Toll-like Receptor 9 (TLR9) activation and the innate immune response to microbial and human DNA

Dissertation for Obtaining the Doctoral Degree

of Natural Sciences (Dr. rer. nat.)

Faculty of Natural Sciences

University of Hohenheim

Institute of Nutritional Sciences

Submitted by

Emily Ping Fang Hsu

From *Taipei, Taiwan*

2023

Acknowledgement

I would first like to express my gratitude to Prof. Dr. Florian Fricke from the Department of the Microbiome and Applied Bioinformatics for his supervision and guidance throughout this PhD project, and for providing the space and facilities to carry out experiments. His timely suggestions aided tremendously in giving this thesis the structure and overarching theme it needs, and expanded my understanding of how experimental results can lend to fresh insights on innate immunity within the human gut environment.

I would also like to single out Dr. Daniel Podlesny, who performed bioinformatics analyses on microbial genomes and generated the K-mer database, therefore allowing me to quickly select relevant microbial genomes for experimental use. Additionally helpful to me is Monica Schumacher for her efficient management of lab supplies and work environment, thus streamlining many of the complex DNA processing procedures required for this project.

Additionally, I would like to thank the following current and former members of the Department of the Microbiome and Applied Bioinformatics: Elizabeth Dörner, Alena Bubeck, and Kathrin Stempfle. Also, the secretaries Kirsten Schubert-Wagner and Anne Jordan for administrative helps.

Prof. Dr. Thomas Kufer from the Department of Immunology helped me tremendously with this project, by lending his laboratory space, supplying bacterial and cell cultures, and giving me personal advice on immunological concepts and relevant experimental procedures. I would also give my thanks to Yvonne Postma for teaching me the varied immunological techniques used in Prof. Kufer's group.

I would finally like to thank Prof. Lutz Grevé from the Department of Biochemistry and Dr. Siegfried Preuß from the Department of Animal Genetics and Breeding for lending their laboratory space and supplies for relevant experiments, and Prof. Schmidt from the Department of Food Biotechnology for providing bacterial cultures.

Abstract

The human Toll-like Receptor 9 (TLR9) is an endosomal Pattern Recognition Receptor (PRR) that recognizes DNA sequences containing the unmethylated Cytosine-Guanine (CpG) dimers, which are present in greater abundance in most bacterial genomes compared to those of vertebrates. Specific CpG-containing sequences are strongly stimulatory of human TLR9, as shown in published studies using synthetic oligonucleotides (ODN) and DNA from bacterial species of varying genomic CpG concentration.

Human TLR9 activation was experimentally examined in this thesis using DNA extracted from different bacterial sources, human DNA from Caco-2 cells, known immunostimulatory ODN, and short ODN. *In vitro* assays using fragment length-standardized microbial genomic DNA on HEK-Dual TLR9 Cells and human peripheral blood mononuclear cells (PBMCs) revealed that TLR9 activation strongly correlated to CpG concentration of the input DNA, with an additional influence of CpG-containing 5-mer TCGTT concentration.

When DNA of varying origins and fragment lengths were used together, however, complex dynamics of TLR9 activation, co-activation, and repression were observed, which were less predictable than expected from genomic CpG concentration alone. DNase I-treated microbial DNA fragments of less than 15 bp of length were non-activating on their own, but co-activated human TLR9 together with ODN-2006 in Ramos Blue (B) cells. Similarly, human DNA fragments at the length of 50-200 bp co-activated human TLR9 with both ODN-2006 and *Escherichia coli* DNA in HEK-dual TLR9 cells. In contrast, large human DNA fragments at over 10000 bp of length repressed TLR9 activation by ODN-2006 in Ramos Blue cells.

Finally, a preliminary study was conducted in HT-29 cells on the effect of TLR9 activation on the invasion of *Fusobacterium nucleatum*, an opportunistic gut pathogen with a

2

very low genomic CpG concentration at 0.296%, using ODN-2006 and human DNA as TLR9 activators. While increased presence of intracellular *Fusobacterium nucleatum* upon treatment with both ODN-2006 and human DNA was noted, more studies are needed to confirm TLR9 activation as a cause of greater bacterial invasion.

The human colon is the location of the largest microbial population of the human body, which provides a rich source of non-human DNA in contact with human TLR9 present in intestinal epithelial cells, plasmacytoid dendritic cells (pDCs), and B lymphocytes. Additionally, the daily mass shedding and death of human intestinal epithelial cells provide large amounts of human DNA, which when combined with microbial DNA could result in co-activation and possible autoimmunity. The thesis thus provided an *in vitro* model of TLR9 activation by complex DNA of varying origins and fragment lengths likely to present in the human gut environment, and prepared a working basis for future studies of TLR9 activation by human fecal metagenomic DNA.

Zusammenfassung

Der menschliche Toll-like-Rezeptor 9 (TLR9) ist ein endosomaler Mustererkennungsrezeptor (PRR), der DNA-Sequenzen erkennt, die unmethylierte Cytosin-Guanin-Dimere (CpG) enthalten. Diese sind in den meisten bakteriellen Genomen in größerer Menge vorhanden als im Vergleich zu denen von Wirbeltieren. Spezifische CpG-haltige Sequenzen stimulieren den menschlichen TLR9 unterschiedlich stark, wie in veröffentlichten Studien mit synthetischen Oligonukleotiden (ODN) und DNA von Bakterienarten mit unterschiedlicher genomischer CpG-Konzentration zeigten.

In der vorliegenden Arbeit wurde die Aktivierung des menschlichen TLR9 unter Verwendung von DNA aus verschiedenen bakteriellen Quellen, menschliche DNA aus Caco-2-Zellen, bekannte immunstimulierende ODN und kurze ODN experimentell untersucht. In *in-vitro*-Assays, in denen die TLR9 Aktivierung durch, mit Fragmentlängen-standardisierte mikrobielle genomische DNA in HEK-Dual TLR9-Zellen und humanen peripheren mononukleären Blutzellen (PBMCs) untersucht würde, zeigten dass die TLR9-Aktivierung stark mit der CpG-Konzentration Input-DNA korreliert und die CpG-haltige 5-mer TCGTT-Konzentration einen zusätzlichen Einfluss hat.

Wurden jedoch DNAs unterschiedlicher Herkunft und Fragmentlängen zusammen verwendet, konnte eine komplexe Dynamik der TLR9-Aktivierung, Koaktivierung und Unterdrückung beobachtet werden, die weniger vorhersehbar war als von der genomischen CpG-Konzentration allein erwartet. Mikrobielle DNA-Fragmente mit einer Länge von weniger als 15 Basenpaaren (bp), welche mit DNase I behandelte wurden, besaßen alleine keine aktivierende Funktion. Im Zusammenspiel mit ODN-2006, konnte jedoch eine Aktivierung zusammen des menschlichen TLR9 in Ramos Blue (B)-Zellen beobachtet werden. In ähnlicher Weise ko-

aktivierten menschliche DNA-Fragmente mit einer Länge von 50-200 bp den menschlichen TLR9, sowohl unter Zugabe von ODN-2006, als auch *Escherichia coli* DNA in HEK-dual TLR9-Zellen. Im Gegensatz dazu unterdrückten große menschliche DNA-Fragmente mit einer Länge von über 10000 bp die TLR9-Aktivierung durch ODN-2006 in Ramos Blue-Zellen.

Schließlich wurde eine vorläufige Studie über die Auswirkungen der TLR9-Aktivierung in HT-29-Zellen auf die Invasion von *Fusobacterium nucleatum*, einem opportunistischen Darmpathogen mit einer sehr niedrigen genomischen CpG-Konzentration von 0,296%, durchgeführt. Als TLR9-Aktivatoren wurden dabei ODN-2006 und menschliche DNA verwendet. Während eine erhöhte Präsenz von intrazellulärem *Fusobacterium nucleatum* nach der Behandlung sowohl mit ODN-2006 als auch mit menschlicher DNA festgestellten wurde, sind weitere Studien erforderlich, um die TLR9-Aktivierung als Ursache für eine stärkere bakterielle Invasion zu bestätigen.

Der menschliche Dickdarm beherbergt die größte mikrobielle Population des menschlichen Körpers. Diese stellt eine reichhaltige Quelle nicht-menschlicher DNA dar, die mit menschlichem TLR9 in intestinalen Epithelzellen, plasmazytoiden dendritischen Zellen (pDCs) und B-Lymphozyten in Kontakt kommt. Darüber hinaus liefern die täglichen Ausscheidungen und das Absterben menschlicher Darmepithelzellen große Mengen menschlicher DNA, die in Kombination mit mikrobieller DNA zu einer Koaktivierung und möglicher Autoimmunität führen könnten. Die Arbeit lieferte somit ein *in-vitro*-Modell der TLR9-Aktivierung durch komplexe DNA unterschiedlicher Herkunft und Fragmentlänge, wie sie in der menschlichen Darmumgebung vorkommen kann, und bereitete eine Arbeitsgrundlage für künftige Studien zur TLR9-Aktivierung durch menschliche fäkale metagenomische DNA.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ALT	Alanine transaminase
AP-1	Activator protein 1
APS	Ammonium persulfate
BCG	Bacillus Calmette-Guérin, vaccine made from attenuated live bovine tuberculosis bacillus *Mycobacterium bovis*
bp	Base pairs
BSA	Bovine serum albumen
CFU	Colony forming unit
cGAMP	Cyclic guanosine monophosphate–adenosine monophosphate
cGAS	Cyclic GMP-AMP synthase
CpG	Cytosine-phosphate-Guanine dimer (5'-C-phosphate-G-3')
C-section	Caesarian section
CXCL1	Chemokine (C-X-C motif) ligand 1
DENV	Dengue virus
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
DNA, mtDNA, cfDNA	Deoxyribonucleic acid, mitochondrial DNA, cell-free DNA
DOTAP	1-(2,3-Dioleoyloxy)-N,N,N-trimethylammonium-propane methyl sulfate
DSMZ	German Collection of Microorganisms and Cell Culture GmBH
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
ER	Endoplasmic reticulum
FCS	Fetal calf serum
FDR	False discovery rate
FITC	Fluorescein isothiocyanate
HEK	Human embryonic kidney cells
HMGB-1	High mobility group box 1
IBD	Inflammatory bowel disease
IFN-α, IFN-β	Interferon
IKKα, IKKβ	Inhibitor of nuclear factor kappa-B kinase subunit α, β
IL-6, 8, 10, 18, 1β	Interleukin 6, 8, 10, 18, 1β
IMDM	Iscove's Modified Dulbecco's Medium
IRAK	Interleukin-1 receptor (IL-1R) associated kinase
IRF7	Interferon regulatory factor 7
LB	Luria Broth
LL-37	Active form of cathelicidin antimicrobial peptide
LPS	Bacterial lipopolysaccharides, protein on the surface of gram-negative bacteria and TLR4 agonist
LRR	Leucine rich repeat
MAMP	Microbe-associated molecular patterns
MCP-1	Monocyte chemoattractant protein 1
MOI	Multiplicity of infection

