Direct Detection and Quantification of Bacterial Genes Associated with Inflammation in DNA Isolated from Stool

Ramón Gómez-Moreno[1,2], **Iraida E. Robledo**[3], and **Abel Baerga-Ortiz**[1,2]

Abel Baerga-Ortiz: abel.baerga@upr.edu

[1]Department of Biochemistry, University of Puerto Rico, Medical Sciences Campus, San Juan, Puerto Rico

[2]Molecular Sciences Building, University of Puerto Rico, Medical Sciences Campus, San Juan, Puerto Rico

[3]Department of Microbiology and Medical Zoology, University of Puerto Rico, Medical Sciences Campus, San Juan, Puerto Rico

Abstract

Although predominantly associated with health benefits, the gut microbiota has also been shown to harbor genes that promote inflammation. In this work, we report a method for the direct detection and quantification of these pro-inflammatory bacterial genes by PCR and qPCR in DNA extracted from human stool samples. PCR reactions were performed to detect (i) the *pks* island genes, (ii) *tcpC*, which is present in some strains of *Escherichia coli* and (iii) *gelE* presented in some strains of *Enterococcus faecalis*. Additionally, we screened for the presence of the following genes encoding cyclomodulins that disrupted mammalian cell division: (iv) *cdt* (which encodes the cytolethal distending toxin) and (v) *cnf*-1 (which encodes the cytotoxic necrotizing factor-1). Our results show that 20% of the samples (N = 41) tested positive for detectable amounts of *pks* island genes, whereas 10% of individuals were positive for *tcpC* or *gelE* and only one individual was found to harbor the *cnf*-1 gene. Of the 13 individuals that were positive for at least one of the pro-inflammatory genes, 5 were found to harbor more than one. A quantitative version of the assay, which used real-time PCR, revealed the pro-inflammatory genes to be in high copy numbers: up to 1.3 million copies per mg of feces for the *pks* island genes. Direct detection of specific genes in stool could prove useful toward screening for the presence of pro-inflammatory bacterial genes in individuals with inflammatory bowel diseases or colorectal cancer.

Keywords

Microbiota; Stool Samples; Gut Inflammation; PCR; Microbial Biomarkers

1. Introduction

The gut microbiota contains between 5 and 8 million active genes distributed over a great diversity of mostly uncultivable and novel species [1]–[3]. This complex community of bacteria has been shown to promote a healthy gut homeostasis as it is tasked with housekeeping functions like catabolism of complex polysaccharides, biosynthesis of vitamins and other essential nutrients, degradation of xenobiotics and the maintenance of serotonin levels in plasma [4]–[6]. Additionally, it is believed that a healthy commensal gut microbiota aids in the clearance of pathogenic bacteria and in the prevention of intestinal inflammation [7] [8]. Shifts in the composition of the resident microbiota, or microbial dysbioses, have been shown to lead to a variety of intestinal disorders [9] [10]. With the advent of powerful tools for parallel DNA sequencing, there has been a renewed interest in exploring the possible involvement of the gut microbiota in the development of intestinal disorders such as inflammatory bowel Disease (IBD), colorectal cancer (CRC) and gut-brain associated diseases [11]–[13].

Clinical correlations have been established that link chronic bacterial infections and intestinal disorders. For instance, the human pathogen *Helicobacter pylori* is widely recognized as a strong cofactor for the development of gastric adenocarcinoma [14]. Similarly, gut colonization with *Clostridium difficile* is followed by development of colitis, toxic megacolon and perforation [15].

In addition to the known gastrointestinal pathogens, there are the typically non-pathogenic members of the microbiota, which may harbor genes associated with inflammation. For instance, certain strains of commensal *Escherichia coli* have been found to harbor a genomic island encoding a colibactin polyketide synthase multienzyme (*pks* island) that induces DNA damage followed by the activation of DNA repair pathways [16]. The presence of the *pks* island induces mammalian chromosome instability in mouse and human colon cells in culture [17]. Strains of *E. coli* harboring the *pks* island were found in higher frequency in mucosal samples from individuals suffering from CRC and IBD compared with the control group [18].

Other pro-inflammatory gene products found in certain strains of *E. coli* include *tcpC*, which has been found to impair toll-like receptor (TLR) function and the cytotoxic necrotizing factor-1 (CNF-1), an activator of oncogenic Rho GTPases [19]–[21]. Additionally, there are certain strains of commensal *Enterococcus* species that are not pathogenic but contain a subset of the virulence genes including *gelE*—a gene encoding a collagen-degrading metalloprotease which has been found to induce inflammation in a mouse model [22].

It has been established that these pro-inflammatory genes promote the colonization of pathogenic strains of bacteria. Substantially less clear is the effect of these pro-inflammatory genes when harbored by non-pathogenic commensal bacteria in the microbiota over a period of time [23]. Additionally, the distribution and frequency of most of these pro-inflammatory genes in the human population are not known and methods have not been implemented for the quick and direct detection of these genes in easily obtained stool samples.

In this work, we have developed a rapid assay for the detection of pro-inflammatory genes directly in a small stool sample (0.5 g). The assay can be performed in diagnostic format (classical PCR) or a quantitative format (rtPCR). Our results show that pro-inflammatory genes are highly prevalent (up to 20% of the samples tested) and in high enough copy number (up to 1.3 million copies per mg stool) for direct detection in a small stool sample. The assays described in this report could prove useful in future explorations of the relationship between the presence of these pro-inflammatory genes and intestinal diseases.

2. Materials and Methods

2.1. Stool Samples

Anonymous human stool samples were donated by a clinical laboratory (Laboratorio Clínico El Monte, San Juan, Puerto Rico, N = 41). Samples were stored at −80°C.

