterology*, *23*(27), 4986–5003. https://doi.org/10.3748/wjg.v23.i27.4986

Darfeuille-Michaud, A., Boudeau, J., Bulois, P., Neut, C., Glasser, A.-L., Barnich, N., … Colombel, J.-F. (2004). High prevalence of adherent-invasive Escherichia coli associated with ileal mucosa in Crohn's disease. *Gastroenterology*, *127*(2), 412–421.

David, L. A., Maurice, C. F., Carmody, R. N., Gootenberg, D. B., Button, J. E., Wolfe, B. E., … Turnbaugh, P. J. (2014). Diet rapidly and reproducibly alters the human gut microbiome. *Nature*, *505*(7484), 559–563. https://doi.org/10.1038/nature12820

de Bruyn, M., & Vermeire, S. (2017). NOD2 and bacterial recognition as therapeutic targets for Crohn's disease. *Expert Opinion on Therapeutic Targets*, *21*(12), 1123–1139. https://doi.org/10.1080/14728222.2017.1397627

de Moura, C. S., Lollo, P. C. B., Morato, P. N., Carneiro, E. M., & Amaya-Farfan, J. (2013). Whey protein hydrolysate enhances the exercise-induced heat shock protein (HSP70) response in rats. *Food Chemistry*, *136*(3–4), 1350–1357. https://doi.org/10.1016/j.foodchem.2012.09.070

de Souza, H. S. P., & Fiocchi, C. (2016). Immunopathogenesis of IBD: Current state of the art. *Nature Reviews. Gastroenterology & Hepatology*, *13*(1), 13–27. https://doi.org/10.1038/nrgastro.2015.186

de Souza, H. S. P., Fiocchi, C., & Iliopoulos, D. (2017). The IBD interactome: An integrated view of aetiology, pathogenesis and therapy. *Nature Reviews Gastroenterology & Hepatology*, *14*(12), 739–749. https://doi.org/10.1038/nrgastro.2017.110

de Souza, N. (2018). Organoids. *Nature Methods*, *15*(1), 23–23. https://doi.org/10.1038/nmeth.4576

Dheer, R., Santaolalla, R., Davies, J. M., Lang, J. K., Phillips, M. C., Pastorini, C., … Abreu, M. T. (2016). Intestinal Epithelial Toll-Like Receptor 4 Signaling Affects Epithelial Function and Colonic Microbiota and Promotes a Risk for Transmissible Colitis. *Infection and Immunity*, *84*(3), 798–810. https://doi.org/10.1128/IAI.01374-15

Diaz Heijtz, R., Wang, S., Anuar, F., Qian, Y., Björkholm, B., Samuelsson, A., … Pettersson, S. (2011). Normal gut microbiota modulates brain development and behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(7), 3047–3052. https://doi.org/10.1073/pnas.1010529108

Donaldson, G. P., Lee, S. M., & Mazmanian, S. K. (2016). Gut biogeography of the bacterial microbiota. *Nature Reviews. Microbiology*, *14*(1), 20–32. https://doi.org/10.1038/nrmicro3552

Dotan, I., Fishman, S., Dgani, Y., Schwartz, M., Karban, A., Lerner, A., … Halpern, Z. (2006). Antibodies against laminaribioside and chitobioside are novel serologic markers in Crohn's disease. *Gastroenterology*, *131*(2), 366–378. https://doi.org/10.1053/j.gastro.2006.04.030

Dunn, K. A., Moore-Connors, J., MacIntyre, B., Stadnyk, A., Thomas, N. A., Noble, A., … Van Limbergen, J. (2016). The Gut Microbiome of Pediatric Crohn's Disease Patients Differs from Healthy Controls in Genes That Can Influence the Balance Between a Healthy and Dysregulated Immune Response: *Inflammatory Bowel Diseases*, *22*(11), 2607–2618. https://doi.org/10.1097/MIB.0000000000000949

Economou, M., Trikalinos, T. A., Loizou, K. T., Tsianos, E. V., & Ioannidis, J. P. A. (2004). Differential effects of NOD2 variants on Crohn's disease risk and phenotype in diverse populations: A metaanalysis. *The American Journal of Gastroenterology*, *99*(12), 2393–2404. https://doi.org/10.1111/j.1572-0241.2004.40304.x

Feerick, C. L., & McKernan, D. P. (2017). Understanding the regulation of pattern recognition receptors in inflammatory diseases - a 'Nod' in the right direction. *Immunology*, *150*(3), 237–247. https://doi.org/10.1111/imm.12677

Ferrand, A., Al Nabhani, Z., Tapias, N. S., Mas, E., Hugot, J.-P., & Barreau, F. (2019). NOD2 Expression in Intestinal Epithelial Cells Protects Toward the Development of Inflammation and Associated Carcinogenesis. *Cellular and Molecular Gastroenterology and Hepatology*, *7*(2), 357–369. https://doi.org/10.1016/j.jcmgh.2018.10.009

Filippone, R. T., Sahakian, L., Apostolopoulos, V., & Nurgali, K. (2019). Eosinophils in Inflammatory Bowel Disease. *Inflammatory Bowel Diseases*. https://doi.org/10.1093/ibd/izz024

Finlayson-Trick, E., Connors, J., Stadnyk, A., & Van Limbergen, J. (2018). Regulation of Antimicrobial Pathways by Endogenous Heat Shock Proteins in Gastrointestinal Disorders. *Gastrointestinal Disorders*, *1*(1), 39–56. https://doi.org/10.3390/gidisord1010005

Forbes, A., Escher, J., Hébuterne, X., Kłęk, S., Krznaric, Z., Schneider, S., … Bischoff, S. C. (2017). ESPEN guideline: Clinical nutrition in inflammatory bowel disease. *Clinical Nutrition*, *36*(2), 321–347. https://doi.org/10.1016/j.clnu.2016.12.027

Frank, D. N., Robertson, C. E., Hamm, C. M., Kpadeh, Z., Zhang, T., Chen, H., … Li, E. (2011). Disease phenotype and genotype are associated with shifts in intestinal-associated microbiota in inflammatory bowel diseases. *Inflammatory Bowel Diseases*, *17*(1), 179–184. https://doi.org/10.1002/ibd.21339

Freeman, H. J. (2002). Familial Crohn's Disease in Single or Multiple First-Degree Relatives: *Journal of Clinical Gastroenterology*, *35*(1), 9–13. https://doi.org/10.1097/00004836-200207000-00004

Frolkis, A., Dieleman, L. A., Barkema, H. W., Panaccione, R., Ghosh, S., Fedorak, R. N., … Alberta IBD Consortium. (2013). Environment and the inflammatory bowel diseases. *Canadian Journal of Gastroenterology = Journal Canadien De Gastroenterologie*, *27*(3), e18-24.

Fukata, M., & Arditi, M. (2013). The role of pattern recognition receptors in intestinal inflammation. *Mucosal Immunology*, *6*(3), 451–463. https://doi.org/10.1038/mi.2013.13

Furrie, E., Macfarlane, S., Thomson, G., Macfarlane, G. T., Microbiology & Gut Biology Group, & Tayside Tissue & Tumour Bank. (2005). Toll-like receptors-2, -3 and -4 expression patterns on human colon and their regulation by mucosal-associated bacteria. *Immunology*, *115*(4), 565–574. https://doi.org/10.1111/j.1365-2567.2005.02200.x

Garcie, C., Tronnet, S., Garénaux, A., McCarthy, A. J., Brachmann, A. O., Pénary, M., … Martin, P. (2016). The Bacterial Stress-Responsive Hsp90 Chaperone (HtpG) Is Required for the Production of the Genotoxin Colibactin and the Siderophore Yersiniabactin in *Escherichia coli*. *Journal of Infectious Diseases*, *214*(6), 916–924. https://doi.org/10.1093/infdis/jiw294

Gatti, S., Galeazzi, T., Franceschini, E., Annibali, R., Albano, V., Verma, A., … Catassi, C. (2017). Effects of the Exclusive Enteral Nutrition on the Microbiota Profile of Patients with Crohn's Disease: A Systematic Review. *Nutrients*, *9*(8), 832. https://doi.org/10.3390/nu9080832

Gerasimidis, K., Bertz, M., Hanske, L., Junick, J., Biskou, O., Aguilera, M., … Edwards, C. A. (2014). Decline in Presumptively Protective Gut Bacterial Species and Metabolites Are Paradoxically Associated with Disease Improvement in Pediatric Crohn's Disease During Enteral Nutrition: *Inflammatory Bowel Diseases*, *20*(5), 861–871. https://doi.org/10.1097/MIB.0000000000000023

Geremia, A., & Arancibia-Cárcamo, C. V. (2017). Innate Lymphoid Cells in Intestinal Inflammation. *Frontiers in Immunology*, *8*, 1296. https://doi.org/10.3389/fimmu.2017.01296

Gevers, D., Kugathasan, S., Denson, L. A., Vázquez-Baeza, Y., Van Treuren, W., Ren, B., … Xavier, R. J. (2014). The Treatment-Naive Microbiome in New-Onset Crohn's Disease. *Cell Host & Microbe*, *15*(3), 382–392. https://doi.org/10.1016/j.chom.2014.02.005

Gibson, F. C., Onderdonk, A. B., Kasper, D. L., & Tzianabos, A. O. (1998). Cellular mechanism of intraabdominal abscess formation by Bacteroides fragilis. *Journal of Immunology (Baltimore, Md.: 1950)*, *160*(10), 5000–5006.

Gilbert, J. A., Blaser, M. J., Caporaso, J. G., Jansson, J. K., Lynch, S. V., & Knight, R. (2018). Current understanding of the human microbiome. *Nature Medicine*, *24*(4), 392–400. https://doi.org/10.1038/nm.4517

Grimes, C. L., Ariyananda, L. D. Z., Melnyk, J. E., & O'Shea, E. K. (2012). The Innate Immune Protein Nod2 Binds Directly to MDP, a Bacterial Cell Wall Fragment. *Journal of the American Chemical Society*, *134*(33), 13535–13537. https://doi.org/10.1021/ja303883c

Grudniak, A. M., Klecha, B., & Wolska, K. I. (2018). Effects of null mutation of the heat-shock gene *htpG* on the production of virulence factors by *Pseudomonas aeruginosa*. *Future Microbiology*, *13*(1), 69–80. https://doi.org/10.2217/fmb-2017-0111

Grudniak, A. M., Markowska, K., & Wolska, K. I. (2015). Interactions of Escherichia coli molecular chaperone HtpG with DnaA replication initiator DNA. *Cell Stress & Chaperones*, *20*(6), 951–957. https://doi.org/10.1007/s12192-015-0623-y

Grudniak, A. M., Pawlak, K., Bartosik, K., & Wolska, K. I. (2013). Physiological consequences of mutations in the htpG heat shock gene of Escherichia coli. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *745–746*, 1–5. https://doi.org/10.1016/j.mrfmmm.2013.04.003

Halme, L. (2006). Family and twin studies in inflammatory bowel disease. *World Journal of Gastroenterology*, *12*(23), 3668. https://doi.org/10.3748/wjg.v12.i23.3668

Hart, A. L., Al-Hassi, H. O., Rigby, R. J., Bell, S. J., Emmanuel, A. V., Knight, S. C., … Stagg, A. J. (2005). Characteristics of intestinal dendritic cells in inflammatory bowel diseases. *Gastroenterology*, *129*(1), 50–65.