MRS	De Man, Rogosa, and Sharpe medium
MUC2	Mucin 2
MyD88	Myeloid differentiation primary response 88
NAFLD	Non-alcoholic fatty liver disease
Nalp3	NOD-like receptor family pyrin domain containing 3
NCBI	National Center for Biotechnology Information
NEC	Necrotizing enterocolitis
NET	Neutrophil extracellular trap, DNA-protein matrix released by neutrophils to trap bacteria
NF-kB	Nuclear factor kappa-light-chain-enhancer of activated B cells
OD	Optical density
ODN, sODN, iODN	Synthetic oligonucleotides, short synthetic oligonucleotides, inhibitory synthetic oligonucleotides
PAGE	Polyacrylamide gel electrophoresis
PAMP	Pathogen-associated molecular patterns
PBMC	Peripheral blood mononuclear cells
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
pDC, BM-DC	Plasmacytoid dendritic cells, bone marrow derived dendritic cells
PRR	Pattern Recognition Receptor
RNA	Ribonucleic acid
RPMI	Roswell Park Memorial Institute medium
s.d.	Standard deviation
s.e.m.	Standard error of the mean
SEAP	Secreted embryonic alkaline phosphatase
SLE	Systemic lupus erythematosus, an autoimmune disease
STAT6	Signal transducer and activator of transcription 6
STING	Stimulator of interferon genes
TAE	Tris base, acetic acid, and EDTA buffer solution
TBE	Tris base, boric acid, and EDTA buffer solution
TEMED	Tetramethyl ethylenediamine
T_h1, T_h2	Type 1 T helper cell, Type 2 T helper cell
TIR domain	Toll/Interleukin-1R domain
TLR	Toll-Like Receptor
TNF	Tumor necrosis factor
TRAF3, TRAF6	TNF (tumor necrosis factor) receptor associated factors 3 and 6
UV	Ultraviolet Light

1. Introduction

1.1 The Toll-Like Receptor 9 (TLR9) and the Innate Immune System

The Toll-like Receptor 9 (TLR9) belongs to the Toll-like receptor (TLR) family, a category of Pattern Recognition Receptors (PRRs) within the innate immune system[1,2]. The innate immune system has an older evolutionary lineage compared to the adaptive immune system[3,4], and is the only type of immunity found in invertebrates[5]. Evolutionary development of innate immunity had been studied in various organisms including *Drosophila*[6], jawless fish[7], and Placozoans[8], the latter of which the most primitive animal in which innate immunity was identified. For organisms possessing both the innate and the adaptive immune system, the innate immune system forms the first line of defense[9], and gives rise to a fast-acting immune response which confronts a diversity of pathogens regardless of origin[10]. The innate immune system can also signal the immune cells responsible for the adaptive immune response, such as Antigen Presenting Cells, via the Type I interferon (IFN) signaling pathway[11,12].

Contrary to the adaptive immune system, where immune cells generate pathogen-specific antibodies in a lock-and-key interaction with bacterial and viral antigens via memory of prior exposures[13], the innate immune system recognizes generalized features of invading pathogens such as bacterial lipopolysaccharides (LPS)[14] or viral nucleic acid sequences[15]. PRRs, variously located at the surface, cytosolic, or endosomal regions of a cell[16], respond to a variety of features deemed potentially harmful to the host[16,17]. PRRs recognize foreign bodies distinguishable from the self, as well as patterns associated with tissue damage[18]. Activation of PRRs leads to inflammation[17], a complex physiological process which includes production of cytokines[17] and Nitrous Oxide (NO)[19, 20], vasodilation, and infiltration of immune cells into damaged tissues[21].

14

The TLR family is a family of PRR that recognizes patterns expressed by intruding microorganisms including bacteria, viruses, and fungi, collectively known as pathogen-associated molecular pattern (PAMP) or microbe-associated molecular pattern (MAMP)[23-25]. The receptor family received its name from the protein product of Toll gene in *Drosophila melanogaster*, discovered by the Nobel laureate Christiane Nüsslein-Volhard, which not only controls early larval development of dorsoventral polarity in *Drosophila*[26], but also their innate immune defense against fungal infection[27]. The vertebrate equivalents of the *Drosophila* Toll, which similarly induce an NF-kB signaling cascade upon activation, is therefore named Toll-like receptors[28].

The TLR family are Type I transmembrane glycoproteins consisting of an extracellular domain, a single transmembrane domain, and an intracellular domain[29, 30]. The extracellular domain consists of modules containing leucine rich repeat (LRR) amino acid sequences sandwiched between N-and-C termini and form a characteristic horseshoe shape, which is responsible for recognizing specific microbial motifs[29]. In mammals, the TLR family recognizes MAMPs via direct binding of the ligand with the extracellular domain of the TLR. Upon ligand binding, TLR dimerizes and recruits adaptor proteins responsible for downstream signaling[30]. The structure of extracellular domains of endosomal nucleic acid sensors TLR7, TLR8 and TLR9 are similar to each other, all of which include a long Z-shaped insertion loop between LRR14 and LRR15[31]. The Z-loop was found to be crucial for DNA binding for TLR9[31].

The TLR family is present in both vertebrates and invertebrates, and represents one of the oldest, most evolutionarily conserved innate immunity receptor families[32]. While they recognize patterns expressed by bacteria and viruses, the patterns they recognize are not pathogen-specific, which leads to problems when they react against commensal microorganisms[33]. Thirteen members of the TLR family were identified in human and mice together, though some TLRs,

such as TLR10, 11, 12 and 13, are present in mice but not in humans. Ten TLRs were identified in humans[34].

1.2 Location and Structure of the TLR9

TLR9, similar to other TLR specialized in nucleic acid sensing, are not located on the cell surface. Instead, they are endosomal receptors, thus the nucleic acid ligand must be transported to the endosome before an interaction takes place[35]. To activate TLR9, a ligand whether in the form of a bacterium, virus, or synthetic oligonucleotide, must first be introduced into the cell[35].

TLR9 recognizes DNA patterns characteristic of invading bacteria and viruses, namely unmethylated Cytosine-Guanine (CpG) dimers, which are present at higher frequencies in most bacterial genomes compared to those of vertebrates[36]. The resting TLR9 resides on the Endoplasmic Reticulum (ER)[37]. However, they translocate to the endo-lysosomes where they bind to DNA[38]. Translocation of the TLR9 from the ER to the endo-lysosome is constitutive for a fraction of TLR9[39], where they traffic through the Golgi complex upon exiting the ER[38, 39]. More TLR9 were subsequently recruited upon an initial activation event of local TLR9, augmenting the TLR9 response[39]. The endosomal location of TLR9 prevents TLR9 from reacting to self-DNA[40]. Self-DNA are rarely endocytosed and remain in the extracellular domain, where they are rapidly degraded[41]. Engineered relocation of TLR9 onto the cell surface was found to induce powerful reactivity to self-DNA and autoimmunity[40].

The inactive TLR9 is a monomer and contains a Z-loop at its extracellular domain. The orientation of the Z-loop for TLR9 is oriented differently from the RNA sensor TLR8, with the TLR9 Z-loop extending toward the C-terminus and capable of interacting with LRR 15-21. The ring structure of TLR9 has more unoccupied inner space compared to TLR8 and therefore allows

ligand binding on the concave inner surface. Upon addition of TLR9-activating ligands, two receptors bind to each other, forming a dimer around the ligand[31].

With the addition of a strand of CpG-containing Oligonucleotides (ODN), local conformational changes of the loop regions LRR8, LRR11 and LRR18 occur, allowing for TLR9 dimerization. Both members of the dimer recognize the same sequence: The CpG bases are recognized via multiple amino acids and hydrogen bonds, and the ODN forms a bent shape wedged between the TLR9 dimers. During dimerization, LRRNT, LRR1 and LRR2 of one receptor recognizes CpG dimer bases of the stimulatory ODN strand with additional hydrogen binding capacities for flanking purine sequences, while LRR20 to LRR22 of the second receptor binds to the backbone, resulting in opposing orientation of the two receptors. When inhibitory ODN (iODNs) are present, the iODN sequence that binds to a TLR9 monomer partially overlaps with the ODN, competitively inhibiting ODN binding, resulting in an antagonistic effect to TLR9 activation[31].

1.3 Downstream Signaling upon TLR9 Activation

Upon dimerization, TLR9 signals Toll-interacting proteins (TIR domain) attached to MyD88, an adaptor protein. MyD88 then signals the protein kinases IRAK1, IRAK2 and IRAK4, which are attached to TRAF3 and TRAF6. Two different pathways are activated depending on cell type: the NF-kB/AP1 pathway through TAK1, or the Type-I interferon pathway through IRF7[42]. In the first pathway, exemplified by human B cells, TRAF6 signals of the protein complex TAB2, TAB3 and TAK1, the same complex activated by the surface TLR such as TLR1, 2, 5 and 6. Three downstream pathways of the protein complex were identified: IKKα and IKKβ signals NF-kB, MIKK4 or MIKK7 signals JNK which signals AP-1, and MIKK3 or MIKK6 signals p38 which signals CREB. NF-kB, AP-1 and CREB are transcription factors

occupying transcription sites of pro-inflammatory genes, resulting in production of inflammatory

cytokines such as IL-6, IL-8, and TNF-α[24]. The second pathway is characteristic of plasmacytoid

dendritic cells and also involves IKKα, which signals IRF7, a transcription factor for the

production of Type-I interferons such as IFN-α and IFN-β[24]. This pathway is primarily signaled

by endosomal nucleic acid sensors[43], and thus can be used to distinguish TLR9 activation by

bacterial DNA from surface TLR activation by residue bacterial components such as LPS in

TLR9 activation experiments using microbial DNA.

1.4 TLR9 Recognizes CpG-Containing DNA Sequences

The CpG dimer is one of the 16 possible dimer combinations within a genome. However,

due to rules of reverse complement, all dimers necessitate fixed sequences at its complementary

strand, i.e., AA must be complemented with TT, and so forth, thus by chance alone, CpG dimers

should occur ca. 1/8 (12.5%) of the time in a typical double-stranded DNA. The CpG

concentration of bacteria, whether human-associated or purely environmental, varies greatly

between species but typically, the CpG concentration is lower than 12.5%. A reason is that the

concentration of G+C is lower than 50% for most bacterial species[44]. The phenomenon may be

energetic in nature. Transcription of different bases requires different amounts of energy, with

Guanine as the costliest base followed by Cytosine[43]. Therefore, positive selection for the use of

A and T bases in transcribed regions of the genome occurred when multiple codon options are

available for a given amino acid occurs, depressing the G+C concentration[45].

However, a lower CpG concentration than expected even within a given % G+C is found

in many commensals and pathogens[46], as well as in typical vertebrate genomic DNA[47]. Among

bacteria of the same species such as *Clostridioides difficile*, more virulent and invasive strains

exhibit additional CpG depression compared to less virulent strains[48]. In addition, some DNA

viruses are also lower in CpG concentration than expected[49], suggesting evolutionary selection for evasion from TLR9 recognition.