2.2. DNA Extraction

Bacterial DNA was extracted from 41 human stool samples using the QIAamp® DNA Stool Mini Kit (QIAGEN). Approximately 0.2 g stool was resuspended in 1.4 mL of ASL buffer. The bacterial DNA was extracted following the manufacturer's protocol and eluted in 200 μl of the buffer included in the kit. A total of 5 μl of DNA extract was used as template for the ensuing PCR reactions.

2.3. Detection of Bacterial Genes by PCR

A volume of 5 ul of undiluted fecal DNA extract (from 0.2 mg stool) was used and the PCR conditions are summarized in Table 1. All PCR reactions begin with an initial denaturation of 1 minute at 94 C followed by 30 cycles of denaturation (30 seconds at 94°C), annealing (30 seconds at different temperatures shown in Table 1) and extension (3 minutes at 68°C). All reactions were finalized with a long extension step of 10 minutes at 72°C to ensure maximum detection. The detection of *pks island* genes was performed using the primer for the amplification of *clbN* genes that were described previously [23]. The detection of *gelE* was carried out using gelE-specific primers that were also described previously [24]. For the detection of *tcpC* and *cnf*-1, oligonucleotide primers were designed using the software Primer 3 (http://simgene.com/Primer3) on the nucelotide sequences for *E. coli* ECOR63 (GenBank: GQ903014.1) and *E. coli* MS 153-1 (NCBI: ZP_16446903.1). For the detection of *cdt*, primers were designed using the *cdtB* subunit from *Campylobacter jejuni subsp. jejuni* 81–176 (GenBank: AAB06708.1).

2.4. Purification of PCR Products

The PCR products were separated on a 2% agarose gel and extracted from the agarose gel using the QIAquick Gel Extraction Kit (QIAGEN).

2.5. DNA Sequence Confirmation by the Sanger Method

The DNA sequence for the amplified fragments was verified using the Sanger method for DNA sequencing at the RCMI Center for Genomics in Health Disparities and Rare Diseases at the University of Puerto Rico. The purified amplicon was sent for sequencing using the forward oligonucleotide primer for each of the products.

2.6. Estimation of Gene Copy Number by Quantitative PCR (qPCR)

In order to provide a quantitative estimate on the number of copies of each of the pro-inflammatory genes, we measured the PCR reaction in real time and compared the results for the DNA extracted from stool samples with known amounts of DNA extracted from isolated bacteria known to contain the genes. The qPCR reaction was performed with 6.8 μl of stool DNA extract and the QuantiTect SYBR Green PCR kit (Qiagen) and was read on a Eppendorf Mastercycler Realplex.

2.7. Calibration Curve

A calibration curve was built for each gene by performing real-time PCR on the DNA extracts from a bacterial isolate known to contain at least a copy of the bacterial gene of interest. In order to identify strains that could serve as positive controls (or method calibrants) for the *pks* island, *tcpC*, and *cnf* genes, we screened a collection of clinical isolates of *E. coli* that were part of a nosocomial infection surveillance study [25] and found several *E. coli* isolates that tested positive for one of the genes. The DNA extracted from these strains was used as positive control in the PCR reaction, which was done in triplicate. The positive control for *gelE* was *Enterococcus faecalis* strain H32, an environmental isolate known to contain the *gelE* virulence gene and kindly donated by Dr. Luis Ríos-Hernández from University of Puerto Rico, Mayagüez Campus. The genome of *E. coli* contains 4.6 million base pairs and weighs 5.1×10^{-15} g. The genome of *E. faecalis*, somewhat smaller, weighs 3.6×10^{-15} g. Thus, the DNA copy number conversion can be determined from a measurement of absorbance at 260nm according to the following relationship:

$$\#\mathrm{DNA\,copies}/\mu l = A260 * (50 \times 10^{-9}\,\mathrm{g\,DNA}/\mu l)/(GW) \quad (1)$$

where A260 is the absorbance of the sample at 260 nm and GW is the weight in grams of the entire genome of the organism (5.1×10^{-15} g for *E. coli* and 3.6×10^{-15} g for *E. faecalis*).

For example, the *pks* island gene standards were made with known amounts of *E. coli* DNA. Samples containing 384 ng/μl (75,294,120 copies per μl), 100 ng/μl (19,607,843 copies per μl), 10 ng/μl (1,960,784 copies per μl), 1 ng/μl (196,078 copies per μl) and 0.1 ng/μl (19,607 copies per μl) were used as templates of the qPCR reaction monitored in real time.

The correlation coeffcient (χ^2) for each calibration curve was determined by the least squares method and the percentage efficiency of the PCR reaction calculated by implementing the equation:

$$\mathrm{Efficiency} = -1 + 10^{(-1/\mathrm{slope})} \quad (2)$$

Typical positive values for the *pks* island range between 2800 – 1,300,000 copies per μl.

3. Results

3.1. Detection and Prevalence of Pro-Inflammatory Genes in Stool Samples

A total of 41 anonymous human stool samples were analyzed for the presence of the genes for *pks* island, *tcpC*, *gelE*, and *cnf-1*. The presence of bacteria was confirmed in all of the samples by amplification of the 16S-rRNA gene (data not shown). The presence of the *pks* island genes was established by the amplification of a DNA fragment of 733 bp (Figure 1(a)). A total of 8 samples out of 41 were found to contain the gene for *pks* island (20%). The other pro-inflammatory gene, *tcpC*, was detected in a total of 7 samples (17%) as evidenced by the presence of a band of 216 bp (Figure 1(b)) and the enterococcal *gelE* was found in 3 samples easily detected by a band of 213 bp (7%). The gene for cytotoxic necrotizing factor-1 (*cnf*-1) was only found in a single sample while the gene for the cytolethal distending toxin (*cdt*) was not found in this group of samples. The results from the PCR screens are summarized in Table 2.