Homer, C. R., Richmond, A. L., Rebert, N. A., Achkar, J.-P., & McDonald, C. (2010). ATG16L1 and NOD2 interact in an autophagy-dependent antibacterial pathway implicated in Crohn's disease pathogenesis. *Gastroenterology*, *139*(5), 1630–1641, 1641.e1-2. https://doi.org/10.1053/j.gastro.2010.07.006

Honoré, F. A., Méjean, V., & Genest, O. (2017). Hsp90 Is Essential under Heat Stress in the Bacterium Shewanella oneidensis. *Cell Reports*, *19*(4), 680–687. https://doi.org/10.1016/j.celrep.2017.03.082

Hoytema van Konijnenburg, D. P., & Mucida, D. (2017). Intraepithelial lymphocytes. *Current Biology*, *27*(15), R737–R739. https://doi.org/10.1016/j.cub.2017.05.073

Hu, S., Wang, Y., Lichtenstein, L., Tao, Y., Musch, M. W., Jabri, B., … Chang, E. B. (2010). Regional differences in colonic mucosa-associated microbiota determine the physiological expression of host heat shock proteins. *American Journal of Physiology. Gastrointestinal and Liver Physiology*, *299*(6), G1266-1275. https://doi.org/10.1152/ajpgi.00357.2010

Huang, C., Haritunians, T., Okou, D. T., Cutler, D. J., Zwick, M. E., Taylor, K. D., … Kugathasan, S. (2015). Characterization of genetic loci that affect susceptibility to inflammatory bowel diseases in African Americans. *Gastroenterology*, *149*(6), 1575–1586. https://doi.org/10.1053/j.gastro.2015.07.065

Huang, L., Zhao, L., Liu, W., Xu, X., Su, Y., Qin, Y., & Yan, Q. (2019). Dual RNA-Seq Unveils Pseudomonas plecoglossicida htpG Gene Functions During Host-Pathogen Interactions With Epinephelus coioides. *Frontiers in Immunology*, *10*, 984. https://doi.org/10.3389/fimmu.2019.00984

Hugot, J.-P., Chamaillard, M., Zouali, H., Lesage, S., Cézard, J.-P., Belaiche, J., … Thomas, G. (2001). Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature*, *411*(6837), 599–603. https://doi.org/10.1038/35079107

Hugot, J.-P., Laurent-Puig, P., Gower-Rousseau, C., Olson, J. M., Lee, J. C., Beaugerie, L., … Thomas, G. (1996). Mapping of a susceptibility locus for Crohn's disease on chromosome 16. *Nature*, *379*(6568), 821–823. https://doi.org/10.1038/379821a0

Irrazábal, T., Belcheva, A., Girardin, S. E., Martin, A., & Philpott, D. J. (2014). The Multifaceted Role of the Intestinal Microbiota in Colon Cancer. *Molecular Cell*, *54*(2), 309–320. https://doi.org/10.1016/j.molcel.2014.03.039

Islam, K. B. M. S., Fukiya, S., Hagio, M., Fujii, N., Ishizuka, S., Ooka, T., … Yokota, A. (2011). Bile acid is a host factor that regulates the composition of the cecal microbiota in rats. *Gastroenterology*, *141*(5), 1773–1781. https://doi.org/10.1053/j.gastro.2011.07.046

Ivanov, I. I., Frutos, R. de L., Manel, N., Yoshinaga, K., Rifkin, D. B., Sartor, R. B., … Littman, D. R. (2008). Specific microbiota direct the differentiation of IL-17-producing T-helper cells in the mucosa of the small intestine. *Cell Host & Microbe*, *4*(4), 337–349. https://doi.org/10.1016/j.chom.2008.09.009

Johnson, J. L. (2012). Evolution and function of diverse Hsp90 homologs and cochaperone proteins. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1823*(3), 607–613. https://doi.org/10.1016/j.bbamcr.2011.09.020

Johnson, R. M. (2004). Murine Oviduct Epithelial Cell Cytokine Responses to Chlamydia muridarum Infection Include Interleukin-12-p70 Secretion. *Infection and Immunity*, *72*(7), 3951–3960. https://doi.org/10.1128/IAI.72.7.3951-3960.2004

Joossens, M., Huys, G., Cnockaert, M., De Preter, V., Verbeke, K., Rutgeerts, P., … Vermeire, S. (2011). Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. *Gut*, *60*(5), 631–637. https://doi.org/10.1136/gut.2010.223263

Jovel, J., Patterson, J., Wang, W., Hotte, N., O'Keefe, S., Mitchel, T., … Wong, G. K.-S. (2016). Characterization of the Gut Microbiome Using 16S or Shotgun Metagenomics. *Frontiers in Microbiology*, *7*, 459. https://doi.org/10.3389/fmicb.2016.00459

Jung, H. C., Eckmann, L., Yang, S. K., Panja, A., Fierer, J., Morzycka-Wroblewska, E., & Kagnoff, M. F. (1995). A distinct array of proinflammatory cytokines is expressed in human colon epithelial cells in response to bacterial invasion. *Journal of Clinical Investigation*, *95*(1), 55–65. https://doi.org/10.1172/JCI117676

Kaakoush, N. O., Day, A. S., Leach, S. T., Lemberg, D. A., & Mitchell, H. M. (2016). Reduction in Gut Microbial Diversity as a Mechanism of Action of Exclusive Enteral Nutrition: *American Journal of Gastroenterology*, *111*(7), 1033. https://doi.org/10.1038/ajg.2016.61

Kabi, A., & McDonald, C. (2012). FRMBP2 directs NOD2 to the membrane. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(52), 21188–21189. https://doi.org/10.1073/pnas.1219395110

Kamm, M. A. (2017). Rapid changes in epidemiology of inflammatory bowel disease. *The Lancet*, *390*(10114), 2741–2742. https://doi.org/10.1016/S0140-6736(17)32669-7

Kamphuis, J. B. J., Mercier-Bonin, M., Eutamène, H., & Theodorou, V. (2017). Mucus organisation is shaped by colonic content; a new view. *Scientific Reports*, *7*(1), 8527. https://doi.org/10.1038/s41598-017-08938-3

Kaparakis-Liaskos, M., & Ferrero, R. L. (2015). Immune modulation by bacterial outer membrane vesicles. *Nature Reviews Immunology*, *15*(6), 375–387. https://doi.org/10.1038/nri3837

Kaplan, G., Benchimol, E., Bernstein, C., Bitton, A., Murthy, S., Nguyen, G., … Weizman, A. (2018). *Impact of Inflammatory Bowel Disease in Canada* (p. 232).

Karatzas, P. S., Gazouli, M., Safioleas, M., & Mantzaris, G. J. (2014). DNA methylation changes in inflammatory bowel disease. *Annals of Gastroenterology*, *27*(2), 125–132.

Kennedy, N. A., Lamb, C. A., Berry, S. H., Walker, A. W., Mansfield, J., Parkes, M., … Lees, C. W. (2018). The Impact of NOD2 Variants on Fecal Microbiota in Crohn's Disease and Controls Without Gastrointestinal Disease. *Inflammatory Bowel Diseases*, *24*(3), 583–592. https://doi.org/10.1093/ibd/izx061

Khor, B., Gardet, A., & Xavier, R. J. (2011). Genetics and pathogenesis of inflammatory bowel disease. *Nature*, *474*(7351), 307–317. https://doi.org/10.1038/nature10209

Kim, Y.-G., Kamada, N., Shaw, M. H., Warner, N., Chen, G. Y., Franchi, L., & Núñez, G. (2011). The Nod2 sensor promotes intestinal pathogen eradication via the chemokine CCL2-dependent recruitment of inflammatory monocytes. *Immunity*, *34*(5), 769–780. https://doi.org/10.1016/j.immuni.2011.04.013

Kim, Y.-G., Park, J.-H., Shaw, M. H., Franchi, L., Inohara, N., & Núñez, G. (2008). The cytosolic sensors Nod1 and Nod2 are critical for bacterial recognition and host defense after exposure to Toll-like receptor ligands. *Immunity*, *28*(2), 246–257. https://doi.org/10.1016/j.immuni.2007.12.012

Kojima, K., Musch, M. W., Ropeleski, M. J., Boone, D. L., Ma, A., & Chang, E. B. (2004). Escherichia coli LPS induces heat shock protein 25 in intestinal epithelial cells through MAP kinase activation. *American Journal of Physiology. Gastrointestinal and Liver Physiology*, *286*(4), G645-652. https://doi.org/10.1152/ajpgi.00080.2003

Lallès, J. P., & David, J. C. (2011). Fasting and refeeding modulate the expression of stress proteins along the gastrointestinal tract of weaned pigs. *Journal of Animal Physiology and Animal Nutrition*, *95*(4), 478–488. https://doi.org/10.1111/j.1439-0396.2010.01075.x

Lasry, A., Zinger, A., & Ben-Neriah, Y. (2016). Inflammatory networks underlying colorectal cancer. *Nature Immunology*, *17*(3), 230–240. https://doi.org/10.1038/ni.3384

Lawley, M., Wu, J. W., Navas-López, V. M., Huynh, H. Q., Carroll, M. W., Chen, M., … Wine, E. (2018). Global Variation in Use of Enteral Nutrition for Pediatric Crohn Disease: *Journal of Pediatric Gastroenterology and Nutrition*, *67*(2), e22–e29. https://doi.org/10.1097/MPG.0000000000001946

Lee, J., Mo, J.-H., Katakura, K., Alkalay, I., Rucker, A. N., Liu, Y.-T., … Raz, E. (2006). Maintenance of colonic homeostasis by distinctive apical TLR9 signalling in intestinal epithelial cells. *Nature Cell Biology*, *8*(12), 1327–1336. https://doi.org/10.1038/ncb1500

Lee, K.-H., Biswas, A., Liu, Y.-J., & Kobayashi, K. S. (2012). Proteasomal Degradation of Nod2 Protein Mediates Tolerance to Bacterial Cell Wall Components. *Journal of Biological Chemistry*, *287*(47), 39800–39811. https://doi.org/10.1074/jbc.M112.410027

Lenaerts, K., Renes, J., Bouwman, F. G., Noben, J.-P., Robben, J., Smit, E., & Mariman, E. C. (2007). Arginine deficiency in preconfluent intestinal Caco-2 cells modulates expression of proteins involved in proliferation, apoptosis, and heat shock response. *Proteomics*, *7*(4), 565–577. https://doi.org/10.1002/pmic.200600715

Lesage, S., Zouali, H., Cézard, J.-P., Colombel, J.-F., Belaiche, J., Almer, S., … Hugot, J.-P. (2002). CARD15/NOD2 Mutational Analysis and Genotype-Phenotype Correlation in 612 Patients with Inflammatory Bowel Disease. *The American Journal of Human Genetics*, *70*(4), 845–857. https://doi.org/10.1086/339432

Levine, A. P., Pontikos, N., Schiff, E. R., Jostins, L., Speed, D., NIDDK Inflammatory Bowel Disease Genetics Consortium, … Segal, A. W. (2016). Genetic Complexity of Crohn's Disease in Two Large Ashkenazi Jewish Families. *Gastroenterology*, *151*(4), 698–709. https://doi.org/10.1053/j.gastro.2016.06.040

Lewis, J. D., & Abreu, M. T. (2017). Diet as a Trigger or Therapy for Inflammatory Bowel Diseases. *Gastroenterology*, *152*(2), 398-414.e6. https://doi.org/10.1053/j.gastro.2016.10.019

Ley, R. E., Hamady, M., Lozupone, C., Turnbaugh, P. J., Ramey, R. R., Bircher, J. S., … Gordon, J. I. (2008). Evolution of mammals and their gut microbes. *Science (New York, N.Y.)*, *320*(5883), 1647–1651. https://doi.org/10.1126/science.1155725

Li, Z., Menoret, A., & Srivastava, P. (2002). Roles of heat-shock proteins in antigen presentation and cross-presentation. *Current Opinion in Immunology*, *14*(1), 45–51.