Among microbial species associated with human diseases, a wide range of CpG concentrations are represented. *Mycobacterium tuberculosis*, the bacterium that causes the lung disease tuberculosis, has 12.7% CpG[50]. *Salmonella* Typhi has 8.48% CpG[50], while *Shigella flexneri*'s CpG concentration of 7.27%[46] is very similar to that of *Escherichia coli*, some strains of which are also invasive. Others, however, have low CpG concentration, such as the food-borne pathogen *Campylobacter jejuni* (1.44%)[50] and *Clostridium perfringens* (0.482%)[46]. Significantly, whole genera of bacteria adapted to strictly intracellular lifestyle are low in CpG. *Borrelia bergadorfi*, the causative agent of Lyme disease, has 1.01% CpG[46]. *Mycoplasma* spp., a parasitic intracellular genus causing a variety of diseases, have genomic CpG concentration as low as 0.209%[51]. The low CpG concentration of intracellular species suggested possible adaptation to escape the TLR9-dependent antimicrobial response cascade, allowing for extended survival in the intracellular environment[51].

For human gut commensals, CpG concentration is generally lower than 12.5%: *Escherichia coli* has 7.47% CpG, *Bacteroides fragilis* has 4.72% CpG, while *Fusobacterium nucleatum*, an oral commensal also found in the human gut, has a very low CpG concentration of 0.296%[46]. Bioinformatics estimate for the average CpG concentration of a typical adult microbiota is at circa 4.85%[50]. The microbiota of a nursing infant, due to the prominent presence of the high-CpG genus *Bifidobacterium*, has a higher CpG concentration than an adult's[52].

The human genome has a very low CpG concentration: at 0.9%, the CpG concentration is at one-fourths of expected from genomic % G+C of 38%[53]. The low CpG concentration of human DNA, like those of other vertebrate species, could represent an evolutionary adaptation against autoimmunity[47].

The human genomic DNA is additionally heavily methylated. Typically, methylation occurs at locations rich in CpG dimers, on the Cytosine base[54]. ODNs with methylated CpG bases are non-activating of TLR9[55], suggesting that the lack of TLR9 activation by human DNA is not due solely to low CpG concentration. Methylation is another means by which human TLR9 distinguish self-DNA from microbial intruders[55]. Nonetheless, CpG methylation is exploited by some bacteria. An enzyme involved with CpG-DNA methylation, the methyltransferase M.Mpel, was found in *Mycoplasma penetrans* which, along with its very low genomic CpG concentration, is likely to further help it evade human TLR9 detection and promote their survival inside host cells[51].

1.5 Sequence Patterns Influencing Human TLR9 Activation

Elucidation of molecular patterns activating human and mouse TLR9 were carried out using synthetic CpG-containing oligonucleotides, or ODNs[56]. The design of CpG-containing ODN originated from studies of genetic sequences of bacterial extracts such as those from the mycobacterium Bacillus Calmette-Guerin (BCG)[57] and antisense sequences to retroviruses[58]. Antisense ODN were found to induce B cell proliferation and immunoglobin secretion[59]. Via cloning of the bacterial and viral genes, it was found that the self-complimentary palindrome CpG dimers within these genes are responsible for innate immune response to bacterial DNA[60]. Hartmann and Krieg (2009) additionally noted that antisense ODN alone do not explain immune activation, but specific sequence motifs do, and the flanking sequences around the CpG dimer play a major role[61].

Three classes of CpG-ODN for humans were designed leading to somewhat different responses from TLR9-containing immune cells. Class A CpG-ODN is of mixed phosphorothioate- and phosphodiester backbone, and the two CpG-containing sequence motifs

have a phosphodiester backbone, and are palindromic and self-complementary. The sequence, named ODN-2216, has the sequence: 5'-ggGG<u>GACGA:TCGTC</u>gggggg-3' (20-mer), where capital letters indicate phosphodiester bases and underlined letters indicate the self-complementary loop. The sequence induces strong production of Interferon-alpha (IFN-α) from human plasmacytoid dendritic cells (pDCs) but weak B-cell response. Class B ODN, ODN-2006, is a linear ODN of completely phosphorothioate backbone with the sequence: 5'-tcgtcgttttgtcgttttgtcgtt-3' (24-mer). The sequence is highly stimulatory of human B-cells, inducing TLR9-dependent NF-kB signaling and proliferation, but is weak in inducing pDC IFN-α production. Finally, Class C ODN, ODN-2395, is a duplex DNA of phosphorothioate backbone with a 3' palindrome with the sequence: 5'-tcgtcgtttt<u>cggcgc:gcgccg</u>-3' (22-mer), where the underlined letters indicate palindromic sequences through which two sequences complement each other and form a duplex. ODN-2395 stimulate both human and murine TLR9, and induce both pDC and B-cell response. For all classes of ODN, a switch of the CpG dimer with GpC eliminates the TLR9 response.[62]

In a series of ingenious experiments using a self-designed ODN (called minH75, a modified ODN sequence based on ODN-2006 with a phosphodiester backbone) and short ODNs (sODN) of less than 10 bases long, Pohar et al. illustrated an additional dynamic of human TLR9 activation. sODN sequences containing the CpG dimer flanked by Tyrosine bases as 4-mers and 5-mers, TCGT and TCGTT, were non-activating on their own of Ramos Blue cells (an engineered B lymphocyte reporter cell line), but intensely promoted TLR9 activation together with low concentrations of minH75. The phenomenon of sODN co-activation of human TLR9 with ODN occurred when sequences of different lengths are present, though the CpG dimer remains central: when the CpG dimer on the sODN sequence TCGTT was switched with GpC, or when the sODN consist of Tyrosine bases alone, no co-activation occurred, similar to how ODN-

2006GC (ODN-2006 with all CpG dimers replaced by GpC) is non-activating of human TLR9. CpG dimer without flanking tyrosine bases as sODN also co-activated TLR9 with minH75 on Ramos Blue cells, but to a weaker extent compared to TCGT and TCGTT.[63]

In addition, they showed TCGTT to co-activate human TLR9 with ODN sequences containing a single CpG dimer, which are non-activating on their own. TCGTT also co-activated TLR9 with DNA extracted from bacteria, mouse fibroblast DNA and their unmethylated PCR products, methylated ODN, and calf thymus DNA, thus lower the sequence specificity needed for TLR9 activation by a given DNA fragment. Via FITC labeling, Pohar et al. confirmed colocalization of sODN with ODN within the endosome of mouse bone marrow dendritic cells (BM-DC), explaining the synergistic effect of ODN and sODN in contact with TLR9.[63]

Co-activation experiments implied that while human DNA, with its low CpG concentration and extensive methylation, is itself not a potent TLR9 activator, they can act in synergy with small CpG-containing DNA fragments of microbial origins to induce TLR9 response. The interaction suggests that when human and microbial DNA of different lengths are co-present in the endosome, cross-interaction can result in increased TLR9 activation that cannot be explained by direct activation alone.[63]

Aside from co-activation of ODN by sODN, a class of ODNs are known to inhibit TLR9 activation, called inhibitory ODN (iODN). Even at the ratio of 1:10, iODN suppressed TLR9 activation by ODN, indicating a potent inhibitory effect[64]. A potent TLR9-inhibitory sequence is found in telomeric hexamer repeats at the end of the human chromosome TTAGGG[65].

1.6 TLR9 Response to Self-DNA

The human genomic DNA is both low in CpG dimer and where CpG dimers are present, heavily methylated[54]. Therefore, human TLR9 is less inclined to be activated by its own DNA,

and can distinguish DNA sequences of microbial intruders versus its own. However, an exception exists regarding mitochondrial DNA[66]. Evolutionarily considered, the mitochondria originated from the ingestion of bacteria by ancient eukaryotic cells, thus retaining an independent genome with its own DNA sequence (mtDNA) with "bacteria-like" properties[67]. Compared to nuclear DNA, mtDNA contains high concentration of hypomethylated CpG sequences, thus is significantly more activating of TLR9[66]. Therefore, mtDNA can be recognized similarly to the DNA of bacteria, leading to a possible autoimmune response when they are damaged and recognized by the TLR9 receptor[66]. The role of mtDNA in TLR9 activation had been illustrated in a Dengue Fever Virus (DENV) model where the RNA virus, not a TLR9 ligand itself, caused mitochondrial permeability and release of mtDNA into the cytosol, which induced TLR9 response to mtDNA and release of inflammatory cytokines by pDCs[68,69].

DNA from apoptotic cells can also be a TLR9 ligand, as had been shown in mouse models on obesity[70] and acetaminophen-induced hepatoxicity[71]. In an obesity model, Nishimoto et al. noted that adipocytes degenerate due to chronic inflammation induced by high-fat diet. Apoptotic adipocytes released cell-free DNA (cfDNA) into the bloodstream, promoting monocyte chemo-attractant protein (MCP-1) expression in macrophages[70]. In the hepatoxicity model, when liver injury and liver cell apoptosis was induced by the drug acetaminophen, free DNA released from apoptotic liver cell was found to activate TLR9, which signaled transcription of pro-inflammatory genes IL-1β and IL-18 and the Nalp3 inflammasome[71]. Increased cfDNA in the bloodstream was also found in patients with obesity[72], insulin resistance[70], metastatic cancer[73], and Systemic Lupus Erythematosus (SLE)[74].

Self-DNA can be complexed to proteins e.g., high-mobility group box 1 (HMGB-1)[75] and antimicrobial peptides e.g., LL-37[76], forming a ribonucleoprotein complex. These complexes are formed in the event of DNA release from burst cells undergoing necroptosis[75], and during the

formation of the neutrophil extracellular trap (NET)[76], a matrix consisting of cellular DNA and proteins released by neutrophils as a means to capture invasive microorganisms such as type A *Streptococcus*[77]. Self-DNA-protein complex is a source of TLR9-associated autoimmune response in autoimmune diseases such as SLE. SLE is characterized by defective clearance of dead cells, which are insufficiently recognized and removed by professional phagocytes, thus becoming autoantigens consisting of self-DNA complexed to proteins that can bind to antibodies on the surface of pDCs and B cells[78]. The complexes are transported by surface antibodies into the endosome, which comes into contact with TLR9[79], whose activation leads to production of large amounts of IFNα and TNFα by pDCs and proliferation of B cells characteristic of autoimmunity[80].

1.7 TLR9 Activation and Innate Immunity in the Human Gut

The mucosal layer is a major defensive barrier of the gut segregating the intestinal microbiota from the cell layer, confining bacteria to the lumen of the large intestine. A gel-like matrix called the mucus, consisting of mucin, digestive enzymes, antimicrobial peptides and immunoglobins, lies between the cell layer and the lumen[33]. Mucin is a glycoprotein secreted by goblet cells, an elongated epithelial cell type present in the human intestine and the lungs[81]. The mucus is divided into two layers in the colon: a loose upper layer and a packed lower layer[82]. The upper layer of the intestinal mucus is in direct contact with the microbiota. Antimicrobial peptides, secreted by Paneth cells in the intestinal crypt, are concentrated at the lower layer of the mucus, and preclude bacteria from direct contact with the intestinal epithelium[82]. Mucus also protects against bacterial adhesion to the epithelium via transportation of bacteria down the digestive tract and shedding[83].