A total of thirteen (13) samples were found to contain at least one of the four pro-inflammatory genes. Of these 13 samples, a total of eight (8) samples were found to contain only one of the genes and four samples were found to contain two pro-inflammatory genes. One result, corresponding to anonymous sample "Hu037" was found to contain three genes: *pks* island, *tcpC* and *gelE* (Table 3). Future work will be aimed at establishing whether these individuals harboring multiple pro-inflammatory bacterial genes have a higher risk of developing intestinal disorders.

The DNA sequences for the PCR amplified pro-inflammatory bacterial genes were confirmed by performing Sanger sequencing on the gel-purified PCR products. By comparing the DNA sequences found in individual samples, we were able to make a preliminary assessment of the sequence variability for the pro-inflammatory bacterial genes. While the DNA sequences for the amplicons generated from *tcpC* and *gelE* were exactly the same in all positive stool samples, the DNA sequence for the *pks* island amplicon showed variability around sequence positions 8312 and 8383 (numbering according to GeneBank accession No. AM229678). The amplicon generated from the *pks* island in samples Hu033, Hu037, Hu038 contains three distinct variants (Figure 2). Stool sample labeled Hu033 features A and G in positions 8312 and 8383, respectively. Meanwhile, the stool samples Hu037 and Hu038 have a G-A and A-A in those positions.

3.2. Quantification of Pro-Inflammatory Genes in Stool

From the PCR results visualized on an agarose gel, it was clear that some samples registered a more intense band than others, raising the possibility that some individuals harbor a higher content of bacterial pro-inflammatory genes than others. In order to provide a quantitative estimate of the number of copies of the pro-inflammatory genes per patient, we performed the PCR assay with real-time detection (Figure 3). A standard curve that was made with known amounts of DNA copies as described in the methods section, enabled the quantification of DNA copy number for each of the genes in the 41 stool samples. The number of copies of *pks* island among those who tested positive fluctuated between 2800 and 1.3 million copies per mg of stool analyzed and the number of *gelE* copies fluctuated

between 6000 – 260,000 copies per mg of stool. The amount of *tcpC* DNA was less variable fluctuating between 100,000 – 800,000 copies per mg of stool. Future work will be aimed at establishing whether a correlation exists between a high copy number for these pro-inflammatory genes and a higher risk of developing gastrointestinal disorders.

4. Discussion

The elucidation of the role of the gut microbiota in human health is an active area of research [26]–[28]. A number of efforts worldwide have served to elucidate important associations between the microbiota and health [29]–[31]. For instance, an unequivocal link has been established between the gut microbiota and obesity [32] [33]. Additionally, the presence of pathogens such as *Helicobacter pylori* and *Clostridium difficile* has been found to correlate with inflammation and ultimately with the development of gastric cancer [14] [15]. More recently, there have been reports that the gram-negative bacterium of the genus *Fusobacterium* is more highly represented in colorectal tumor tissue than in normal mucosal tissue from the same patients [34] [35].

In this report, instead of focusing on the presence of bacterial species, we have developed a quantitative assay to detect specific bacterial genes that have been associated with inflammation, directly in a small stool sample. Some of these genes in this report were found to be present in as many as 20% of the stool samples and in copy numbers as high as 1.3 million copies per mg of stool. The assay described in this report provides a quick method for the screening of stool samples to ascertain the presence of pro-inflammatory genes in case-control studies involving individuals with colorectal cancer or inflammatory bowel diseases. The genes in this report, namely *pks* island, *tcpC*, *cnf*-1, *gelE* and *cdt*, have all been the subject of scrutiny for their possible crosstalk with inflammatory processes. Here we discuss each of them separately:

pks Island

Of the genes tested, the *pks* island was the most prevalent and was found in high copy numbers ($>10^6$ copies per mg stool). The *pks* genomic island encodes a number of multidomain enzymes termed polyketide synthases that typically synthesize secondary metabolites in numerous microbial species, although the actual compound made by this multienzyme has not been chemically characterized [16]. The *pks* island has been found in pathogenic and in non-pathogenic strains of *E. coli* [23]. Although the *pks* island is not pro-inflammatory per se, the presence of this genomic island in some strains of *E. coli* has been implicated in the formation of colorectal tumors in inflammation-prone mice, which lack the inteleukin-10 gene [18]. That study also reported that the *pks* island was more highly represented in the intestinal mucosa from colorectal cancer (CRC) patients than in non-CRC controls [18]. In the present study, the *pks* island genes were found in 20% of the stool samples analyzed, a frequency consistent with previous findings reporting the *pks* island to be present in 20% of strains isolated from human mucosa samples from non-cancer donors [18] and in 32% of *E. coli* strains isolated from the feces of hospitalized veterans [23]. Although our assay is carried out on total DNA isolated from a stool sample, without strain isolation or cultivation, our results are similar to those obtained using the more conventional culture-dependent method.

tcpC

Similarly, the *tcpC* gene has been found in strains of *E. coli* associated with severe kidney infection [36]. *tcpC* subverts toll-like receptor-dependent inflammation by binding directly to the receptor, resulting in the formation of kidney abscesses in mice infected with strains of uropathogenic *E. coli* [19]. However, there are no reports of the *tcpC* gene in the commensal *E. coli* strains that inhabit the GI tract. Much to our surprise, seven (7) of our fecal samples were found to harbor the tcpC gene in high enough copy number to detect by PCR.