Liu, J. Z., & Anderson, C. A. (2014). Genetic studies of Crohn's disease: Past, present and future. *Best Practice & Research Clinical Gastroenterology*, *28*(3), 373–386. https://doi.org/10.1016/j.bpg.2014.04.009

Liu, J. Z., van Sommeren, S., Huang, H., Ng, S. C., Alberts, R., Takahashi, A., … Weersma, R. K. (2015). Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nature Genetics*, *47*(9), 979–986. https://doi.org/10.1038/ng.3359

Lotz, M., Gütle, D., Walther, S., Ménard, S., Bogdan, C., & Hornef, M. W. (2006). Postnatal acquisition of endotoxin tolerance in intestinal epithelial cells. *The Journal of Experimental Medicine*, *203*(4), 973–984. https://doi.org/10.1084/jem.20050625

Maaser, C., Egan, L. J., Birkenbach, M. P., Eckmann, L., & Kagnoff, M. F. (2004). Expression of Epstein-Barr virus-induced gene 3 and other interleukin-12-related molecules by human intestinal epithelium. *Immunology*, *112*(3), 437–445. https://doi.org/10.1111/j.1365-2567.2004.01895.x

MacLellan, A., Connors, J., Grant, S., Cahill, L., Langille, M., & Van Limbergen, J. (2017). The Impact of Exclusive Enteral Nutrition (EEN) on the Gut Microbiome in Crohn's Disease: A Review. *Nutrients*, *9*(5), 0447. https://doi.org/10.3390/nu9050447

Magalhaes, J. G., Fritz, J. H., Le Bourhis, L., Sellge, G., Travassos, L. H., Selvanantham, T., … Philpott, D. J. (2008). Nod2-dependent Th2 polarization of antigen-specific immunity. *Journal of Immunology (Baltimore, Md.: 1950)*, *181*(11), 7925–7935.

Marina-Garcia, N., Franchi, L., Kim, Y.-G., Hu, Y., Smith, D. E., Boons, G.-J., & Nunez, G. (2009). Clathrin- and Dynamin-Dependent Endocytic Pathway Regulates Muramyl Dipeptide Internalization and NOD2 Activation. *The Journal of Immunology*, *182*(7), 4321–4327. https://doi.org/10.4049/jimmunol.0802197

Martinez-Porchas, M., Villalpando-Canchola, E., Ortiz Suarez, L. E., & Vargas-Albores, F. (2017). How conserved are the conserved 16S-rRNA regions? *PeerJ*, *5*, e3036. https://doi.org/10.7717/peerj.3036

Maunder, R. (2000). Mediators of stress effects in inflammatory bowel disease: Not the usual suspects. *Journal of Psychosomatic Research*, *48*(6), 569–577.

Meyer, A. M., Ramzan, N. N., Heigh, R. I., & Leighton, J. A. (2006). Relapse of inflammatory bowel disease associated with use of nonsteroidal anti-inflammatory drugs. *Digestive Diseases and Sciences*, *51*(1), 168–172. https://doi.org/10.1007/s10620-006-3103-5

Mirkov, M. U., Verstockt, B., & Cleynen, I. (2017). Genetics of inflammatory bowel disease: Beyond NOD2. *The Lancet. Gastroenterology & Hepatology*, *2*(3), 224–234. https://doi.org/10.1016/S2468-1253(16)30111-X

Mo, J., Boyle, J. P., Howard, C. B., Monie, T. P., Davis, B. K., & Duncan, J. A. (2012). Pathogen Sensing by Nucleotide-binding Oligomerization Domain-containing Protein 2 (NOD2) Is Mediated by Direct Binding to Muramyl Dipeptide and ATP. *Journal of Biological Chemistry*, *287*(27), 23057–23067. https://doi.org/10.1074/jbc.M112.344283

Mohanan, V., & Grimes, C. L. (2014). The Molecular Chaperone HSP70 Binds to and Stabilizes NOD2, an Important Protein Involved in Crohn Disease. *Journal of Biological Chemistry*, *289*(27), 18987–18998. https://doi.org/10.1074/jbc.M114.557686

Mow, W. S., Vasiliauskas, E. A., Lin, Y.-C., Fleshner, P. R., Papadakis, K. A., Taylor, K. D., … Targan, S. R. (2004). Association of antibody responses to microbial antigens and complications of small bowel Crohn's disease. *Gastroenterology*, *126*(2), 414–424.

Mowat, A. M., & Agace, W. W. (2014). Regional specialization within the intestinal immune system. *Nature Reviews. Immunology*, *14*(10), 667–685. https://doi.org/10.1038/nri3738

Moya, A., & Ferrer, M. (2016). Functional Redundancy-Induced Stability of Gut Microbiota Subjected to Disturbance. *Trends in Microbiology*, *24*(5), 402–413. https://doi.org/10.1016/j.tim.2016.02.002

Murch, S. H., Braegger, C. P., Walker-Smith, J. A., & MacDonald, T. T. (1993). Location of tumour necrosis factor alpha by immunohistochemistry in chronic inflammatory bowel disease. *Gut*, *34*(12), 1705–1709. https://doi.org/10.1136/gut.34.12.1705

Nakamura, N., Lill, J. R., Phung, Q., Jiang, Z., Bakalarski, C., de Mazière, A., … Mellman, I. (2014). Endosomes are specialized platforms for bacterial sensing and NOD2 signalling. *Nature*, *509*(7499), 240–244. https://doi.org/10.1038/nature13133

Nam, S. Y., Kim, N., Kim, J. S., Lim, S. H., Jung, H. C., & Song, I. S. (2007). Heat shock protein gene 70-2 polymorphism is differentially associated with the clinical phenotypes of ulcerative colitis and Crohn's disease. *Journal of Gastroenterology and Hepatology*, *22*(7), 1032–1038. https://doi.org/10.1111/j.1440-1746.2007.04927.x

Negroni, A., Pierdomenico, M., Cucchiara, S., & Stronati, L. (2018). NOD2 and inflammation: Current insights. *Journal of Inflammation Research*, *Volume 11*, 49–60. https://doi.org/10.2147/JIR.S137606

Ng, S. C., Shi, H. Y., Hamidi, N., Underwood, F. E., Tang, W., Benchimol, E. I., … Kaplan, G. G. (2017). Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: A systematic review of population-based studies. *The Lancet*, *390*(10114), 2769–2778. https://doi.org/10.1016/S0140-6736(17)32448-0

Ng, S. C., Tsoi, K. K. F., Kamm, M. A., Xia, B., Wu, J., Chan, F. K. L., & Sung, J. J. Y. (2012). Genetics of inflammatory bowel disease in Asia: Systematic review and meta-analysis: *Inflammatory Bowel Diseases*, *18*(6), 1164–1176. https://doi.org/10.1002/ibd.21845

Nimmo, E. R., Prendergast, J. G., Aldhous, M. C., Kennedy, N. A., Henderson, P., Drummond, H. E., … Satsangi, J. (2012). Genome-wide methylation profiling in Crohn's disease identifies altered epigenetic regulation of key host defense mechanisms including the Th17 pathway: *Inflammatory Bowel Diseases*, *18*(5), 889–899. https://doi.org/10.1002/ibd.21912

O'Brien, J., Hayder, H., Zayed, Y., & Peng, C. (2018). Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Frontiers in Endocrinology*, *9*, 402. https://doi.org/10.3389/fendo.2018.00402

Ohno, H. (2016). Intestinal M cells. *Journal of Biochemistry*, *159*(2), 151–160. https://doi.org/10.1093/jb/mvv121

Okamoto, K., Fujiya, M., Nata, T., Ueno, N., Inaba, Y., Ishikawa, C., … Kohgo, Y. (2012). Competence and sporulation factor derived from Bacillus subtilis improves epithelial cell injury in intestinal inflammation via immunomodulation and cytoprotection. *International Journal of Colorectal Disease*, *27*(8), 1039–1046. https://doi.org/10.1007/s00384-012-1416-8

Okumura, R., & Takeda, K. (2017). Roles of intestinal epithelial cells in the maintenance of gut homeostasis. *Experimental & Molecular Medicine*, *49*(5), e338. https://doi.org/10.1038/emm.2017.20

Parkes, G. C., Whelan, K., & Lindsay, J. O. (2014). Smoking in inflammatory bowel disease: Impact on disease course and insights into the aetiology of its effect. *Journal of Crohn's and Colitis*, *8*(8), 717–725. https://doi.org/10.1016/j.crohns.2014.02.002

Pearl, L. H., & Prodromou, C. (2006). Structure and mechanism of the Hsp90 molecular chaperone machinery. *Annual Review of Biochemistry*, *75*, 271–294. https://doi.org/10.1146/annurev.biochem.75.103004.142738

Penagini, F., Dilillo, D., Borsani, B., Cococcioni, L., Galli, E., Bedogni, G., … Zuccotti, G. V. (2016). Nutrition in Pediatric Inflammatory Bowel Disease: From Etiology to Treatment. A Systematic Review. *Nutrients*, *8*(6). https://doi.org/10.3390/nu8060334

Peterson, L. W., & Artis, D. (2014). Intestinal epithelial cells: Regulators of barrier function and immune homeostasis. *Nature Reviews. Immunology*, *14*(3), 141–153. https://doi.org/10.1038/nri3608