The microbiota is required for normal mucus development. Germ-free mice, which do not have a microbiota, showed anomaly in mucus composition. In a study using bacteria-sized beads flushing through the colon of germ-free and conventional mice, the beads were trapped within the upper mucosal layer of germ-free mice, which suggests a greater permeability of the mucus of germ-free mice to bacteria, while the mucus of conventional mice is fully impenetrable. Proteomics analyses showed greater presence of MUC2 in the mucus of conventional mice compared to germ-free mice.[84]

In addition to microbial DNA, the intestinal mucus contains human DNA, likely to originate from daily massive shedding and death of human intestinal epithelial cells. The human intestinal epithelial cells are some of the shortest-lived cells of the human tissues, with an extremely high turnover rate. At normal rates, all human intestinal epithelial cells are replaced in four days[85]. A typical intestinal epithelial cell in the colon is generated from the stem cells at the bottom of the intestinal crypt, pushed up alongside the crypt toward the tip of the villi, and finally shed from the tip[86]. The shed cells then die of apoptosis in the mucosal region[86].

In inflammatory bowel diseases and parasite infections, increased cell turnover occurs, suggesting greater apoptosis in the gut[87-89]. In addition, some gut cells die of necrosis, a cell death pathway that does not follow the standard apoptosis program[87]. In necrosis, the cells burst and spill their cytosolic materials into the surrounding environment, including DNA, which can be complexed with proteins into ribonucleoprotein complexes activating TLR9[90]. Increased cell death and premature cell death prior to shedding also compromise the gut barrier, resulting in greater contact of gut microbiota with immune cells[91]. Inflammatory bowel diseases such as Crohn's disease and ulcerative colitis are characterized by disruption of the gut barrier[92], where an increase of apoptotic and necrotic DNA in the gut, together with a compromised gut barrier

that permits more members of the microbiota to invade intestinal tissues[93], exposes resident

immune cells of the gut to an increased amount of DNA of both microbial and human origins,

aggravating the inflammatory response[94].

1.8 TLR9 Activation Can be Anti-Inflammatory

Activation of PRRs leads to the production of chemokines and cytokines[17]. While most

cytokines are pro-inflammatory, some, such as IL-10, are anti-inflammatory cytokines that

counter the inflammatory response[95]. Activation of different PRR's result in the release of

different cytokines types depending on the cell types affected and the receptors involved[17].

Anti-inflammatory roles played by TLR9 activation were studied on neonatal Necrotizing

Enterocolitis (NEC)[96, 97] and Non-Alcoholic Fatty Liver Disease (NAFLD) models[98], which

characterized the antagonistic relationship between TLR4 and TLR9 activation, where the former

is pro-inflammatory but the latter anti-inflammatory. In both models, the TLR4-activating agent

LPS aggravated inflammatory responses: In NEC, TLR4 activation induced gut cell apoptosis

and production of the inflammatory cytokine IL-6[99]; in NAFLD, it induced liver stenosis and

worsening of the liver disease indicator ALT[98]. The disease manifestations in both examples

could be reversed by the addition of ODN-1826, the mouse-specific TLR9 agonist. In a mouse

and piglet model on neonatal NEC induced by *Enterobacteriaceae*, Good et al. (2013) found that

TLR9 activation by *Lactobacillus rhamnosus* extracts could protect the intestine from NEC. The

same effect could be shown when ODN-1826 was used[99]. Therefore, TLR9 activation by gut

commensal DNA might play a role in immune homeostasis of the gut environment, and can be

anti-inflammatory in specific circumstances.

Additional evidence of TLR9 playing an inflammatory role was described in lupus-prone

MRL/Mplpr/lpr mouse models representing SLE. Compared to mice with wild-type TLR9, TLR9

deficient mice showed a greater number of activated B and pDC cells and increased production of inflammatory cytokines with concordant severer skin and kidney disease manifestations. TLR9 activation can thus serve a protective role in promoting clearance of dead cell debris and reducing inflammation in SLE patients.[100]

1.9 Goal of This Thesis

The gut microbiota is a rich supplier of DNA in the human gut environment and interacts with the human TLR9 in direct and indirect ways. As the human large intestine harbors a vast population of human microbiota[101] and short-lived human intestinal epithelial cells[85], sources for high concentration of DNA of both microbial and human origin, it is an ideal environment for human TLR9 activation to occur. The project aims to clarify how specific combinations of DNA of various origins—synthetic ODNs, microbial DNA and human DNA—, that might be present in the human gut in both normal and pathological conditions, affect human TLR9 activation. While it is known that microbial DNA activates human TLR9, and TLR9 activation depends on CpG-containing sequences, which are more concentrated in most microbial genomes compared to that of the human genome, human TLR9 activation by microbial and human DNA co-present in the gut is likely to play out in a more complex manner that affects human gut homeostasis.

In addition, we aim to develop experimental models using DNA of varying origins and fragment lengths to provide insights into TLR9 activation applicable to situations such as infant gut development, inflammatory bowel diseases, bacterial infection, and colorectal cancer. While the overall innate immune response in the gut in the aforementioned situations is likely to be far more complex, various normal and disease conditions are predicted to cause TLR9-containing cells of both epithelial and leukocytic origins to come into contact with microbial and human

DNA. The role played by human DNA, likely to originate from apoptotic and necrotic intestinal epithelial cells, in the context of human TLR9 response in the gut, remains unclear.

Finally, preliminary experiments on *Fusobacterium nucleatum* invasion of intestinal epithelial cells are included to identify the role of TLR9 activation on bacterial penetrance. As the species has a very low CpG concentration and is an opportunistic pathogen often found in the gut of patients with colorectal cancer[102], an invasion model could provide hints on TLR9 activation within the tumor microenvironment.

2. Methods

2.1 Bacteria Cultivation

2.1.1 Source of Bacteria Stock

Micrococcus luteus (DSM 1790), *Bifidobacterium bifidum* (DSM 20239), *Bifidobacterium breve* (DSM 20213), *Lactobacillus ruminis* (DSM 20403), *Bacteroides dorei* (DSM 17855), *Enterococcus faecalis* (DSM 20478) , *Staphylococcus epidermidis* (DSM 20044), *Lactobacillus salivarius* (DSM 20555), *Clostridioides difficile* (DSM 1296), *Clostridium perfringens* (DSM 756), and *Fusobacterium nucleatum* (DSM 15643) were purchased as freeze-fried stocks from the German Collection of Microorganisms and Cell Culture GmBH (DSMZ). *Escherichia* K12/C600 and *Campylobacter jejuni* (DSM 4688) were a gift from Department of Food Biotechnology (Prof. Herbert Schmidt), and *Escherichia coli* DH5α were a gift from the Department of Immunology (Prof. Thomas Kufer) from the University of Hohenheim, Stuttgart, Germany).

2.1.2 Overview of Medium and Conditions for Bacterial Cultures

Table 1. Cultivation Medium, Growth and Freezing Conditions of the Bacteria Species Used

Bacterial Species (strain)	Growth Medium	Growth Environment	Propagation Time	Freezing medium
Escherichia coli (K12/C600)	LB Medium and LB Agar (Carl Roth)	37°C, aerobic	16 h	70% LB medium 15% Glycerol (Sigma) 15% sterile H_2O
Escherichia coli (DH5α)	LB Medium and LB Agar Columbia Agar with 5% defibrinated sheep blood (Biomerieux)	37°C, aerobic	20 h	70% LB medium 15% Glycerol 15% sterile H_2O

Campylobacter jejuni (DSM 4688)	Columbia Agar with 5% defibrinated sheep blood	37°C, microaerophilic	72 h	70% Columbia Broth (BD) 15% Glycerol 15% sterile H_2O
Micrococcus luteus (DSM 1790)	Nutrient Agar: 5g/L peptone (Carl Roth), 3g/L meat extract (Carl Roth), 15g/L Agar (Carl Roth)	30°C, aerobic	96 h	70% nutrient broth (5g/L Peptone, 3g/L Meat Extract) 15% Glycerol 15% sterile H_2O
Bifidobacterium bifidum (DSM 20239)	Bifidobacterium medium: MRS medium (Carl Roth) with 40mL/L salt solution*, 5g/L NaCl and 0.05% Cysteine-HCl (Carl Roth), pH 6.8, Bifidobacterium agar (+15g/L Agar)	37°C, anaerobic	48 ~ 72 h	70% Bifidobacterium medium 15% Glycerol 15% Sterile H_2O
Bifidobacterium breve (DSM 20213)	Bifidobacterium medium: MRS medium with 40mL/L salt solution(Note 1), 5g/L NaCl and 0.05% cysteine-HCl, pH 6.8, Bifidobacterium agar (+15g/L Agar)	37°C, anaerobic	48 ~ 72 h	70% Bifidobacterium medium 15% Glycerol 15% Sterile H_2O
Lactobacillus ruminis (DSM 20403)	MRS medium with additional 2g/L meat extract, 1g/L yeast extract, and 0.05% Cysteine-HCl pH 6.2~6.5 MRS Agar (+15g/L Agar)	30°C, anaerobic	48 h	70% MRS medium 15% Glycerol 15% Sterile H_2O
Bacteroides dorei (DSM 17855)	Columbia Agar with 5% defibrinated sheep blood	37°C, anaerobic	48 h	20% Powdered Skim Milk (Carl Roth)
Enterococcus faecalis (DSM 20478)	Columbia Agar with 5% defibrinated sheep blood	37°C, microaerophilic	48 h	70% Columbia Broth (BD) 15% Glycerol 15% sterile H_2O
Staphylococcus epidermidis (DSM 20044)	Columbia Agar with 5% defibrinated sheep blood	37°C, aerobic	48 h	70% Columbia Broth 15% Glycerol 15% Sterile H_2O
Lactobacillus salivarius (DSM 20555)	MRS medium with additional 2g/L meat extract and 1g/L yeast extract, pH 6.2~6.5	37°C, anaerobic	48 h	70% MRS medium 15% Glycerol 15% Sterile H_2O
Clostridioides difficile (DSM 1296)	Brain-Heart Infusion Broth (Sigma) with 5g/L Yeast Extract and 0.05% Cysteine-HCl Columbia Agar with 5% defibrinated sheep blood	37°C, anaerobic	48 ~ 72 h	70% Brain-Heart Infusion Broth (Sigma) with 5g/L Yeast Extract and 0.05% Cysteine-HCl 15% Glycerol 15% Sterile H_2O
Clostridium perfringens (DSM 756)	Columbia Agar with 5% defibrinated sheep blood	37°C, anaerobic	48 h	70% Columbia Broth 15% Glycerol 15% sterile H_2O