cnf-1

The *cnf*-1 gene encodes an enzyme, the cytotoxic necrotizing factor-1, which is found in strains of *E. coli* commonly associated with urinary tract infections [37]. The Cnf enzyme deaminates Rho GTPase in glutamine 63 causing it to be locked in the active GTP-bound state. Although the *cnf*-1 gene has been found in bacterial isolates from feces of children with diarrhea, this is the first report of this gene in high enough copies to be detected directly in stool [38]. Only one (1) sample in the present study was found to contain the *cnf*-1 gene.

gelE

GelE gene encodes a metallo-collagenase from the genus *Enterococcus*. The GelE collagenase has been found to promote chronic inflammation in mice lacking the *interleukin*-10 gene [22]. More recently, *gelE* gene was found in higher frequency in bacterial isolates from the feces of children with IBD vs. children without IBD [39]. Here we report 3 stool samples with detectable levels of *gelE* gene directly in stool.

With the growth of affordable parallel sequencing methods, scientists have gained a whole new view of the diversity of bacterial species (or taxa) in a complex community such as the gut microbiota [40] [41]. One important line of experimentation centers on building a comprehensive list of bacterial species present in the community for comparison between disease cases and controls. However, most community profiling efforts rely on molecular markers (such as the 16S rRNA gene) and do not necessarily detect additional genes that may have been acquired in time. For instance, the bacterial community profile for any stool sample would likely reveal the presence of *E. coli*, but it would not necessarily show whether the *E. coli* strains in the sample contain pro-inflammatory genes (*tcpC*, *pks* island, etc.). In addition, *E. coli* is not the only species that harbor the *pks* island, as the genes can also be carried by other bacteria, such as *Klebsiella pneumoniae*, *Enterobacter aerogenes* and *Ci-trobacter koseri* [42].

The assay described in this report is meant to complement ongoing efforts, whether by bacterial community profiling or by shotgun metagenomics, to ascertain the presence of harmful or protective genes, in fecal samples regardless of which species carries the genes. Clearly, other complementary assays will have to be developed to determine whether the pro-inflammatory bacterial genes are being transcribed in the gut, or whether the gene products are being synthesized in active form. But the development of a quick and quantitative assay for the detection of pro-inflammatory genes in stool samples is a first step

toward the quick interrogation of stool samples in large-scale case-control studies involving IBD and CRC patients.

5. Conclusion

We report the detection, directly in DNA extracted from a fecal sample, of several genes of bacterial origin which had been previously associated with inflammation. Although the presence of these genes had been ascertained in strains of bacteria known to colonize other tissues, this is the first report of these genes being found in the normal commensal flora. The assay described in this report provides a quick and inexpensive method for the analysis of a large number of samples from case-control studies involving patients of IBD or CRC.

Acknowledgments

References

1. Balter M. Taking Stock of the Human Microbiome and Disease. Science. 2012; 336:1246–1247. http://dx.doi.org/10.1126/science.336.6086.1246. [PubMed: 22674333]

2. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA. Diversity of the Human Intestinal Microbial Flora. Science. 2005; 308:1635–1638. http://dx.doi.org/10.1126/science.1110591. [PubMed: 15831718]

3. Blaser M, Bork P, Fraser C, Knight R, Wang J. The Microbiome Explored: Recent Insights and Future Challenges. Nature Reviews Microbiology. 2013; 11:213–217. http://dx.doi.org/10.1038/nrmicro2973.

4. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto J, Hansen T, Le Paslier D, Linneberg A, Nielsen H, Bjorn, Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Li S, Jian M, Zhou Y, Li Y, Zhang X, Li S, Qin N, Yang H, Wang J, Brunak S, Dore J, Guarner F, Kristiansen K, Pedersen O, Parkhill J, Weissenbach J, Bork P, Dusko Ehrlich S, Wang J. A Human Gut Microbial Gene Catalogue Established by Metagenomic Sequencing. Nature. 2010; 464:59–65. http://dx.doi.org/10.1038/nature08821. [PubMed: 20203603]

5. Wikoff WR, Anfora AT, Liu J, Schultz PG, Lesley SA, Peters EC, Siuzdak G. Metabolomics Analysis Reveals Large Effects of Gut Microflora on Mammalian Blood Metabolites. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106:3698–3703. http://dx.doi.org/10.1073/pnas.0812874106. [PubMed: 19234110]

6. Ursell LK, Knight R. Xenobiotics and the Human Gut Microbiome: Metatranscriptomics Reveal the Active Players. Cell Metabolism. 2013; 17:317–318. http://dx.doi.org/10.1016/j.cmet.2013.02.013. [PubMed: 23473028]

7. Sperandio V. Microbiology. Virulence or competition? Science. 2012; 336:1238–1239. http://dx.doi.org/10.1126/science.1223303. [PubMed: 22582015]

8. Mazmanian SK, Round JL, Kasper DL. A Microbial Symbiosis Factor Prevents Intestinal Inflammatory Disease. Nature. 2008; 453:620–625. http://dx.doi.org/10.1038/nature07008. [PubMed: 18509436]

9. DuPont AW, DuPont HL. The Intestinal Microbiota and Chronic Disorders of the Gut. Nature Reviews Gastroenterology and Hepatology. 2011; 8:523–531. http://dx.doi.org/10.1038/nrgastro.2011.133.

10. Manichanh C, Borruel N, Casellas F, Guarner F. The Gut Microbiota in IBD. Nature Reviews Gastroenterology and Hepatology. 2012; 9:599–608.

11. Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR. Molecular-Phylogenetic Characterization of Microbial Community Imbalances in Human Inflammatory Bowel Diseases. Proceedings of the National Academy of Sciences of the United States of America. 2007; 104:13780–13785. http://dx.doi.org/10.1073/pnas.0706625104. [PubMed: 17699621]

12. Sears CL, Garrett WS. Microbes, Microbiota, and Colon Cancer. Cell Host & Microbe. 2014; 15:317–328. http://dx.doi.org/10.1016/j.chom.2014.02.007. [PubMed: 24629338]

13. Kinross JM, Darzi AW, Nicholson JK. Gut Microbiome-Host Interactions in Health and Disease. Genome Medicine. 2011; 3:14. http://dx.doi.org/10.1186/gm228. [PubMed: 21392406]

14. Parsonnet J, Friedman GD, Vandersteen DP, Chang Y, Vogelman JH, Orentreich N, Sibley RK. *Helicobacter pylori* Infection and the Risk of Gastric Carcinoma. New England Journal of Medicine. 1991; 325:1127–1131. http://dx.doi.org/10.1056/NEJM199110173251603. [PubMed: 1891020]

15. Trudel JL. *Clostridium difficile* Colitis. Clinics in Colon and Rectal Surgery. 2007; 20:13–17. http://dx.doi.org/10.1055/s-2007-970195. [PubMed: 20011356]

16. Nougayrède J, Homburg S, Taieb F, Boury M, Brzuszkiewicz E, Gottschalk G, Buchrieser C, Hacker J, Dobrindt U, Oswald E. *Escherichia coli* Induces DNA Double-Strand Breaks in Eukaryotic Cells. Science. 2006; 313:848–851. http://dx.doi.org/10.1126/science.1127059. [PubMed: 16902142]

17. Cuevas-Ramos G, Petit CR, Marcq I, Boury M, Oswald E, Nougayrède J. *Escherichia coli* Induces DNA Damage *in Vivo* and Triggers Genomic Instability in Mammalian Cells. Proceedings of the National Academy of Sciences of the United States of America. 2010; 107:11537–11542. http://dx.doi.org/10.1073/pnas.1001261107. [PubMed: 20534522]

18. Arthur JC, Perez-Chanona E, Mühlbauer M, Tomkovich S, Uronis JM, Fan T, Campbell BJ, Abujamel T, Dogan B, Rogers AB, Rhodes JM, Stintzi A, Simpson KW, Hansen JJ, Keku TO, Fodor AA, Jobin C. Intestinal Inflammation Targets Cancer-Inducing Activity of the Microbiota. Science. 2012; 338:120–123. http://dx.doi.org/10.1126/science.1224820. [PubMed: 22903521]

19. Yadav M, Zhang J, Fischer H, Huang W, Lutay N, Cirl C, Lum J, Miethke T, Svanborg C. Inhibition of TIR Domain Signaling by TcpC: MyD88-Dependent and Independent Effects on *Escherichia coli* Virulence. PLoS Pathogens. 2010; 6:e1001120. http://dx.doi.org/10.1371/journal.ppat.1001120. [PubMed: 20886104]

20. Fabbri A, Travaglione S, Ballan G, Loizzo S, Fiorentini C. The Cytotoxic Necrotizing Factor 1 from *E. coli*: A Janus Toxin Playing with Cancer Regulators. Toxins (Basel). 2013; 5:1462–1474. http://dx.doi.org/10.3390/toxins5081462. [PubMed: 23949007]

21. Garcia TA, Ventura CL, Smith MA, Merrell DS, O'Brien AD. Cytotoxic Necrotizing Factor 1 and Hemolysin from Uropathogenic *Escherichia coli* Elicit Different Host Responses in the Murine Bladder. Infection and Immunity. 2013; 81:99–109. http://dx.doi.org/10.1128/IAI.00605-12. [PubMed: 23090961]

22. Steck N, Hoffmann M, Sava IG, Kim SC, Hahne H, Tonkonogy SL, Mair K, Krueger D, Pruteanu M, Shanahan F, Vogelmann R, Schemann M, Kuster B, Sartor RB, Haller D. *Enterococcus faecalis* Metalloprotease Compromises Epithelial Barrier and Contributes to Intestinal Inflammation. Gastroenterology. 2011; 141:959–971. http://dx.doi.org/10.1053/j.gastro.2011.05.035. [PubMed: 21699778]

23. Johnson JR, Johnston B, Kuskowski MA, Nougayrede J, Oswald E. Molecular Epidemiology and Phylogenetic Distribution of the *Escherichia coli* pks Genomic Island. Journal of Clinical Microbiology. 2008; 46:3906–3911. http://dx.doi.org/10.1128/JCM.00949-08. [PubMed: 18945841]

24. Vankerckhoven V, Van Autgaerden T, Vael C, Lammens C, Chapelle S, Rossi R, Jabes D, Goossens H. Development of a Multiplex PCR for the Detection of *asa*1, *gelE*, *cylA*, *esp*, and *Hyl* Genes in Enterococci and Survey for Virulence Determinants among European Hospital Isolates of

Enterococcus faecium. Journal of Clinical Microbiology. 2004; 42:4473–4479. http://dx.doi.org/10.1128/JCM.42.10.4473-4479.2004. [PubMed: 15472296]

25. Robledo IE, Aquino EE, Vazquez GJ. Detection of the KPC Gene in *Escherichia coli*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, and *Acinetobacter baumannii* during a PCR-Based Nosocomial Surveillance Study in Puerto Rico. Antimicrobial Agents and Chemotherapy. 2011; 55:2968–2970. http://dx.doi.org/10.1128/AAC.01633-10. [PubMed: 21444702]