Petnicki-Ocwieja, T., Hrncir, T., Liu, Y.-J., Biswas, A., Hudcovic, T., Tlaskalova-Hogenova, H., & Kobayashi, K. S. (2009). Nod2 is required for the regulation of commensal microbiota in the intestine. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(37), 15813–15818. https://doi.org/10.1073/pnas.0907722106

Philpott, D. J., Sorbara, M. T., Robertson, S. J., Croitoru, K., & Girardin, S. E. (2014). NOD proteins: Regulators of inflammation in health and disease. *Nature Reviews. Immunology*, *14*(1), 9–23. https://doi.org/10.1038/nri3565

Pierdomenico, M., Cesi, V., Cucchiara, S., Vitali, R., Prete, E., Costanzo, M., … Stronati, L. (2016). NOD2 Is Regulated By Mir-320 in Physiological Conditions but this Control Is Altered in Inflamed Tissues of Patients with Inflammatory Bowel Disease: *Inflammatory Bowel Diseases*, *22*(2), 315–326. https://doi.org/10.1097/MIB.0000000000000659

Pockley, A. G., & Henderson, B. (2018). Extracellular cell stress (heat shock) proteins—immune responses and disease: An overview. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1738), 20160522. https://doi.org/10.1098/rstb.2016.0522

Powell, N., Walker, M. M., & Talley, N. J. (2010). Gastrointestinal eosinophils in health, disease and functional disorders. *Nature Reviews. Gastroenterology & Hepatology*, *7*(3), 146–156. https://doi.org/10.1038/nrgastro.2010.5

Price, A. E., Shamardani, K., Lugo, K. A., Deguine, J., Roberts, A. W., Lee, B. L., & Barton, G. M. (2018). A Map of Toll-like Receptor Expression in the Intestinal Epithelium Reveals Distinct Spatial, Cell Type-Specific, and Temporal Patterns. *Immunity*, *49*(3), 560-575.e6. https://doi.org/10.1016/j.immuni.2018.07.016

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., … Wang, J. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*, *464*(7285), 59–65. https://doi.org/10.1038/nature08821

Quince, C., Ijaz, U. Z., Loman, N., Eren, M. A., Saulnier, D., Russell, J., … Gerasimidis, K. (2015). Extensive Modulation of the Fecal Metagenome in Children With Crohn's Disease During Exclusive Enteral Nutrition: *American Journal of Gastroenterology*, *110*(12), 1718–1729. https://doi.org/10.1038/ajg.2015.357

Rakoff-Nahoum, S., Paglino, J., Eslami-Varzaneh, F., Edberg, S., & Medzhitov, R. (2004). Recognition of Commensal Microflora by Toll-Like Receptors Is Required for Intestinal Homeostasis. *Cell*, *118*(2), 229–241. https://doi.org/10.1016/j.cell.2004.07.002

Ramanan, D., Tang, M. S., Bowcutt, R., Loke, P., & Cadwell, K. (2014). Bacterial sensor Nod2 prevents inflammation of the small intestine by restricting the expansion of the commensal Bacteroides vulgatus. *Immunity*, *41*(2), 311–324. https://doi.org/10.1016/j.immuni.2014.06.015

Reboldi, A., & Cyster, J. G. (2016). Peyer's patches: Organizing B-cell responses at the intestinal frontier. *Immunological Reviews*, *271*(1), 230–245. https://doi.org/10.1111/imr.12400

Reis, A. C. M., Silva, J. O., Laranjeira, B. J., Pinheiro, A. Q., & Carvalho, C. B. M. (2014). Virulence factors and biofilm production by isolates of Bacteroides fragilis recovered from dog intestinal tracts. *Brazilian Journal of Microbiology*, *45*(2), 647–650. https://doi.org/10.1590/S1517-83822014000200037

Rimoldi, M., Chieppa, M., Salucci, V., Avogadri, F., Sonzogni, A., Sampietro, G. M., … Rescigno, M. (2005). Intestinal immune homeostasis is regulated by the crosstalk between epithelial cells and dendritic cells. *Nature Immunology*, *6*(5), 507–514. https://doi.org/10.1038/ni1192

Ruemmele, F.M., Veres, G., Kolho, K. L., Griffiths, A., Levine, A., Escher, J. C., … Turner, D. (2014). Consensus guidelines of ECCO/ESPGHAN on the medical management of pediatric Crohn's disease. *Journal of Crohn's and Colitis*, *8*(10), 1179–1207. https://doi.org/10.1016/j.crohns.2014.04.005

Ruemmele, Frank M. (2016). Role of Diet in Inflammatory Bowel Disease. *Annals of Nutrition and Metabolism*, *68*(1), 33–41. https://doi.org/10.1159/000445392

Saxena, A., Lopes, F., Poon, K. K. H., & McKay, D. M. (2017). Absence of the NOD2 protein renders epithelia more susceptible to barrier dysfunction due to mitochondrial dysfunction. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, *313*(1), G26–G38. https://doi.org/10.1152/ajpgi.00070.2017

Seegert, D., Rosenstiel, P., Pfahler, H., Pfefferkorn, P., Nikolaus, S., & Schreiber, S. (2001). Increased expression of IL-16 in inflammatory bowel disease. *Gut*, *48*(3), 326–332. https://doi.org/10.1136/gut.48.3.326

Segal, A. W. (2018). The role of neutrophils in the pathogenesis of Crohn's disease. *European Journal of Clinical Investigation*, *48*, e12983. https://doi.org/10.1111/eci.12983

Sender, R., Fuchs, S., & Milo, R. (2016). Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans. *Cell*, *164*(3), 337–340. https://doi.org/10.1016/j.cell.2016.01.013

Shelburne, C. E., Coopamah, M. D., Sweier, D. G., An, F. Y.-P., & Lopatin, D. E. (2007). HtpG, the Porphyromonas gingivalis HSP-90 homologue, induces the chemokine CXCL8 in human monocytic and microvascular vein endothelial cells. *Cellular Microbiology*, *9*(6), 1611–1619. https://doi.org/10.1111/j.1462-5822.2007.00897.x

Shelburne, C. E., Shelburne, P. S., Dhople, V. M., Sweier, D. G., Giannobile, W. V., Kinney, J. S., … Lopatin, D. E. (2008). Serum Antibodies to Porphyromonas gingivalis Chaperone HtpG Predict Health in Periodontitis Susceptible Patients. *PLoS ONE*, *3*(4), e1984. https://doi.org/10.1371/journal.pone.0001984

Sidiq, T., Yoshihama, S., Downs, I., & Kobayashi, K. S. (2016). Nod2: A Critical Regulator of Ileal Microbiota and Crohn's Disease. *Frontiers in Immunology*, *7*, 367. https://doi.org/10.3389/fimmu.2016.00367

Siegmund, B., & Zeitz, M. (2011). Innate and adaptive immunity in inflammatory bowel disease. *World Journal of Gastroenterology*, *17*(27), 3178–3183. https://doi.org/10.3748/wjg.v17.i27.3178

Silverberg, M. S., Satsangi, J., Ahmad, T., Arnott, I. D., Bernstein, C. N., Brant, S. R., … Warren, B. F. (2005). Toward an Integrated Clinical, Molecular and Serological Classification of Inflammatory Bowel Disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. *Canadian Journal of Gastroenterology*, *19*(suppl a), 5A-36A. https://doi.org/10.1155/2005/269076

Song, H. Y., Dunbar, J. D., Zhang, Y. X., Guo, D., & Donner, D. B. (1995). Identification of a Protein with Homology to hsp90 That Binds the Type 1 Tumor Necrosis Factor Receptor. *Journal of Biological Chemistry*, *270*(8), 3574–3581. https://doi.org/10.1074/jbc.270.8.3574

Soroosh, A., Koutsioumpa, M., Pothoulakis, C., & Iliopoulos, D. (2018). Functional role and therapeutic targeting of microRNAs in inflammatory bowel disease. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, *314*(2), G256–G262. https://doi.org/10.1152/ajpgi.00268.2017

Spence, J. R., Lauf, R., & Shroyer, N. F. (2011). Vertebrate intestinal endoderm development. *Developmental Dynamics*, *240*(3), 501–520. https://doi.org/10.1002/dvdy.22540

Sproule-Willoughby, K. M., Stanton, M. M., Rioux, K. P., McKay, D. M., Buret, A. G., & Ceri, H. (2010). In vitro anaerobic biofilms of human colonic microbiota. *Journal of Microbiological Methods*, *83*(3), 296–301. https://doi.org/10.1016/j.mimet.2010.09.020

Stewart, L., D M Edgar, J., Blakely, G., & Patrick, S. (2018). Antigenic mimicry of ubiquitin by the gut bacterium Bacteroides fragilis: A potential link with autoimmune disease. *Clinical and Experimental Immunology*, *194*(2), 153–165. https://doi.org/10.1111/cei.13195

Strober, W., Murray, P. J., Kitani, A., & Watanabe, T. (2006). Signalling pathways and molecular interactions of NOD1 and NOD2. *Nature Reviews. Immunology*, *6*(1), 9–20. https://doi.org/10.1038/nri1747

Sun, Y., Li, L., Xia, Y., Li, W., Wang, K., Wang, L., … Ma, S. (2019). The gut microbiota heterogeneity and assembly changes associated with the IBD. *Scientific Reports*, *9*(1), 440. https://doi.org/10.1038/s41598-018-37143-z

Swidsinski, A., Ladhoff, A., Pernthaler, A., Swidsinski, S., Loening-Baucke, V., Ortner, M., … Lochs, H. (2002). Mucosal flora in inflammatory bowel disease. *Gastroenterology*, *122*(1), 44–54.