Fusobacterium nucleatum (DSM 15643)	Columbia Agar with 5% defibrinated sheep blood	37°C, anaerobic	48 ~ 72 h	70% Columbia Broth 15% Glycerol 15% sterile H_2O
Liquid culture (*Fusobacterium nucleatum and Clostridium perfringens*)	Brain-Heart-Infusion Broth with 5 mg/mL Hemin (Sigma) and 1 mg/mL Vitamin K (Sigma)	37°C, anaerobic	Variable (based on growth curve)	N/A
Liquid culture (*Staphylococcus epidermidis*)	Brain-Heart-Infusion Broth	37°C, aerobic	Variable (based on growth curve)	N/A

Note 1: Per 1L of salt solution, the following formula was used: 0.25 g $CaCl_2$ x 2 H_2O, 0.50 g $MgSO_4$ x 7 H_2O, 1.00 g K_2HPO_4, 1.00 g KH_2PO_4, 10.0 g $NaHCO_3$, 2.0 g NaCl. The mixture was filled up to 1L with distilled water and autoclaved using the standard liquid program.[103]

2.1.3 Medium Preparation

With the exception of Columbia Agar with 5% defibrinated sheep blood, which was purchased directly as prepared agar plates from Biomerieux, culture media were prepared from powdered stock, supplemental ingredients, and sterile water according to growth formulas provided by DSMZ. All ingredients, except for Agar, Hemin, Vitamin K and Cysteine-HCl, were added to water directly and left on a magnetic stand with a stirrer until fully dissolved, and pH value was adjusted with HCl and NaOH and a pH meter to +/- 0.1. Agar (if needed) was added to the medium and briefly stirred, and the medium was autoclaved at 121°C for 20 min. Once the autoclaved medium reached "hand-warm" temperature, i.e., when the medium bottles could be held by hand without discomfort, sterile 0.2 µm filtered Hemin, Vitamin K and Cysteine-HCl were added to the desired final concentration. Medium containing agar were poured into sterile 10 cm petri dishes at circa 20 mL per dish, and the plates were left at room temperature with open lids until agar solidified. Liquid medium and agar plates were stored in 4°C until further need. For cultivation of anaerobic organisms, aliquots of liquid medium and agar plates were incubated inside the anaerobic chamber for 24 h prior to use.

2.1.4 Cultivation and Storage of Bacterial Stock from Freeze-Dried Pellets

Purchased freeze-dried bacterial stocks were encased in closed glass tubes within sealed glass vials supplied by DSMZ. To break the glass vial, the sealed end was placed on a gas burner for circa 10 s, and water was pipetted onto the heated glass to crack the vial, with the aid of a pair of forceps. The inner glass tubes were taken out with forceps and moved to a suitable location (sterile bench or anaerobic chamber) depending on growth condition. The mesh seal was removed with forceps and set aside, and 500 µL of suitable medium was added directly to the freeze-dried pellet and left for 30 min with the mesh seal placed loosely on top. 350 µL of the re-suspended bacteria were inoculated in 5 mL of suitable liquid medium, and 30 µL of bacteria were streaked onto an agar plate (five in total) in four dilutions using an inoculation loop. The liquid culture and agar plates were incubated in suitable temperature and atmospheric conditions until visible growth was observed.

Bacterial stocks were prepared from both solid and liquid cultures. For single-colony stocks, one colony was picked from an agar plate and resuspended in cryotubes containing 300 µL of suitable storage medium (**Table 1**). Lawn stocks were generated by scratching a line of bacterial colonies using an inoculation loop and dipping the loop into 1 mL of suitable storage medium. Liquid stocks were generated by pipetting 1050 mL of liquid culture onto 450 µL of 50% glycerol. All bacterial stocks were briefly vortexed, then stored in -80°C until further need.

2.1.5 Recovery and Cultivation of Bacteria on Agar Plates

To recover bacteria, frozen stocks were streaked on suitable agar plates using autoclaved cotton swabs in several dilutions to generate single colonies. Once colonies were visible, single colonies were spread onto fresh agar plates with an inoculation spatula. Anaerobic bacteria were

left to grow either inside the anaerobic chamber, or in anaerobic jars (Oxoid 2.5L Anaerojar) and boxes (Biomerieux 2.5L Anaerobic Box) with suitable gas pack generators (Oxoid and Biomerieux). Microaerophilic bacteria were grown in anaerobic jars and boxes with microaerophilic gas packs (Oxoid and Biomerieux). Detailed protocol link from Oxoid is listed in **Appendix 7.2**.

2.1.6 Cultivation of Bacteria in Liquid Medium and Growth Curve Determination

A single colony of recovered bacteria on agar plates were inoculated into 5mL of pre-prepared liquid medium until visible growth was observed. A dilution of 1:100 (1:500 for *Clostridium perfringens*) was made, and the bacteria left to grow in an incubator or anaerobic chamber. One mL of bacteria was taken out from growth tubes at various times for optical density (OD) measurement at 600nm of wavelength, and 10 µL of bacteria were added to 90 µL of Phosphate Saline Buffer (PBS) (Gibco) in serial dilutions. A total of 50 µL of 10^{-3} to 10^{-6} dilutions were spread onto Columbia Agar plates with 5% defibrinated sheep blood using an inoculation spatula to generate single colonies. Colonies were counted in plates containing target colony number of 30 to 300 per plate, and three graphs were generated: OD versus time, colony forming unit (CFU) versus time, and CFU versus OD, to determine the time required for optimal logarithmic phase growth for the given species.

2.1.7 Collection of Bacterial Pellets from Liquid Medium and Agar Plates

Bacterial pellets were collected from both liquid cultures and agar plates. Liquid culture was centrifuged at 5000 rpm for 15 min twice with 1 PBS wash in between, and carefully removed via a 1000 µL pipette leaving the bacterial pellet intact. Bacterial lawns from agar

plates were swept with inoculation spatulas and inoculation loops and resuspended in 3mL pre-prepared PBS in 5 mL Eppendorf tubes. Two agar plates of bacterial lawns generated from single colonies were collected into one Eppendorf tube, then centrifuged at 5000 rpm for 15 min with one PBS wash in between. PBS was removed, and dry pellets were stored at -20°C for subsequent DNA extraction.

2.2 Human Cell Culture Growth and Processing

2.2.1 Source of Human Cell Culture Stock

HEK-Dual TLR9, HEK Null cells, and Ramos Blue (B) cells were purchased as frozen stock from InvivoGen Europe. HEK-Dual TLR9 and HEK Null cells originated from human embryonic kidney 293 (HEK-293) cells taken from a female fetus and transfected with human adenovirus 5 DNA for immortalization[104], and engineered with two NF-kB/AP-1 downstream reporters: the secreted embryonic alkaline phosphatase (SEAP) and Luminescence under the control of the IL-8 promoter[105,106]. The HEK-Dual TLR9 cell line was additionally engineered with human TLR9 that signals NF-kB upon activation[105]. Ramos Blue cells are human B lymphocytes extracted from a 3-year-old patient suffering from Burkett's Lymphoma[107] and engineered with SEAP reporter downstream of NF-kB, and possess natural TLR9[108]. Caco-2 cells were a gift from the Department of Nutritional Biochemistry (Prof. Lutz Graeve), and HT-29 cells were a gift from the Department of Immunology (Prof. Thomas Kufer), both from University of Hohenheim, Stuttgart, Germany. Both cell lines were originally isolated from human colon carcinoma[109] and are capable of forming a monolayer upon 100% confluence,

which differentiate into specialized cells possessing typical intestinal epithelial cell morphology and functions[110].

2.2.2 Routine Cell Culture Procedures

HEK-Dual TLR9 cells were cultivated in Dulbecco's Modified Eagle's Medium (DMEM) (Gibco) with 10% heat-inactivated Fetal Calf Serum (hi-FCS, pre-inactivated in 56°C liquid bath for 30 min) (Thermo Fisher Scientific) and 1% Penicillin/Streptomycin (Biochrom Ltd.), supplemented with 100 µg/mL Normocin, 50 µg/mL Zeocin, and 100 µg/mL Hygromycin B Gold (all from InvivoGen). HEK-Null Cells were cultivated in DMEM with 10% hi-FCS and 1% Penicillin/Streptomycin, supplemented with 100 µg/mL Normocin. Ramos Blue cells were grown in Iscove's Modified Dulbecco's Medium (IMDM) (Gibco) with 10% heat-inactivated FCS and 1% Penicillin/Streptomycin, supplemented with 100 µg/mL Normocin and 100 µg/mL Zeocin. Caco-2 and HT-29 cells were grown in DMEM with 10% hi-FCS and 1% Penicillin/Streptomycin. All cells were grown in T75 flasks in 37°C and 5% CO_2.

Cells were routinely passaged every 2-3 days. To passage adherent cells, cells were washed once with PBS and gently detached with 2 mL of 0.025 mg/mL Trypsin/EDTA (Biochrom Ltd. and Thermo Fisher Scientific), then stopped with the addition of growth medium and diluted to the desired cell density. Suspension cells were passaged via direct dilution with medium to the desired density. For HEK-Dual TLR9 and HEK-null cells, passage number of 15 or fewer were used for all experiments. Passage number of 11 or fewer was used for Ramos Blue cells. No passage number limit was observed for Caco-2 and HT-29 cells.

All cell stocks were frozen in growth medium supplemented with additional 10% hi-FCS (20% final concentration) and 10% Dimethyl sulfoxide (DMSO) (Carl Roth) in color-coded 2mL

cryotubes in liquid nitrogen until further need. Passage number was marked for all cell types except HT-29 cells.

2.2.3 Isolation of Human Peripheral Blood Mononuclear Cells (PBMCs) from Fresh Human Blood

Blood samples were collected from healthy adults in their 20s and 30s in EDTA-coated tubes. For the extraction of PBMCs, 16 mL of fresh human blood was mixed with 16 mL of pre-prepared growth medium: 80% RPMI-1640 medium (Gibco) with 20% hi-FCS and 1% Non-Essential Amino Acids (Life Technologies). In a fresh 50mL Falcon tube, 16 mL of Ficoll-Histopaque-1077 (Merck) was prepared, and the blood/medium mixture was slowly pipetted on top of Ficoll-Histopaque without disrupting the two layers. The tube was centrifuged at 1400 rpm for 30 min without breaking (actual centrifuge time ~45 min), and the middle opaque layer containing PBMCs were carefully extracted from the tube with a 1000µL pipette and transferred to a fresh 50mL Falcon tube. 30mL of fresh medium were added to the cells, and the cells were counted in a renumeration chamber. The tube was centrifuged at 1400 rpm for an additional 10 min to pellet the cells. Medium was removed, and the cells were re-constituted to the final cell density of 1×10^7 cells/mL with fresh medium inside a T25 flask and incubated in 37°C and 5% CO_2 overnight.

2.2.4 Partial Differentiation and Apoptosis Induction of Caco-2 Cells

For the generation of Caco-2 cell pellets for DNA extraction, Caco-2 cells were left to grow into a confluent monolayer, and incubated for an additional 6 days with routine medium change to allow for partial differentiation prior to apoptosis induction and cell collection. To

induce apoptosis, Staurosporine (Merck) was added to 8 µg/mL of final concentration to partially-differentiated Caco-2 cells for 24 h.