26. Hood L. Tackling the Microbiome. Science. 2012; 336:1209. http://dx.doi.org/10.1126/science.1225475. [PubMed: 22674329]

27. Zaneveld J, Turnbaugh PJ, Lozupone C, Ley RE, Hamady M, Gordon JI, Knight R. Host-Bacterial Coevolution and the Search for New Drug Targets. Current Opinion in Chemical Biology. 2008; 12:109–114. http://dx.doi.org/10.1016/j.cbpa.2008.01.015. [PubMed: 18280814]

28. Kuczynski J, Costello EK, Nemergut DR, Zaneveld J, Lauber CL, Knights D, Koren O, Fierer N, Kelley ST, Ley RE, Gordon JI, Knight R. Direct Sequencing of the Human Microbiome Readily Reveals Community Differences. Genome Biology. 2010; 11:210. [PubMed: 20441597]

29. De Palma G, Collins SM, Bercik P, Verdu EF. The Microbiota-Gut-Brain Axis in Gastrointestinal Disorders: Stressed Bugs, Stressed Brain or Both? Journal of Physiology. 2014; 592:2989–2997. http://dx.doi.org/10.1113/jphysiol.2014.273995. [PubMed: 24756641]

30. Belkaid Y, Hand TW. Role of the Microbiota in Immunity and Inflammation. Cell. 2014; 157:121–141. http://dx.doi.org/10.1016/j.cell.2014.03.011. [PubMed: 24679531]

31. Haiser HJ, Turnbaugh PJ. Is It Time for a Metagenomic Basis of Therapeutics? Science. 2012; 336:1253–1255. http://dx.doi.org/10.1126/science.1224396. [PubMed: 22674325]

32. Bäckhed F, Ding H, Wang T, Hooper LV, Koh GY, Nagy A, Semenkovich CF, Gordon JI. The Gut Microbiota as an Environmental Factor that Regulates Fat Storage. Proceedings of the National Academy of Sciences of the United States of America. 2004; 101:15718–15723. http://dx.doi.org/10.1073/pnas.0407076101. [PubMed: 15505215]

33. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An Obesity-Associated Gut Microbiome with Increased Capacity for Energy Harvest. Nature. 2006; 444:1027–1031. http://dx.doi.org/10.1038/nature05414. [PubMed: 17183312]

34. McCoy AN, Araujo-Perez F, Azcarate-Peril A, Yeh JJ, Sandler RS, Keku TO. *Fusobacterium* Is Associated with Colorectal Adenomas. PLoS ONE. 2013; 8:e53653. http://dx.doi.org/10.1371/journal.pone.0053653. [PubMed: 23335968]

35. Keku TO, McCoy AN, Azcarate-Peril AM. *Fusobacterium* spp. and Colorectal Cancer: Cause or Consequence? Trends in Microbiology. 2013; 21:506–508. http://dx.doi.org/10.1016/j.tim.2013.08.004. [PubMed: 24029382]

36. Cirl C, Wieser A, Yadav M, Duerr S, Schubert S, Fischer H, Stappert D, Wantia N, Rodriguez N, Wagner H, Svanborg C, Miethke T. Subversion of Toll-Like Receptor Signaling by a Unique Family of Bacterial Toll/Interleukin-1 Receptor Domain-Containing Proteins. Nature Medicine. 2008; 14:399–406. http://dx.doi.org/10.1038/nm1734.

37. Travaglione S, Loizzo S, Ballan G, Fiorentini C, Fabbri A. The *E. coli* CNF1 as a Pioneering Therapy for the Central Nervous System Diseases. Toxins (Basel). 2014; 6:270–282. http://dx.doi.org/10.3390/toxins6010270. [PubMed: 24402235]

38. Elliott SJ, Srinivas S, Albert MJ, Alam K, Robins-Browne RM, Gunzburg ST, Mee BJ, Chang BJ. Characterization of the Roles of Hemolysin and Other Toxins in Enteropathy Caused by Alpha-Hemolytic *Escherichia coli* Linked to Human Diarrhea. Infection and Immunity. 1998; 66:2040–2051. [PubMed: 9573087]

39. Golinska E, Tomusiak A, Gosiewski T, Wiecek G, Machul A, Mikolajczyk D, Bulanda M, Heczko PB, Strus M. Virulence Factors of Enterococcus Strains Isolated from Patients with Inflammatory Bowel Disease. World Journal of Gastroenterology. 2013; 19:3562–3572. http://dx.doi.org/10.3748/wjg.v19.i23.3562. [PubMed: 23801857]

40. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R. Ultra-High-Throughput Microbial Community Analysis on the Illumina HiSeq and MiSeq Platforms. ISME Journal. 2012; 6:1621–1624. http://dx.doi.org/10.1038/ismej.2012.8. [PubMed: 22402401]

41. Navas-Molina JA, Peralta-Sanchez JM, Gonzalez A, McMurdie PJ, Vazquez-Baeza Y, Xu Z, Ursell LK, Lauber C, Zhou H, Song SJ, Huntley J, Ackermann GL, Berg-Lyons D, Holmes S, Caporaso JG, Knight R. Advancing Our Understanding of the Human Microbiome Using QIIME. Methods in Enzymology. 2013; 531:371–444. http://dx.doi.org/10.1016/B978-0-12-407863-5.00019-8. [PubMed: 24060131]