Takahashi, S., Andreoletti, G., Chen, R., Munehira, Y., Batra, A., Afzal, N. A., … Snyder, M. (2017). De novo and rare mutations in the HSPA1L heat shock gene associated with inflammatory bowel disease. *Genome Medicine*, *9*(1), 8. https://doi.org/10.1186/s13073-016-0394-9

Takayama, T., Kamada, N., Chinen, H., Okamoto, S., Kitazume, M. T., Chang, J., … Hibi, T. (2010). Imbalance of NKp44(+)NKp46(-) and NKp44(-)NKp46(+) natural killer cells in the intestinal mucosa of patients with Crohn's disease. *Gastroenterology*, *139*(3), 882–892, 892.e1-3. https://doi.org/10.1053/j.gastro.2010.05.040

Tan, G., Zeng, B., & Zhi, F.-C. (2015). Regulation of human enteric α-defensins by NOD2 in the Paneth cell lineage. *European Journal of Cell Biology*, *94*(1), 60–66. https://doi.org/10.1016/j.ejcb.2014.10.007

Tanca, A., Abbondio, M., Palomba, A., Fraumene, C., Manghina, V., Cucca, F., … Uzzau, S. (2017). Potential and active functions in the gut microbiota of a healthy human cohort. *Microbiome*, *5*(1), 79. https://doi.org/10.1186/s40168-017-0293-3

Tigno-Aranjuez, J. T., & Abbott, D. W. (2012). Ubiquitination and phosphorylation in the regulation of NOD2 signaling and NOD2-mediated disease. *Biochimica Et Biophysica Acta*, *1823*(11), 2022–2028. https://doi.org/10.1016/j.bbamcr.2012.03.017

Ting, H.-A., & von Moltke, J. (2019). The Immune Function of Tuft Cells at Gut Mucosal Surfaces and Beyond. *The Journal of Immunology*, *202*(5), 1321–1329. https://doi.org/10.4049/jimmunol.1801069

Ting, J. P.-Y., & Davis, B. K. (2005). CATERPILLER: A novel gene family important in immunity, cell death, and diseases. *Annual Review of Immunology*, *23*, 387–414. https://doi.org/10.1146/annurev.immunol.23.021704.115616

Torres, J., Mehandru, S., Colombel, J.-F., & Peyrin-Biroulet, L. (2017). Crohn's disease. *The Lancet*, *389*(10080), 1741–1755. https://doi.org/10.1016/S0140-6736(16)31711-1

Traub, S., von Aulock, S., Hartung, T., & Hermann, C. (2006). MDP and other muropeptides--direct and synergistic effects on the immune system. *Journal of Endotoxin Research*, *12*(2), 69–85. https://doi.org/10.1179/096805106X89044

Travassos, L. H., Carneiro, L. A. M., Girardin, S., & Philpott, D. J. (2010). Nod proteins link bacterial sensing and autophagy. *Autophagy*, *6*(3), 409–411.

Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R., & Gordon, J. I. (2007). The human microbiome project. *Nature*, *449*(7164), 804–810. https://doi.org/10.1038/nature06244

Tysk, C., Lindberg, E., Jarnerot, G., & Floderus-Myrhed, B. (1988). Ulcerative colitis and Crohn's disease in an unselected population of monozygotic and dizygotic twins. A study of heritability and the influence of smoking. *Gut*, *29*(7), 990–996. https://doi.org/10.1136/gut.29.7.990

Uhlig, H. H., Schwerd, T., Koletzko, S., Shah, N., Kammermeier, J., Elkadri, A., … Muise, A. M. (2014). The Diagnostic Approach to Monogenic Very Early Onset Inflammatory Bowel Disease. *Gastroenterology*, *147*(5), 990-1007.e3. https://doi.org/10.1053/j.gastro.2014.07.023

van Eden, W., Jansen, M. A. A., Ludwig, I., van Kooten, P., van der Zee, R., & Broere, F. (2017). The Enigma of Heat Shock Proteins in Immune Tolerance. *Frontiers in Immunology*, *8*, 1599. https://doi.org/10.3389/fimmu.2017.01599

Van Limbergen, J., Haskett, J., Griffiths, A. M., Critch, J., Huynh, H., Ahmed, N., … Otley, A. R. (2015). Toward Enteral Nutrition in the Treatment of Pediatric Crohn Disease in Canada: A Workshop to Identify Barriers and Enablers. *Canadian Journal of Gastroenterology and Hepatology*, *29*(7), 351–356. https://doi.org/10.1155/2015/509497

Van Molle, W., Wielockx, B., Mahieu, T., Takada, M., Taniguchi, T., Sekikawa, K., & Libert, C. (2002). HSP70 protects against TNF-induced lethal inflammatory shock. *Immunity*, *16*(5), 685–695.

Vavricka, S. R., Musch, M. W., Chang, J. E., Nakagawa, Y., Phanvijhitsiri, K., Waypa, T. S., … Chang, E. B. (2004). hPepT1 transports muramyl dipeptide, activating NF-kappaB and stimulating IL-8 secretion in human colonic Caco2/bbe cells. *Gastroenterology*, *127*(5), 1401–1409.

Walton, C., Montoya, M. P. B., Fowler, D. P., Turner, C., Jia, W., Whitehead, R. N., … Hunter, J. O. (2016). Enteral feeding reduces metabolic activity of the intestinal microbiome in Crohn's disease: An observational study. *European Journal of Clinical Nutrition*, *70*(9), 1052–1056. https://doi.org/10.1038/ejcn.2016.74

Watt, E., Gemmell, M. R., Berry, S., Glaire, M., Farquharson, F., Louis, P., … Hold, G. L. (2016). Extending colonic mucosal microbiome analysis-assessment of colonic lavage as a proxy for endoscopic colonic biopsies. *Microbiome*, *4*(1), 61. https://doi.org/10.1186/s40168-016-0207-9

Wehkamp, J., Salzman, N. H., Porter, E., Nuding, S., Weichenthal, M., Petras, R. E., … Bevins, C. L. (2005). Reduced Paneth cell alpha-defensins in ileal Crohn's disease. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(50), 18129–18134. https://doi.org/10.1073/pnas.0505256102

Wexler, A. G., & Goodman, A. L. (2017). An insider's perspective: Bacteroides as a window into the microbiome. *Nature Microbiology*, *2*, 17026. https://doi.org/10.1038/nmicrobiol.2017.26

Whitten, K. E., Rogers, P., Ooi, C. K. Y., & Day, A. S. (2012). International survey of enteral nutrition protocols used in children with Crohn's disease: Enteral nutrition for children with CD. *Journal of Digestive Diseases*, *13*(2), 107–112. https://doi.org/10.1111/j.1751-2980.2011.00558.x

Wu, F., Zhang, S., Dassopoulos, T., Harris, M. L., Bayless, T. M., Meltzer, S. J., … Kwon, J. H. (2010). Identification of microRNAs associated with ileal and colonic Crohn's disease†: *Inflammatory Bowel Diseases*, *16*(10), 1729–1738. https://doi.org/10.1002/ibd.21267

Yamada, A., Arakaki, R., Saito, M., Tsunematsu, T., Kudo, Y., & Ishimaru, N. (2016). Role of regulatory T cell in the pathogenesis of inflammatory bowel disease. *World Journal of Gastroenterology*, *22*(7), 2195–2205. https://doi.org/10.3748/wjg.v22.i7.2195

Yamanaka, T., Furukawa, T., Matsumoto-Mashimo, C., Yamane, K., Sugimori, C., Nambu, T., … Fukushima, H. (2009). Gene expression profile and pathogenicity of biofilm-forming Prevotella intermedia strain 17. *BMC Microbiology*, *9*(1), 11. https://doi.org/10.1186/1471-2180-9-11

Yosef, I., Goren, M. G., Kiro, R., Edgar, R., & Qimron, U. (2011). High-temperature protein G is essential for activity of the Escherichia coli clustered regularly interspaced short palindromic repeats (CRISPR)/Cas system. *Proceedings of the National Academy of Sciences*, *108*(50), 20136–20141. https://doi.org/10.1073/pnas.1113519108

Yu, L., Chen, Y., & Tooze, S. A. (2018). Autophagy pathway: Cellular and molecular mechanisms. *Autophagy*, *14*(2), 207–215. https://doi.org/10.1080/15548627.2017.1378838

Zakharzhevskaya, N. B., Vanyushkina, A. A., Altukhov, I. A., Shavarda, A. L., Butenko, I. O., Rakitina, D. V., … Govorun, V. M. (2017). Outer membrane vesicles secreted by pathogenic and nonpathogenic Bacteroides fragilis represent different metabolic activities. *Scientific Reports*, *7*(1), 5008. https://doi.org/10.1038/s41598-017-05264-6

Zeuthen, L. H., Fink, L. N., & Frokiaer, H. (2008). Epithelial cells prime the immune response to an array of gut-derived commensals towards a tolerogenic phenotype through distinct actions of thymic stromal lymphopoietin and transforming growth factor-beta. *Immunology*, *123*(2), 197–208. https://doi.org/10.1111/j.1365-2567.2007.02687.x

Zou, Y., Xue, W., Luo, G., Deng, Z., Qin, P., Guo, R., … Xiao, L. (2019). 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. *Nature Biotechnology*, *37*(2), 179–185. https://doi.org/10.1038/s41587-018-0008-8

Appendix A: Publications during MSc Degree

Reviews

Finlayson-Trick, E., Connors, J., Stadnyk, A., Van Limbergen, J. (2018). Regulation Antimicrobial Pathways by Endogenous Heat Shock Proteins in Gastrointestinal Disorders. *Gastrointestinal Disorders, 1*(1). https://doi.org/10.3390/gidisord1010005

Published Abstracts

Finlayson-Trick, E., Connors, J., Dunn, K., Stadnyk, A., Van Limbergen, J. (2019). Mo1932 Mechanistic Insight Into Bacterial-Host Interactions in Crohn's Disease. *Gastroenterology, 156*(6). doi: 10.1016/S0016-5085(19)319199-1

Finlayson-Trick, E., Connors, J., Whitehouse, S., Stadnyk, A., Van Limbergen, J. (2018). A129 The Role of Bacterial HtpG in Host-Microbiome Interactions in Crohn's Disease. *Journal of the Canadian Association of Gastroenterology, 1*, 194. https://doi.org/10.1093/jcag/gwy009.129

Publications

Thornbury, M., Sicheri, J., Slaine, P., Getz, L., Finlayson-Trick, E., Cook, J., Guinard, C., Boudreau, N., Jackman, D., Rohde, J., McCormick, C. (2019). Characterization of novel lignocellulose-degrading enzymes from the porcupine microbiome using synthetic metagenomics. *PLoS ONE, 14*(1), e0209221. https://doi.org/10.1371/journal.pone.0209221