2.2.5 Collection of Caco-2 Cell Pellets for DNA Extraction

Apoptotic Caco-2 cells were harvested by scraping and centrifuged at 1500 rpm for 10 min to pellet. Supernatant containing apoptotic blebs was transferred to fresh 5mL Eppendorf tubes. The pellet, which contained apoptotic bodies, was washed with PBS and centrifuged for at 1500 rpm for 10 min twice. Supernatant containing apoptotic blebs was centrifuged at 18,000 x g (17,200 rpm) for 50 min to generate bleb pellets. Blebs were washed with PBS and centrifuged at 18000 x g for another 50 min. Intact Caco-2 Cells were treated with 2mL Trypsin/EDTA for 10 min, scraped, reconstituted with growth medium to 10mL, and transferred to fresh 50mL Falcon tubes, then centrifuged at 1200 rpm for 5 min. The pellets were washed with PBS and centrifuged at 1200 rpm for 5 min twice. PBS was removed after washing, and the pellets were frozen at -80°C.

2.3 DNA Extraction

2.3.1 Extraction and Endotoxin Removal of Bacterial DNA from Frozen Pellet

DNA was extracted from bacterial pellets via *Quick*-DNA Fungal and Bacterial Miniprep Kit. Metagenomic DNA from adult and infant stool samples was extracted via Quick-DNA Fecal/Soil Microbe Miniprep Kit (both from Zymo Research). DNA extraction was performed according to standard protocols supplied by the manufacturer. Extracted DNA from the Gram-negative species *Escherichia coli*, *Bacteroides dorei*, *Campylobacter jejuni*, and *Fusobacterium*

nucleatum were purified of contaminating endotoxin using the MiraCLEAN Endotoxin Removal Kit (Mirus Bio LLC). Bacterial DNA was eluted and dissolved in nuclease-free water, and DNA concentration and purity were determined via Nanodrop. Detailed protocol links from manufacturers are listed in **Appendix 7.2.**

2.3.2 Extraction of Human DNA from Intact and Apoptotic Caco-2 Cells

DNA was extracted from frozen Caco-2 cell pellets via DNeasy Blood and Tissue Kit (QIAGEN) with the following protocol modifications: from intact Caco-2 cells, pellets were resuspended with 2 mL of PBS and split into ten reactions. 4 µL of RNase A was added per reaction, then vortexed and rested for 2 min at room temperature. 20 µL of Proteinase K was added, followed by 200 µL of Buffer AL (Lysis buffer) provided in the kit. The reaction was vortexed rigorously for 20 s and incubated in 56°C dry bath for 20 min. 200 µL of 100% Ethanol was then added, and the mixture was vortexed for another 20 s. An additional 600 µL of Ethanol was added to the mix, the tubes briefly vortexed, and the DNA was left to precipitate at -80°C for at least 2 h.

For DNA extraction from apoptotic bodies and apoptotic blebs, Gel Electrophoresis results showed that the smaller fragments found in apoptotic DNA ladders were enriched in the flow-through of the spin columns (**Supplementary Figure S1**). Pellets were suspended in PBS, and the same enzyme digestion and incubation steps were performed as with pellets from intact cells, and combined with 200 µL of 100% Ethanol. The mixture was added to the provided spin columns attached to collection tubes, and centrifuged at 8000 rpm for 1 min. Spin columns are then discarded, and 600 µL of Ethanol was added to the collection tubes containing the

flowthrough, and the mixture transferred to into fresh 1.5mL Eppendorf tubes, vortexed briefly, and left to precipitate at -80°C for at least 2 h.

2.4 DNA Fragmentation

2.4.1 Double-Stranded DNA Fragmentase Treatment of Bacterial and Caco-2 DNA

Time course experiments were conducted for microbial and Caco-2 DNA enzymatically fragmented via dsDNA Fragmentase (NEB; New England Biolabs) to determine the optimal time needed to obtain fragment lengths of 50-200 bp. 3 µg of extracted DNA was diluted with water to 16 µL inside PCR strips (15 µL for human DNA with 1 µL of 200 mM $MgCl_2$). 2 µL of NEB dsDNA Fragmentase Buffer and 2 µL of NEB dsDNA Fragmentase were added to the diluted DNA. The reaction was vortexed, briefly centrifuged, and placed in benchtop dry bath at 37°C for 30, 60, 90, and 120 min, after which 5 µL of 0.5M EDTA (pH 8.0) (Life Technologies) was added to halt the reaction. The fragmentation reaction was loaded onto 1.5% agarose gel and run at 100V with a 100 bp ladder for 60 min. **Figure 1** from **Results 3.1.2** showed the resultant gel image and fragmentation time selection for microbial DNA from different species.

Time course experiments were performed on four bacterial genomes: *Bifidobacterium bifidum*, *Escherichia coli*, *Lactobacillus salivarius* and *Campylobacter jejuni* from 2 to 25 min to obtain target fragment length of 1000 bp. Bacterial DNA was loaded onto multiple wells, and a section of the gel representing DNA fragments of roughly 1000 bp was cut out into strips, and DNA was extracted from gel strips using GeneJET Gel Extraction Kit (Thermo Fisher Scientific) according to the manufacturer's instruction (**Appendix 7.2**).

Caco-2 DNA was also fragmented to 200-6000 bp of length to obtain fragment length distributions similar to that of DNA extracted from apoptotic cells. A time course experiment was performed on extracted Caco-2 DNA for fragmentation times of 2, 5, 7, 10, and 15 min, and the subsequent range of fragment lengths was determined was determined via Agilent Bioanalyzer (**Methods 2.6.3**). 7 min of fragmentation was used to generate 200-6000 bp fragments, analogous to the typical fragment lengths of DNA extracted from apoptotic Caco-2 cells loaded as control (**Supplementary Figure S2**). The precipitated DNA was then processed and purified according to steps outlined in **Methods 2.5.1** and **2.5.3**.

2.4.2 Covaris g-Tube Shearing of Bacterial DNA

Genomic DNA from *Escherichia coli* and *Campylobacter jejuni* was fragmented to 20,000 bp and 6000 bp via Covaris g-Tube (Covaris Inc.). To obtain fragments of 20,000 bp, 30 µg of bacterial DNA was loaded to the upper chamber of Covaris g-tubes and centrifuged at 5100 rpm for 1 min. For 6000 bp of length, bacterial DNA in Covaris g-tubes were centrifuged at 13,200 rpm for 30 s. The resultant DNA Fragment length were verified on a 0.7% agarose gel with Lambda Hind III Ladder (**Supplementary Figure S3**).

2.4.3 DNA Shearing via Covaris S220 Sonication

DNA from *Micrococcus luteus*, *Bifidobacterium bifidum*, *Escherichia coli*, *Lactobacillus ruminis*, and *Lactobacillus salivarius* were fragmented to 3000 bp of length using the Covaris s220 Focused Ultrasonicator (Covaris Inc.). 15 µg of bacterial DNA was diluted to 200 µL and placed in a blue miniTUBE (Covaris Inc.). Water was added to the sonication chamber to the indicated fill line and degassed for 30 min. Ultrasonicator setting was adjusted as followed—

water level: 15, Temperature: 20°C, Target bp: 3000 bp, peak incidence power: 3 Watts, duty factor: 20%, cycles per burst: 1000, treatment time: 600 s. The miniTUBE containing bacterial DNA were loaded to the miniTUBE holder (Covaris Inc.) and placed in the tank, and left to sonicate for 10 min. The resultant DNA fragments, confirmed to be an average of 3000 bp of length via gel electrophoresis (**Supplementary Figure S4**), underwent additional SPRI bead selection (**Methods 2.5.2**) prior to use in stimulation experiments.

2.4.4 DNase I Treatment of Bacterial and Human DNA

Small DNA fragments of < 15 bp length were obtained from bacterial and human genomic DNA digested with DNase I via the Ambion TURBO DNA-*free* Kit (Life Technologies). Per reaction, 7.5 µg of DNA was diluted to 44 µL in a 1.5-mL Eppendorf tube. 5 µL of TURBO DNase I Buffer and 1 µL of TURBO DNase was added, and the reaction was placed in a benchtop 37°C dry bath for 30 min. DNase Removal Reagent was vortexed to homogenize the beads, and 5 µL of the reagent was added to stop the reaction. The tubes were placed at room temperature for 5 min with intermittent vortexing, and centrifuged at 14,000 rpm for 1 min. The supernatant containing digested DNA fragments was transferred to fresh 1.5mL Eppendorf tubes.

2.5 Downstream Processing of DNA Fragments

2.5.1 DNA Clean-Up via Zymo Clean & Concentrator-25 and Oligo Clean & Concentrator

DNA was cleaned up via the Clean-and-Concentrator 25 or the Oligo Clean-and-Concentrator kit (both from Zymo Research) according to the manufacturer's instruction

(**Appendix 7.2**). Clean-and-Concentrator-25 kit was used to purify DNA extracted from intact and apoptotic Caco-2 cells, as well as DNA fragments between 200 bp to over 10,000 bp long. Oligo Clean-and-concentrator kit was used to purify Caco-2 apoptotic bleb DNA, as well as DNA fragments at 50-200 bp of length. Purified DNA fragments were eluted with water and precipitated with ethanol (**Methods 2.5.3**) prior to use in stimulation experiments.

2.5.2 SPRI Bead Selection: Left-Side and Modified Right-Side Selection

Left-side DNA selection was performed on DNA fragmented to 3000 bp length. Magnetic bead solution (Beckman Coulter and Zymo Research) was vortexed, and DNA was combined with beads at a ratio of 1:0.4. The DNA/Beads mixture was placed on a benchtop rotator for 30 min, briefly centrifuged, and placed on a magnetic stand until the solution is clear, circa 1 min. Supernatant was removed, and the beads were washed twice with 85% ethanol on the magnetic stand. Ethanol was removed and the beads were allowed to dry briefly. Nuclease-free water was then added to the beads, the mixture vortexed at top speed for 20 s, and placed on 37°C for 30 min to elute the DNA. The mixture was placed on a magnetic stand again, and the supernatant containing eluted DNA was transferred to a fresh 1.5 mL Eppendorf tube.

For small DNA fragments of 50-200 bp long, a simplified method of reverse selection was used. Here DNA was combined with magnetic beads at the ratio of 1:1.2, and placed on a benchtop rotator for 30 min. The mixture was centrifuged briefly and placed on a magnetic stand, and the supernatant faction was transferred to a fresh 1.5-mL Eppendorf tube. The supernatant containing the DNA fragments were cleaned up afterwards using the Oligo Clean-and-Concentrator kit and ethanol precipitation (**Methods 2.5.1 and 2.5.3**).

2.5.3 Ethanol Precipitation

Extracted and fragmented DNA was cleaned up by ethanol precipitation. Sodium Acetate (Abcam) was added to the DNA to a final concentration of 0.3M, followed by 2-2.5 times of 100% ethanol by volume. The DNA-ethanol mixture was briefly vortexed and placed in -80°C for 2 h to precipitate the DNA. Afterwards, DNA was centrifuged at 17,200 rpm for 30 min. Supernatant was removed, and the DNA pellet was washed twice with 70% ethanol with 15 min of centrifugation after each wash. Ethanol was removed and the pellet was left to dry for 5 to 10 min, then dissolved in nuclease-free water.