42. Putze J, Hennequin C, Nougayrede J, Zhang W, Homburg S, Karch H, Bringer M, Fayolle C, Carniel E, Rabsch W, Oelschlaeger TA, Oswald E, Forestier C, Hacker J, Dobrindt U. Genetic Structure and Distribution of the Colibactin Genomic Island among Members of the Family *Enterobacteriaceae*. Infection and Immunity. 2009; 77:4696–4703. http://dx.doi.org/10.1128/IAI.00522-09. [PubMed: 19720753]
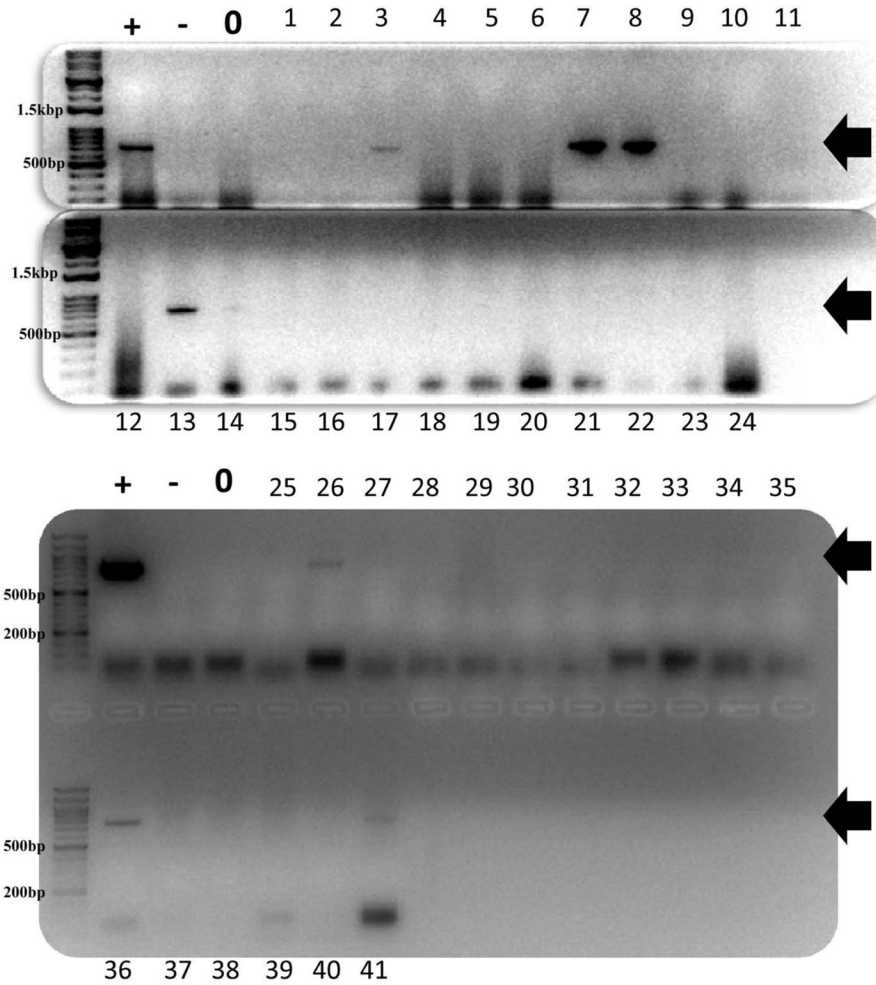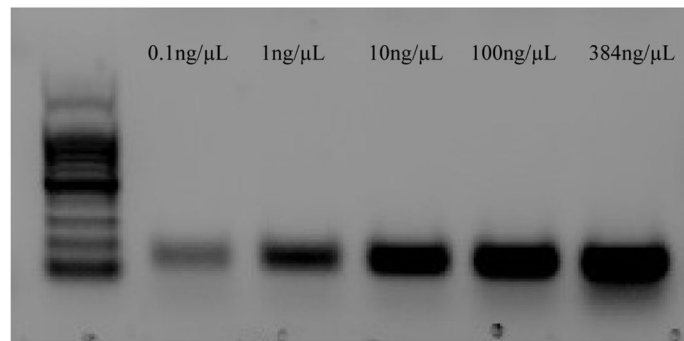
Figure 1.
PCR amplification of the *pks* island genes was visualized on a 2% agarose gel in which the first four lanes contain the DNA size marker, a positive control (+), a negative control (−) and a blank (0), respectively. The DNA size markers used for samples 1 – 24 is 1kb Opti-DNA Marker from ABM band for sample 25 – 41 the 100 bp Opti-DNA Marker by ABM. The presence of the *pks* island gene can be established by the presence of a band at 733 bp (black arrows). The data for the 41 samples is shown in two different gels with eight positives (3, 7, 8, 13, 14, 26, 36 and 41).
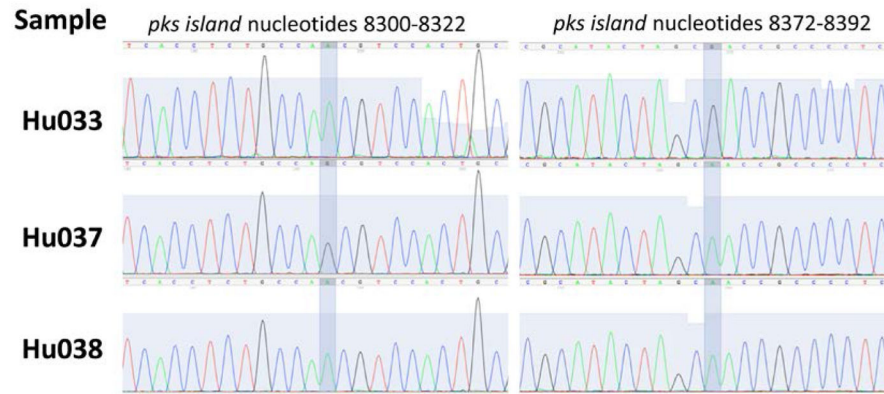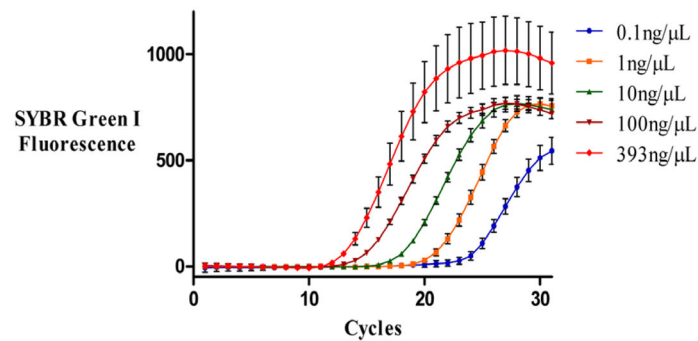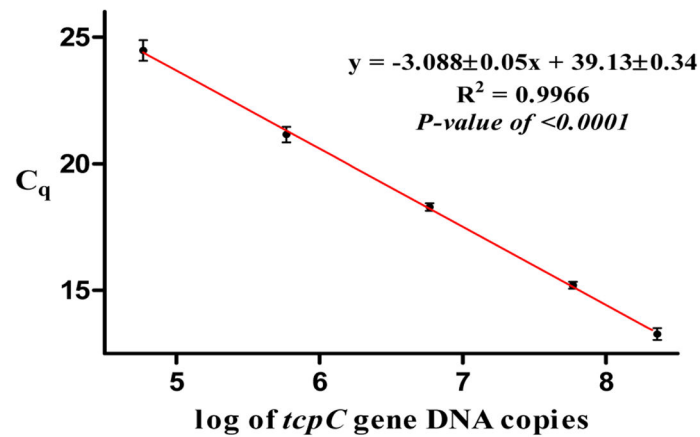
Figure 2.
The DNA sequences for all the fragments generated in this study were confirmed by the Sanger method. Interestingly, three distinct variants of the *pks* island sequence were identified in three different samples (Hu033, Hu037 and Hu038). The sequencing chromatograms corresponding to regions 8300 – 8322 and 8372 – 8392 of the *pks* island gene sequence (Ge-neBank Accession No. AM229678) are shown for all three samples. Sequence variability was observed in positions 8312 and 8383 of this amplicon.