Appendix B: Synthetic *B. fragilis* (Group B) HtpG Construct

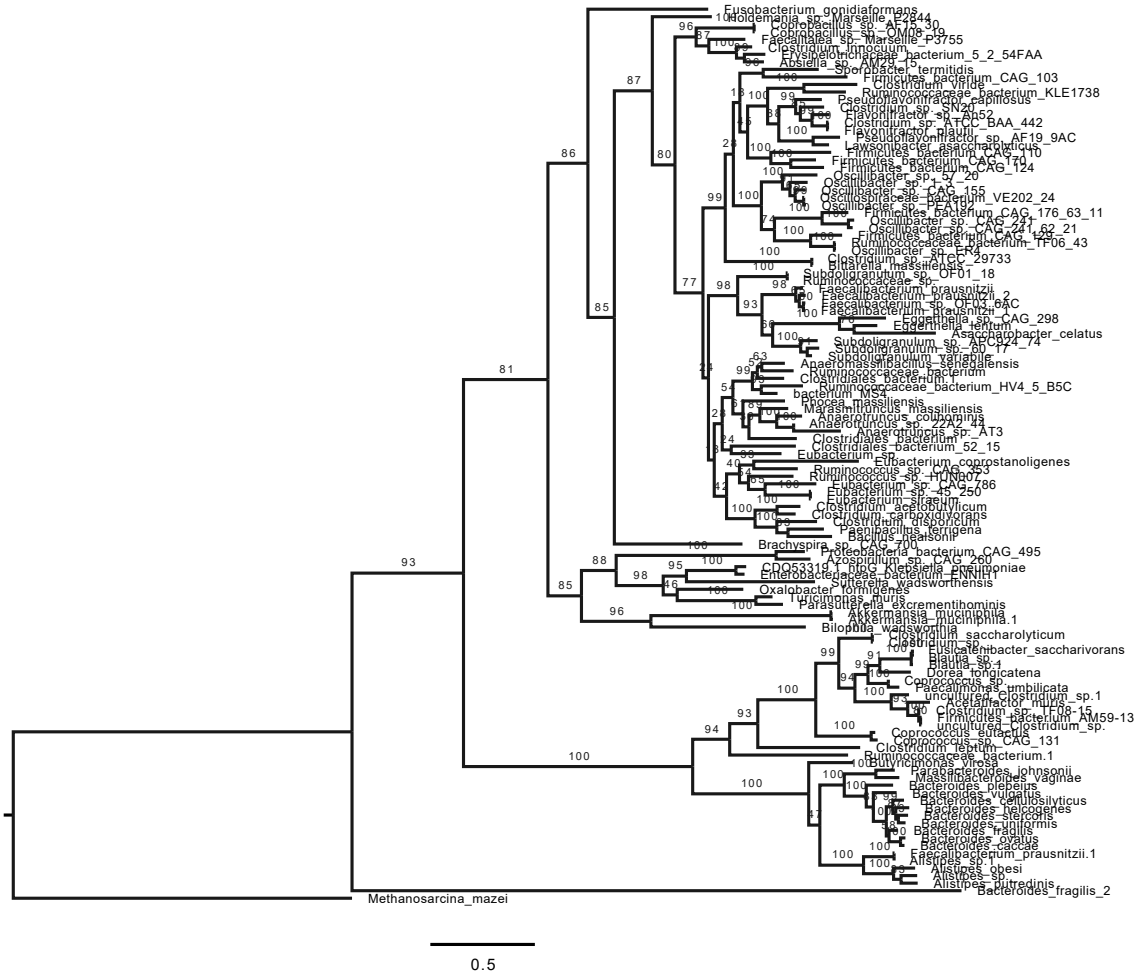
acctgcggatccaggaggcacaaac**atg**caaaaaggtaatattggggttacaacagagaacattttccctatcatcaaaaagtttttgtacag
tgaccatgaaatcttcctgcgggaattagtatccaatgccgttgatgccactcagaagttgaatacattggcttctatcagtgaatttaagggcg
aactgggtgatttgaccgttcacgtttcattaggcaaagacaccattaccatctccgatcgtggtatcggtttgactgctgaagagattgataaa
tatatcaaccagattgccttttcgggggctaacgatttccttgaaaaatataaaaacgatgcgaatgccatcattggacacttcggacttgggtt
ctactctgcattcatggtttccaagaaggttgaaattatcaccaaatcatataaagaaggtgcacaggccgtaaaatggacttgcgacggtagt
ccggagtttacacttgaagaggtggagaaagcggatcgtggtacagatatcgtattgtatattgatgatgattgcaaggagtttcttgaggagt
cacgcatctctgccctcctgaagaaatattgcagcttcctgcccgttcccatcgcttttggtaaaaagaaagagtggaaagacggcaaacaa
gtcgagacggcggaagataatgtcatcaatgacaccattcctttgtggacaaagaaaccgagtgaattgtcggacgaagattataaaaaatt
ctatcgtgagctttatccgatgtcagacgaacctttgttctggattcatttgaatgtagactatccgttccatctgaccggtatcctctacttcccga
aggtaaagagcaatattgatttgaataagaataagattcagttgtattgtaatcaggtttatgttacggattctgtagaaggtattgttccggatttc
cttactctgctccatggtgtgctcgattcaccggatattcctttgaatgtatcccgttcttacctgcaaagtgattcgaacgtgaagaagatctcta
cctatatttcgaaaaaggtatcagaccgtctgcaatctatctttaagaatgatcgcgctcagttcgaagagaagtggaatgatttaaaaatcttta
ttaattatggaatgctcactcaagaggatttctatgataaagcacaaaaattcgccctttcaccgatacggatggcaaatattacacctttgagg
agtaccagactttgattaaagataatcagacagataaagataaaaacctgatctatctgtatgccaataataaggacgaacagtttgcctatatc
gaagctgccaaaaataaaggttacaatgtgctgttgatggacgggcaactggatgtggccatggtaagtatgctcgaacagaaactggaga
aatctcgcttcacccgtgtagacagtgatgttgtcgacaacctgattgtgaaagaagataagaagagcgatgtgcttgaggcttcaaaacaag
aagctctgtcagcagccttcaagagtcagttgccgaaaatggaaaaggttgaatttaatgtcatgactcaggctttaggcgaaaacggctctc
ccgtgatgataacccagagcgaatatatgcgccgtatgaaggaaatggccaatatccaggctggcatgagtttctatggtgaaatgcccgat
atgtttaatctggtattgaattcagaccataaattagtgaaagaagtattggctgatgaagaaaaagagtgcagtgctgccattgctcctataca
gacggaactggaagatgtgacaaaacgtcgtgatgcactcaagaaaaagcaagaaggcaagaaagacgaagatatccctactgtggaga
aagatgaactcaatgatctggataagaaatgggatgagttgaagcagcagaaagattctatttttgccggatatgcaggcaaaaacaaagtg
gtacgtcagttgatcgatctggcgttgttgcaaaacaatatgctgaaaggtgaagcattaaataactttgtaaaaagaagcattgaactgattaa
agccgcc***caccaccaccaccaccac***taagccgccctcgagaccgta

Appendix B1. Group B *B. fragilis* HtpG nucleotide sequence ordered from IDT. Nucleotide sequence for BF2409 was downloaded from the KEGG online database. Restriction sites are underlined, the 6x-His-tag is italicized, and the start codon is bolded.

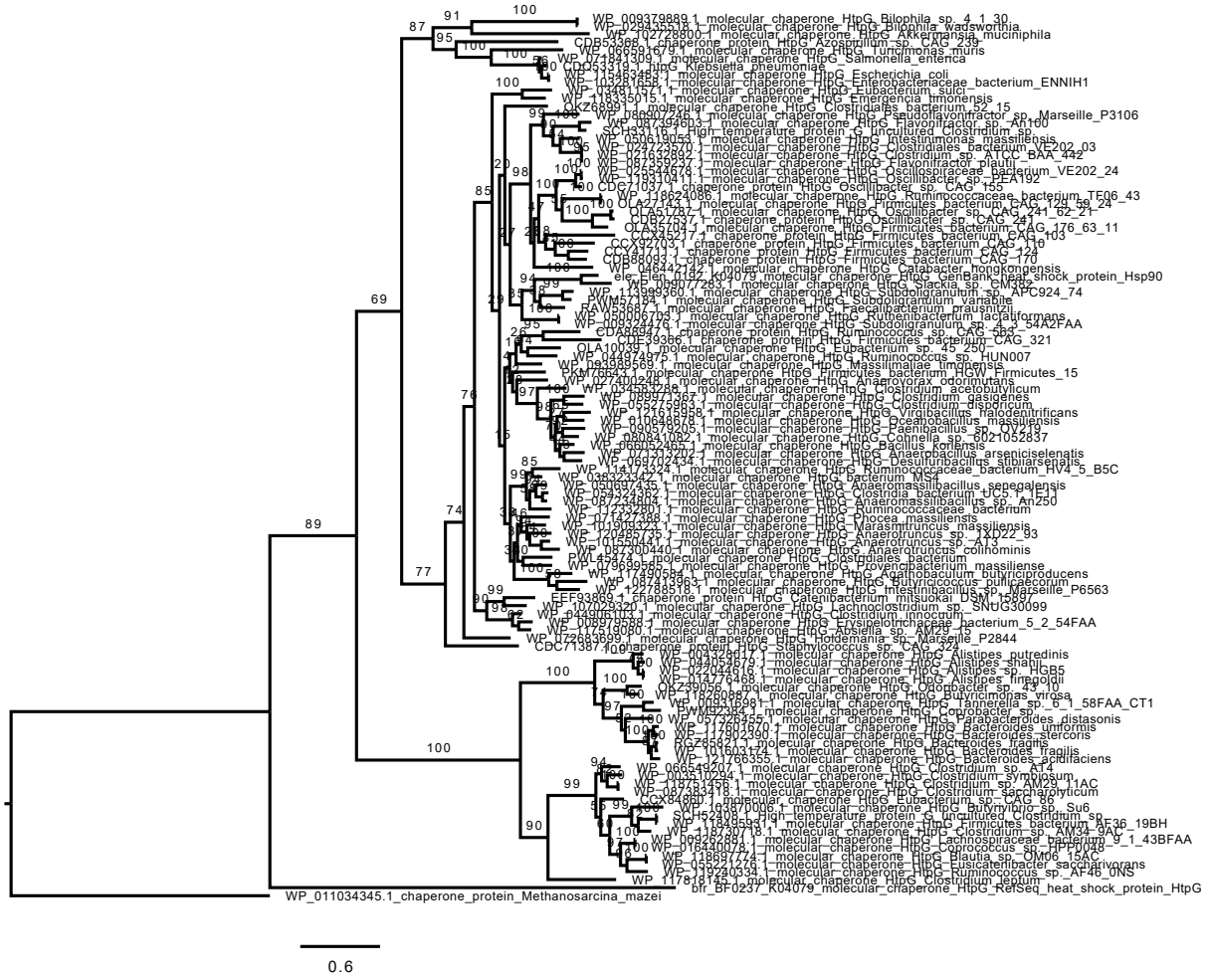
Acctgcggatccaggaggcaccacc**atg**gaaaaagaaggtaataatctgtttcaggtcaatcttaaaggtatgattgccctgttgtcagagc
atatttatagtaatccgaatacttttgtccgggagttattgcaaaatagtgtggatgccatcactgcattgcacaacatcgatgaaaattactccg
gacgtattgatgtcttcctgaatggggatggctcgatggtctttcaggacaatggaatcggactgaaggaagaggaggtataccgtttcctca
cagtgataggtgaaagttcaaagagagacactcccgatgccgacgattttatcggtcggtttggtatcggcttgttgtcttgttttgtggtgacc
aatgaaatcagggtagagagccggtcggcaatgggggggaaatcctgtttgctggtgcggaaaggtggacggtacttatcagactactttcc
ccgatgaagagtgggagatcggttcaagggttgtgttgaggcctaaaaatgaatgggctcatctgttcgaatacgaagtgtttaaaaagatatt
ggtaaattatggagaagtcttgccatatcctgtctatttacatcgtggagaagaagaagagttggttaatactccatcgcccgtctggcttgatc
cgaaagctacccggaaagagttattggattatgggacaaaggtcttccaatcgtctgccctggatgcatttcctatacggacagagcatggac
ggatagaaggtgtactctatgtattgcctttccgtacgcaattctctgtgcgtaattcgcataaagtatatctgaagcggatgttgcttagtgagg
acgattgcaatctgttgccttcttgggcatttttattcgttgcctggtgaatgccgacggactgctttccacagcctcccgcgaatcatttgtcag
caacgattcattgaaagatgcccggaaagagattggggtcgctatcaaggaatatctccgggctttggtgcagaacaatcggtccgttttaat
aaaatactggatgttcatcactttcacattaaagccattgcttcggaagataatgagttacttcgtctgtttatggattatcttccgtttgagacgaat
aagggaataagaagtttcggtagtatccgttcgtcaaataatactatttattatacacgcaatctggaagatttcaggcaagtgcgcagaatag
ccggtgcacaaggcaggctggtagttaatgctgcttatacattcgatgaaacgttgctgaaaaaatatatccggcttaatcaagaactgtcttta
gaggagatttcacctgcccgtctgttggaagagtttgccgaagtagagggtaataaagagcaccgatcttttgaaacgaaagccagtgaact
tttggaacgttttgggtgtatttgccggttgaagcatttcactccggtggatactccggttatattcgtagccgaagaaaaagaagagaacagta
aagtcgccaataatccgttggcggcggtattgggctcggtaaatgccaaaaaacgtttgccacccacattgactttcaatgccgacaatgag
atggtgcagacattgctgagaatccagggagacaacaagttgtttcagcatgttgtgcacattctgtatgtacaatcgcttttgcaaggaaaata
tcctgtgaacagtgaggagatggaacttttcaatcactctctttctgaattgatgactgccaaaatgaatgattttataaactttctcaatgccgcc
gccc*accaccaccaccaccac*taagccgccctcgagaccgta