2.6 Determination of Fragment Length

2.6.1 Agarose Gel Electrophoresis

Gel electrophoresis was performed on extracted and fragmented DNA according to the standard procedure. Large DNA fragments of 6000 to 20,000 bp were loaded onto 0.7% agarose gels, middle-length fragments of 500-6000 bp were loaded onto 1% agarose gels, and small fragments of 50-500 bp were loaded onto 1.5% agarose gels. Agarose Broad Range powder (Carl Roth) was boiled in 1x TAE (diluted from 50x stock from Carl Roth) to dissolve, and stained with either ROTI-stain (Carl Roth) or GelStar Nucleic Acid Gel Stain (Lonza Bioscience) to the final concentration of 0.005%, and poured into a standard gel chamber to polymerize. 200 ng of DNA was diluted to 10 µL, combined with 2 µL of 6x Loading Dye (Carl Roth, Thermo Fisher Scientific). 5 µL of reference DNA ladders (Lambda Hind III, 1 kb, 200 bp or 100 bp ladders) were loaded depending on the expected DNA fragment length. For gels stained with GelStar Nucleic Acid Gel Stain, DNA ladders were first diluted (1 µL ladder, 5 µL Water, 1 µL 6x

Loading Dye). The gel chamber was connected to a power source set to 100 Volts for 45-60 min, and photographed via Life Technologies E-Gel Imager and the E-Gel Imager software (both from Thermo Fisher Scientific). All photos were inverted and exposure times adjusted to highlight DNA fragments while reducing background noise.

2.6.2 PAGE Gel Electrophoresis

Fragment lengths of DNase I-treated bacterial DNA were determined via PAGE gels containing 20% Acrylamide: Bisacrylamide. For a typical run with 2 gels, the following mixture was prepared: 14 mL 30% Acrylamide: Bisacrylamide, 4.55 mL H_2O, 2.1 mL 10x TBE (1M Tris Base, 1M Boric Acid, 0.02M EDTA in distilled H_2O), 350 µL 10% Ammonium Persulphate (APS), 17.5 µL Tetraacetylethylenediamine (TEMED). The mixture was poured into 1.5 mm gel cassettes on a gel caster. Combs were inserted immediately, and the gels were left to polymerize for 20 min. The cassettes were then transferred to a PAGE gel chamber. 1x TBE (diluted from 10x TBE with distilled H_2O) was filled between two cassettes or one cassette plus one buffer dam to the top of the wells, and outside the cassettes to the indicated fill line. Combs were removed, and gel wells were washed with a syringe. 20 µL of DNA sample was combined with 4 µL of 6x Loading Dye (Thermo Fisher Scientific), and loaded in the middle wells. 5 µL of Ultra-Low-Range Ladder (Thermo Fisher Scientific) was loaded at both sides of the sample DNA. The chamber was then connected to the lid containing electrodes, and connected to a power source set to 48 Volts. The gel was left to run to 2/3 of length, around 3 h 30 min.

After the run was complete, the chamber was disconnected. The cassettes were opened and the gels removed from the cassette. Loading wells were removed with a blade, and the gels were transferred to covered dishes containing 80 mL 1x TBE and 4 µL of GelStar Nucleic Acid

Gel Stain. The dishes were wrapped with aluminum foil, and placed on a benchtop shaker set to low speed for 30 min. The gels were then briefly de-stained with water, around 5 min, and placed under the Bio-Rad Chemi-Doc XRS+ Gel Imaging System UV Camera to confirm fragment length. Photos were inverted and exposure times adjusted to highlight DNA fragments and reduce background noise.

For DNA extracted from apoptotic blebs of Caco-2 cells, and bacterial and human DNA undergoing NEB dsDNA Fragmentase treatment, identical experimental procedures were carried out using gels containing 15% Acrylamide.

2.6.3 Agilent Bioanalyzer

A time course of NEB dsDNA Fragmentase treatment was performed on Caco-2 genomic DNA. 1 µL (diluted to concentration of 100 ng/µL when applicable) of fragmentation products after 2, 5, 6, 7, 8, 9, 10 min of NEB dsDNA Fragmentase treatment, plus DNA extracted from Caco-2 apoptotic bodies, were combined with 10 µL of Genomic DNA Sample Buffer (Agilent technologies, part no. 5067-5366). 1 µL of Genomic DNA ladder (part no. 5067-5366) was combined with 10 µL of Genomic DNA Sample Buffer. The mixture was transferred to Optical Tube 8x Strips (part no. 401428) labeled A1-H1 and A2, centrifuged briefly, vortexed at 2000 rpm for 1 min, then centrifuged again. The strips were uncovered and loaded onto a Genomic DNA ScreenTape, which was then inserted into the Agilent 2200 TapeStation System and read via the 2200 TapeStation Controller Software (see **Appendix 7.2** for detailed protocol provided by the manufacturer).

2.7 TLR9 Activation Experiments

2.7.1 Synthetic Oligonucleotides (ODN) and Small Synthetic Oligonucleotides (sODN)

Human TLR9-specific CpG-containing synthetic oligonucleotides ODN-2006 (5'-tcgtcgttttgtcgttttgtcgtt-3') and ODN-2216 (5'-ggGGGACGA:TCGTCgggggg-3'), their negative control GpC counterparts ODN-2006GC (5'-tgctgcttttgtgcttttgtgctt-3') and ODN-2216GC (5'-ggGGGAGCA:TGCTCgggggg-3'), and the inhibitory ODN INH-18 (5'-cctgaatgggaacttaccgctgca-3') were purchased directly from InvivoGen and Eurofins Genomics. (Bases with phosphodiester backbone are denoted by capital letters and bases with phosphonothioate backbone by small letters. Palindromic sequences are indicated with an underline with colon indicating the location of the stem loop.) Human TLR9-specific short synthetic oligonucleotides (sODN), 5'-TCGTT-3', and its negative control 5'-TTTTT-3', with phosphodiester backbone, were purchased from Eurofins Genomics.

2.7.2 Stimulation of Human Cells with ODNs, Bacterial DNA and Caco-2 DNA

HEK-Dual TLR9 cells, Ramos Blue cells, or human PBMCs were plated at the density of $1x10^5$, $5x10^4$ or $2x10^5$, and $4x10^5$ cells/well, 90 μL cells/well in 96-well plates in antibiotics-free medium supplemented with 10% FCS. To intensify the positive signal for specific TLR9 co-activation and repression experiments (**Results 3.7.1-3.7.3**), Ramos Blue cells were pre-starved for 16 h in IMDM with 0.2% Bovine Serum Albumen (BSA) and plated in the starvation medium prior to stimulation. ODN-2006, ODN-2006GC, sODN, bacterial DNA or human DNA of varying concentrations and fragment lengths were diluted with serum-free medium, or complexed with a fixed amount of the transfection agent DOTAP (Carl Roth) pre-diluted with

serum-free medium for 15 min, and 10 µL of the DNA mixture was added to cells in duplicates

to make for 100µL of final plating volume per well. For all experiments that included DOTAP,

DOTAP concentrations were standardized by plate to one nontoxic final concentration.

Stimulated cells were incubated overnight in 37°C, 5% CO_2.

2.7.3 Measurement of TLR9 Activation in HEK-Dual TLR9 and Ramos Blue Cells

TLR9 activation from HEK-Dual TLR9 and Ramos Blue cells was measured as the

release of engineered secreted embryonic alkaline phosphatase (SEAP) in the supernatant, 18-20

h after stimulation. 20 µL of supernatant was collected for the measurement. Ramos Blue cells, a

suspension cell line, was first centrifuged at 300 rpm for 15 min prior to supernatant collection to

avoid cell take-up during supernatant collection. Two additional wells consisting of growth

medium (or starvation medium) without antibiotics were added as blanks, depending on

stimulation condition. Quanti-Blue Solution (InvivoGen) were reconstituted according to the

manufacturer's instruction (**Appendix 7.2**), and 180 µL of Quanti-Blue was added to the

supernatant per well and the mixture incubated in 37°C, 5% CO_2 for 3 h. The plates were read at

the wavelength of 605-650 nm and TLR9 activation normalized to internal positive controls of

cells treated with fixed concentrations of ODN-2006.

2.7.4 Enzyme-Linked Immunosorbent Assay (ELISA)

Supernatant from stimulated PBMCs and HT-29 cells were collected for ELISA. IFN-β,

which was produced in response to ODN-2006 but not to bacterial LPS (**Supplementary Figure

S5**), was used to determine TLR9 activation from PBMC stimulation experiments. Human IL-6

ELISA was used to determine the effect of bacterial LPS. Human IL-8 ELISA was performed on

supernatant collected from HT-29 cells stimulated with ODNs and Caco-2 DNA, and from bacterial invasion experiments. Human IFN-β, IL-6, and IL-8 ELISA kits were purchased from R & D Systems, and ELISA was carried out according to the manufacturer's instruction (**Appendix 7.2**). 1:10 and 1:50 dilutions were made using pre-prepared Reagent Diluent on cell supernatant to obtain cytokine concentrations within standard ranges.

2.8 Invasion Assay

2.8.1 Preparation of Bacterial and Human HT-29 Cells for Invasion

HT-29 cells were prepared in the following way: 3 days prior to the invasion assay, HT-29 cells were plated at 50% confluence, in a 24-well plate with 500μL per well of DMEM medium supplemented with 10% FCS and 1% Penicillin/Streptomycin and left to grow into a confluent monolayer. 24 h prior to invasion, the cells were washed twice with PBS and replaced with 500 μL DMEM with 10% FCS but without antibiotics. The prepared cells were examined under the microscope to confirm proper attachment and 100% confluence. Cell count at 100% confluence was pre-determined to be at 8×10^5 cells/well.

For preparation of active bacterial cultures, all agar plates and liquid medium were pre-incubated in the anaerobic chamber for at least 24 h prior to cultivation of *Fusobacterium nucleatum* and *Clostridium perfringens*, or warmed up in a 37°C incubator for cultivation of *Escherichia coli* DH5α and *Staphylococcus epidermidis*. Frozen stocks of bacteria were streaked onto Columbia blood agar plates in several dilutions with a cotton swab to obtain single colonies. Then, a single colony was added to liquid medium (LB medium for *Escherichia coli* DH5α, Brain-Heart Infusion medium supplemented with 5 μg/mL Hemin and 1 μg/mL Vitamin K for

Fusobacterium nucleatum, *Clostridium perfringens* and *Staphylococcus epidermidis*) until

visible growth was observed. The bacteria were then diluted 1:100 and incubated for an

appropriate amount of time (**Methods 2.1.6**) to obtain log phase growth at time of invasion.