(a)

(b)

$$y = -3.088 \pm 0.05x + 39.13 \pm 0.34$$
$$R^2 = 0.9966$$
$$P\text{-value of} < 0.0001$$

(c)

Figure 3.

For the quantification of DNA copy number, a standard curve was generated for each of the pro-inflammatory genes as described in the methods section. Standards containing 393 ng/µl (77,058,825 copies per µl), 100 ng/µl (19,607,843 copies per µl), 10 ng/µl (1,960,784 copies per µl), 1 ng/µl (196,078 copies per µl) and 0.1 ng/µl (19,607 copies per µl) were used to make the calibration curve. Typical copy number estimates for *tcpC*-positive samples ranged from 100,000 – 800,000 copies per µl. In this figure we show (a) the agarose gel electrophoresis for the PCR amplification of *tcpC* gene using different known amounts of *E.*

coli DNA in the reaction mixture, (b) the qPCR reaction for the standards monitored in real time and (c) the calibration curve for the quantification of *tcpC* copies.

Table 1

Summary of PCR conditions for the detection of a panel of pro-inflammatory bacterial genes.

Bacterial Gene	Annealing Temp	Primer Sequences (5′ to 3′)	Product Size (bp)
pks island	63°C	F: TCGATATAGTCACGCCACCA R: GTCAAGCGAGCATACGAACA	733
tcpC	56°C	F: TCGGCGATAGCTTAAGGAGA R: CCGCCAAATAATGGCTGTAT	216
gelE [24]	56°C	F: TATGACAATGCTTTTTGGGAT R: AGATGCACCCGAAATAATATA	213
cnf-1	56°C	F: AGCGTGCAATCTATCCGTATTT R: TGGAATTTCCCCAGTATAGGTG	173
cdt	50°C	F: AGATATCTTAATGATACAAGAAGCA R: TGAAACTATAGCTAAATTTACACG	196

Table 2

Summary of the results from the PCR screening of pro-inflammatory bacterial genes from anonymous stool samples (N = 41). A DNA calibration curve was made by real-time PCR measurements of samples with known DNA concentration and each point was measured in triplicate.

Bacterial Gene	# Positives	%	Calibration Curve (χ^2)	Efficiency (%)	Copy Number in Positives per mg Stool
pks island	8	20	0.9991	114.092	2800 – 1,300,000
tcpC	7	17	0.9966	99.09	100,000 – 800,000
gelE	3	7	0.9875	58.75	6000 – 260,000
cnf-1	1	2	0.9987	104.69	43,000
cdt	0	0	N/A	N/A	N/A

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 3

Stools samples positive for pro-inflammatory bacterial genes.

Samples	*pks* Island	*tcpC*	*gelE*	*cnf-1*	*cdt*
HUO32	+	–	–	–	–
HUO37	+	+	+	–	–
HUO38	+	–	–	–	–
HUO40	–	+	–	–	–
HUO33	+	+	–	–	–
HUO47	+	–	–	–	–
PR014	–	–	+	–	–
LEMB3_2	+	+	–	–	–
LEMB3_6	–	–	+	–	–
LEMB3_8	–	+	–	–	–
LEMB3_12	+	–	–	+	–
LEMB3_15	–	+	–	–	–
LEMB3_17	+	+	–	–	–