Appendix B2. Group C *B. fragilis* HtpG nucleotide sequence ordered from IDT. Nucleotide sequence for BF0237 was downloaded from the KEGG online database. Restriction sites are underlined, the 6x-His-tag is italicized, and the start codon is bolded.

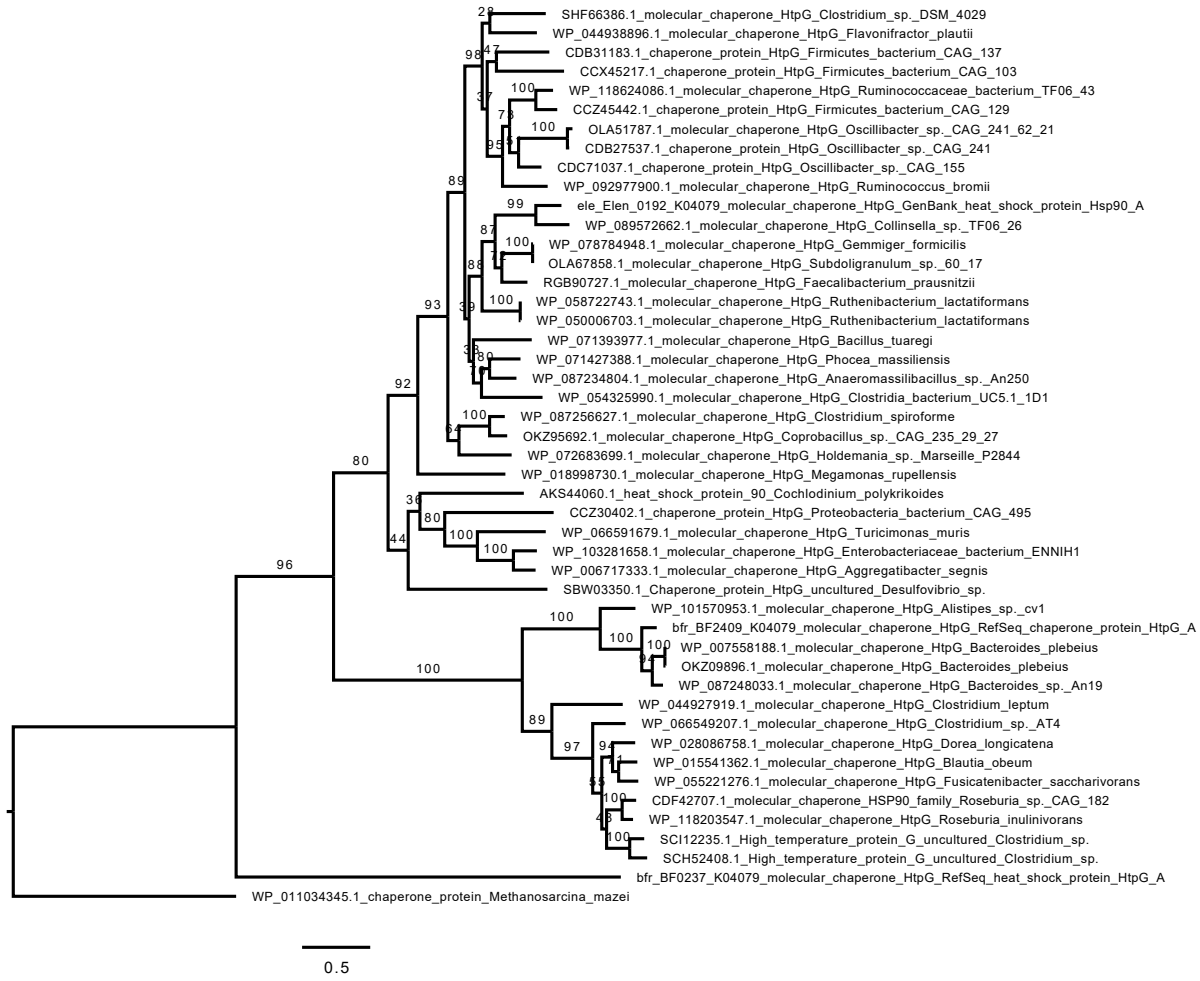
Appendix C: HtpG Phylogenetic trees



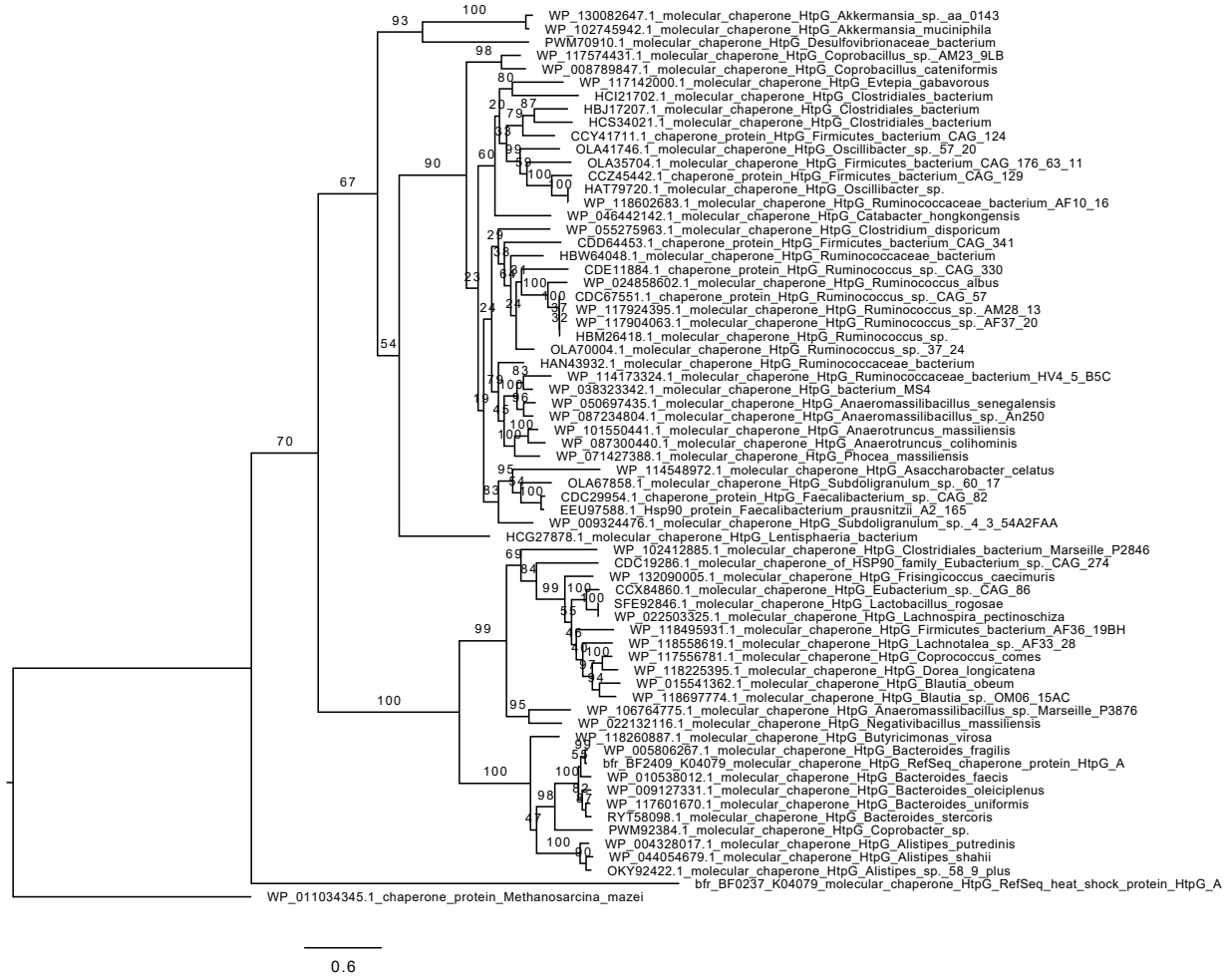
Appendix C1. Group C HtpG not identified in metagenomic sequencing of six CD patient fecal samples at baseline timepoint. From the metagenomic dataset collected for the Dunn et al. (2016) study, *HtpG* nucleotide fragments were identified from six CD patients (CD4, CD6, CD8, CD10, CD13, and CD14) at baseline. Nucleotide fragments were exposed to the blastx database and complete amino acid sequences were collected. MAFFT was used to align amino acid sequences and RAxML was used to generate gene trees. RAxML was run with GAMMA models using the LG exchangeability matrix and frequencies estimated from the data. Bootstrap support was estimated from 100 bootstrap replicates and values are shown on the tree. Bacteria contributing HtpG signal are shown on the tree. Trees are rooted with HtpG from *M. mazei.* The scale for the branch lengths gives the mean number of substitutions per amino acid.



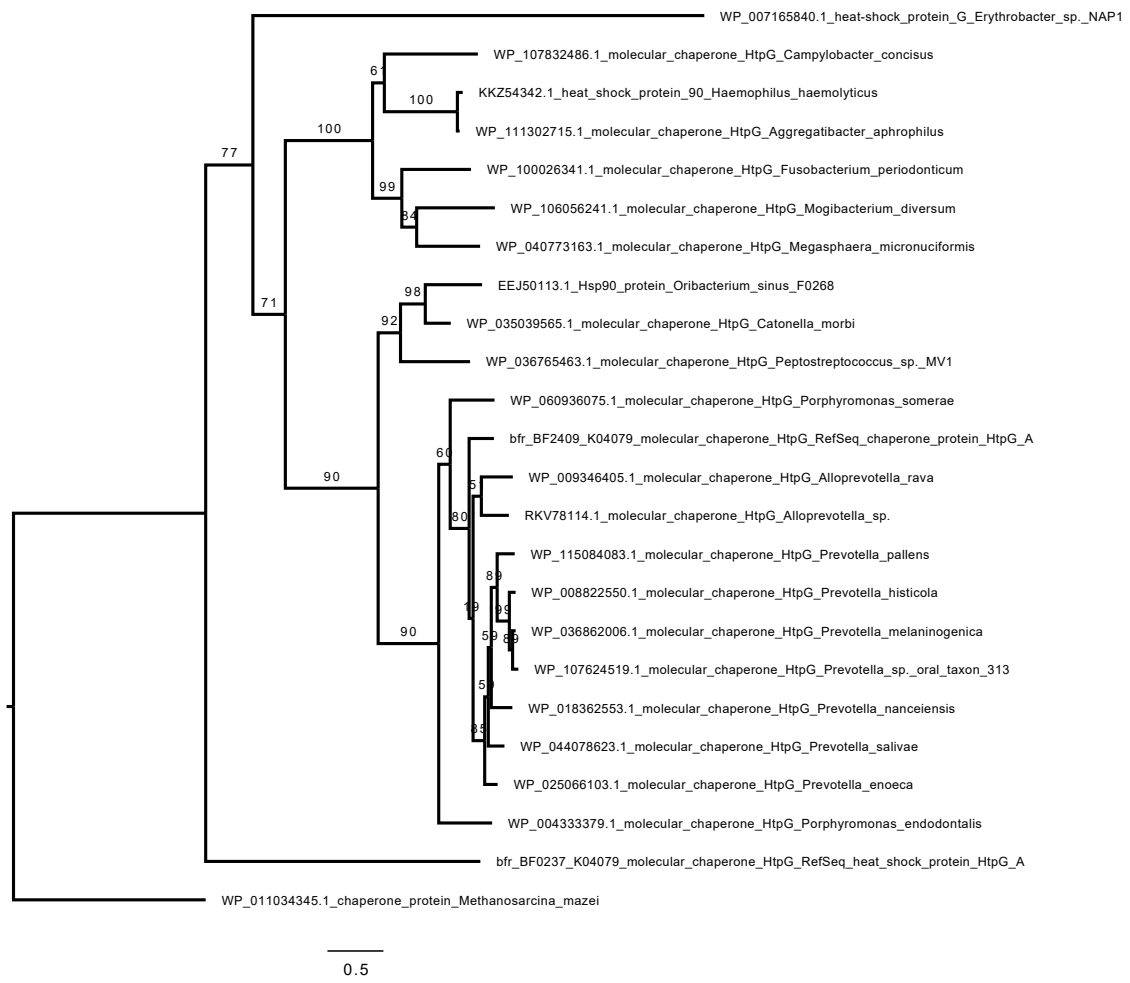
Appendix C2. Group C HtpG not identified in metagenomic sequencing of six CD patient fecal samples at 12-week timepoint. From the metagenomic dataset collected for the Dunn et al. (2016) study, *HtpG* nucleotide fragments were identified from six CD patients (CD4, CD6, CD8, CD10, CD13, and CD14) at 12-weeks. Nucleotide fragments were exposed to the blastx database and complete amino acid sequences were collected. MAFFT was used to align amino acid sequences and RAxML was used to generate gene trees. RAxML was run with GAMMA models using the LG exchangeability matrix and frequencies estimated from the data. Bootstrap support was estimated from 100 bootstrap replicates and values are shown on the tree. Bacteria contributing HtpG signal are shown on the tree. Trees are rooted with HtpG from *M. mazei.* The scale for the branch lengths gives the mean number of substitutions per amino acid.



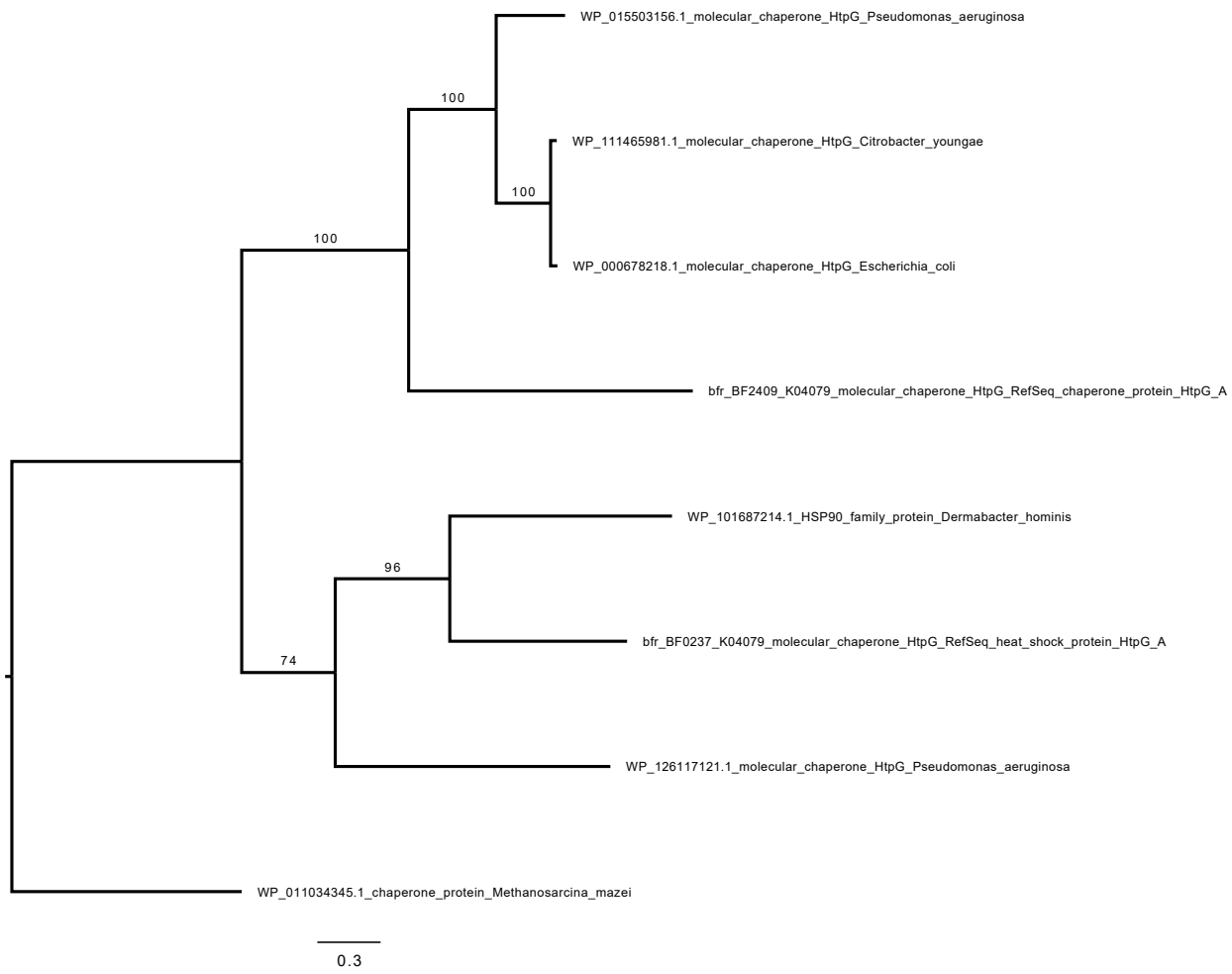
Appendix C3. Group C HtpG not identified in metagenomic sequencing of six CD patient fecal samples at 24-week timepoint. From the metagenomic dataset collected for the Dunn et al. (2016) study, *HtpG* nucleotide fragments were identified from six CD patients (CD4, CD6, CD8, CD10, CD13, and CD14) at 24-weeks. Nucleotide fragments were exposed to the blastx database and complete amino acid sequences were collected. MAFFT was used to align amino acid sequences and RAxML was used to generate gene trees. RAxML was run with GAMMA models using the LG exchangeability matrix and frequencies estimated from the data. Bootstrap support was estimated from 100 bootstrap replicates and values are shown on the tree. Bacteria contributing HtpG signal are shown on the tree. Trees are rooted with HtpG from *M. mazei.* The scale for the branch lengths gives the mean number of substitutions per amino acid.



Appendix C4. Group C HtpG not identified in metagenomic sequencing of healthy control fecal sample (sibling of patient CD10). From the metagenomic dataset collected for the Dunn et al. (2016) study, *HtpG* nucleotide fragments were identified from the healthy control. Nucleotide fragments were exposed to the blastx database and complete amino acid sequences were collected. MAFFT was used to align amino acid sequences and RAxML was used to generate gene trees. RAxML was run with GAMMA models using the LG exchangeability matrix and frequencies estimated from the data. Bootstrap support was estimated from 100 bootstrap replicates and values are shown on the tree. Bacteria contributing HtpG signal are shown on the tree. Trees are rooted with HtpG from *M. mazei.* The scale for the branch lengths gives the mean number of substitutions per amino acid.



Appendix C5. Group C HtpG not identified in metagneomic sequencing of healthy control saliva sample (Human Microbiome Project database). *HtpG* nucleotide sequences were identified from saliva samples donated by five healthy individuals on the Human Microbiome Project online database. Nucleotide fragments were exposed to the blastx database and complete amino acid sequences were collected. MAFFT was used to align amino acid sequences and RAxML was used to generate gene trees. RAxML was run with GAMMA models using the LG exchangeability matrix and frequencies estimated from the data. Bootstrap support was estimated from 100 bootstrap replicates and values are shown on the tree. Bacteria contributing HtpG signal are shown on the tree. Trees are rooted with HtpG from *M. mazei.* The scale for the branch lengths gives the mean number of substitutions per amino acid.



Appendix C6. Group C HtpG identified in metagneomic sequencing of healthy control skin sample (Human Microbiome Project database). *HtpG* nucleotide sequences were identified from skin samples donated by four healthy individuals on the Human Microbiome Project online database. Nucleotide fragments were exposed to the blastx database and complete amino acid sequences were collected. MAFFT was used to align amino acid sequences and RAxML was used to generate gene trees. RAxML was run with GAMMA models using the LG exchangeability matrix and frequencies estimated from the data. Bootstrap support was estimated from 100 bootstrap replicates and values are shown on the tree. Bacteria contributing HtpG signal are shown on the tree. Trees are rooted with HtpG from *M. mazei.* The scale for the branch lengths gives the mean number of substitutions per amino acid.