2.8.2 Determination of Antibiotics Sensitivity of Select Bacterial Strains

Strains from four different species of bacteria, *Escherichia coli* DH5α, *Fusobacterium*

nucleatum (DSM 15643), *Clostridium perfringens* (DSM 756) and *Staphylococcus epidermidis*

(DSM 20044) were selected for antibiotics sensitivity determination. A range of antibiotics were

used, including 1 % Penicillin and Streptomycin, 300 µg/mL of Gentamicin and 200 µg/mL of

Metronidazole, a cocktail combination of all four antibiotics, and 100 µg/mL of Cefoxitin, in

duplicates. Bacteria were plated at 8 x 10^6 bacteria in wells with or without HT-29 cells in 500

µg/mL of antibiotics-free DMEM medium with 10% FCS, and incubated at 37°C, 5% CO_2 for 4

h. Supernatant was then collected and centrifuged at 5000 rpm for 10 min, washed twice with

PBS with centrifugation in between, and resuspended in medium containing the appropriate

antibiotics. The wells containing HT-29 cells were washed twice with PBS and replaced with

medium containing antibiotics. The centrifuged bacteria from the supernatant with antibiotics

medium were placed in a fresh plate. The plates containing cells and the fresh plate containing

bacteria were incubated in 37°C for 1 h. Negative controls wells without antibiotics were

performed in an identical manner except that the bacteria and the cells containing bacteria were

replaced with fresh DMEM without antibiotics.

Afterwards, the supernatant with bacteria was removed, centrifuged for 10 min at 5000

rpm, washed once with PBS, re-centrifuged, and resuspended with 500 µL of PBS. For wells

with HT-29 cells, the medium was removed, the cells washed twice with PBS, and 500 µL of

water was added per well to lyse the cells for 45 min. The lysed cells were collected and centrifuged at 5000 rpm for 10 min, then resuspended in 500 µL of PBS. Bacteria from the supernatant and the cell-containing faction were serially diluted 1:10 to four dilutions: 10^0, 10^{-1}, 10^{-2} and 10^{-3}, and 25 µL of each dilution was streaked onto half of a pre-incubated Columbia blood agar plate, with duplicates of the same condition and dilution factor plated onto each half of the plate. The plates were stored in 37°C incubator (for *Escherichia coli* and *Staphylococcus epidermidis*) or anaerobic chamber set at 37°C (for *Fusobacterium nucleatum* and *Clostridium perfringens*) until visible colony growth was observed.

2.8.3 Combined Invasion and TLR9 Activation Experiment on HT-29 Cells

Fusobacterium nucleatum (DSM 15643) was selected as the invasive bacterial species for combined invasion and TLR9 activation experiments, with *Escherichia coli* DH5α as the non-invasive controls. HT-29 cells and bacterial cultures were prepared via procedures outlined in **Methods 2.8.1**. Directly before invasion, 1mL of active bacteria culture was taken from the liquid culture for OD measurement, and invasion was performed at a MOI of 5:1 bacterial to HT-29 cells. Bacteria was pre-diluted with DMEM and 10% FCS such that 4×10^6 bacteria were in 450µL of medium. Cells were washed twice with PBS prior to the bacteria addition.

The remaining 50 µL to be added consisted of serum-free DMEM for the negative control condition, 5 µg/mL of ODN-2006, 20 µg/mL of Caco-2 DNA, or a combination of ODN-2006 and Caco-2 DNA. The four conditions were used in wells containing *Fusobacterium nucleatum*, plus an extra plate consisting of HT-29 cells without addition of bacteria, in duplicates. For wells containing *Escherichia coli* DH5α, only the negative control condition

consisting of serum-free DMEM was used. Both plates (one with bacteria, one without) were centrifuged for 15 min at 100 x g, and placed in the 37°C incubator for 4 h.

Afterwards, supernatants containing bacteria were passed through a 0.2µm filter and collected onto a fresh 24-well plate. Supernatants were also collected from the plate with cells but no bacteria. Collected supernatant was frozen in -20°C in 24-well plates for human IL-8 ELISA (**Methods 2.7.4**). The plate without bacteria was discarded, while the plate with bacteria was washed twice with PBS, then replaced with 500µL medium supplemented with 300 µg/mL Gentamicin and 200 µg/mL Metronidazole, and incubated for 1 h. Cells were washed twice again with PBS, and lysed with 1% Saponin (500 µL/well) for 10 min. The lysed cells with any invaded bacteria were then centrifuged at 5000 rpm for 15 min. Saponin was removed and replaced with 500µL PBS, and serial dilutions of 10^0 to 10^{-3} were made. 50µL of bacterial dilution was spread onto pre-prepared Columbia Blood Agar plates in appropriate atmospheric conditions to determine colony forming units (CFU) for all conditions (*Escherichia coli* DH5α negative control; *Fusobacterium nucleatum* negative control, ODN-2006, Caco-2 DNA, and combined ODN-2006 and Caco-2 DNA).

2.9 Bioinformatics and Data Analyses

2.9.1 Determination of CpG and K-mer Concentration in Bacteria

Concentration of bacterial CpG and other sequence patterns, including that of CpG-containing k-mer sequences up to 8-mer and potentially repressive k-mer sequences, genome size, and significant ratios such as % CpG / % G+C, were determined by Dr. Daniel Podlesny from the Department of the Microbiome and Applied Bioinformatics at the University of

Hohenheim via custom Python scripts[46]. Full bacterial FASTA genomes were downloaded from the NCBI RefSeq database (https://www.ncbi.nlm.nih.gov/refseq/), including chromosomal and plasmid sequences, and the results collected into a comprehensive Excel database. The type strain of each species was used for k-mer determination.

To determine k-mer concentrations, only one strand of the two complementary strands were used, and overlapping sequences were counted—that is, for k-mers with a length of >5 bp, a sequence with Five bases would contain five, typically distinct, overlapping 5-mers. For selection of bacterial genomes used for stimulation, bacteria spanning a range of CpG concentrations were primarily considered, followed by their significance as members of adult and infant gut commensals or pathogens, and finally the concentration of human TLR9-stimulatory 4-mer TCGT and 5-mer TCGTT.

2.9.2 Statistical Analyses of Technical and Biological Replicates

Technical replicates were generated in each TLR9 activation and bacterial invasion experiment, and average and standard deviation were calculated for all conditions. TLR9 activation was normalized to a standard concentration of positive control ODN-2006, pre-determined via a titration curve of cells treated with serially diluted ODN-2006. In TLR9 co-activation experiments, additional controls of fixed concentrations of ODN-2006 together with short ODNs TCGTT were included, and for TLR9 repression experiment, an additional control of ODN-2006 with the inhibitory ODN INH-18 was included.

Statistical significance of biological replicates was determined for all conditions where more than three independent experiments were carried out. Significance of correlation was determined via Spearman test, while comparison of individual conditions was determined via

either Wilcoxon Mann-Whitney test or Student's T-test, adjusted with Benjamini-Hochberg correction to 5% false discovery rate (FDR). Plots were generated via R's ggplot function (from the /tidyverse package), and statistical analyses via standard R scripts.

2.9.3 Correlation of TLR9 Activation with Bacterial Genomic CpG and Significant K-mer Concentration

Activation and ODN-2006 co-activation of human TLR9 by bacterial genomic DNA was plotted as a x-y scatterplot via R's ggplot2 function, where the x-axis represents genomic CpG or CpG-containing k-mer concentration and y-axis represents TLR9 activation compared to a standard reference of a fixed concentration of ODN-2006. Correlation coefficient and significance of correlation were determined via Spearman test.

3. Results

3.1 TLR9 Activation of HEK-Dual TLR9 Cells Correlated to Bacterial Genomic CpG Concentration

3.1.1 CpG and CpG-Containing K-Mer Concentrations of Bacterial Genomes

Bacterial CpG and CpG-containing k-mer concentrations were calculated via bioinformatics methods (**Methods 2.9.1**) on raw FASTA sequences downloaded from the RefSeq database of National Center for Biotechnology Information (NCBI) by Dr. Daniel Podlesny[46], and 13 bacterial species were selected for TLR9 activation experiments (**Table 2**). Bacterial strains were chosen to represent a wide range of genomic CpG and CpG-containing k-mer concentrations due to the sequences' predicted effect on human TLR9 activation[46,50], and for their relevance as members of adult and infant gut commensals or pathogens. Type strains for a given bacterial species were chosen and purchased from the German Collection of Microorganisms and Cell Culture GmBH (DSMZ) unless otherwise noted in **Methods 2.1.1**.

Within this set, genome sizes varied widely between specie, ranging from 1.64×10^6 bp for *Campylobacter jejuni* to 5.49×10^6 bp for *Bacteroides dorei*[46]. The concentration of Guanine and Cytosine bases (% G+C) ranged from 73.00% for *Micrococcus luteus* to 27.15% for *Fusobacterium nucleatum*[46]. Strong correlation between genomic % G+C and CpG concentration was found: except for *Staphylococcus Epidermidis* and *Lactobacillus salivarius*, higher % C+G was associated with higher CpG concentration. Correlations were weaker between % G+C with the CpG-containing 4-mer % TCGT (CpG flanked by one Tyrosine base at both ends) and 5-mer % TCGTT (CpG flanked by one Tyrosine base at 5' end and two Tyrosine bases at the 3' end, previously determined as relevant for TLR9[56].

CpG concentration of the selected species ranged from 14.97% for *Micrococcus luteus* to 0.2958% for *Fusobacterium nucleatum*, a difference of greater than 50-fold[46]. While *Micrococcus luteus* DNA had the highest CpG concentration of them all and was used in several of our simulation experiments, it was known primarily as an environmental bacterium and skin commensal as opposed to a gut commensal or pathogen[111,112]. Among the common gut commensals, *Bifidobacterium bifidum* had the highest CpG concentration at 12.40%. It was known to be present at high level in the developing infant gut but represented only a minority in the adult gut[52]. Among the low-CpG bacteria, *Fusobacterium nucleatum* was known as a common oral bacterial species and an opportunistic gut pathogen with invasive characteristics [102,113-115], while *Campylobacter jejuni*[116], *Clostridioides difficile*[117], and *Clostridium perfringens*[118] were implicated in pathogenic conditions such as gastroenteritis or colitis.

Table 2. List of Bacterial Genomes Used for TLR9 Activation Experiments and Their Significant K-mer Concentrations[46]. Bacterial genomes were ordered by genomic CpG concentration specific to the strain used for TLR9 activation experiments. Other relevant genomic characteristics, such as % G+C, % GpC, and % CpG-containing 4-mers and 5-mers, are included.

Species (Strain)	Genome Size(Mb)	% G+C	% CpG	% GpC	% TCGT	% TCGTT	CpG/GpC
Micrococcus luteus (DSM 1790)	2.50	72.99525	14.97232	12.9634	1.121989	0.079965	1.154969
Bifidobacterium bifidum (DSM 20239)	2.21	62.66644	12.40225	10.6771	1.154763	0.191452	1.161574
Bifidobacterium breve (DSM 20213)	2.27	58.89029	9.932873	9.632134	0.958706	0.209878	1.031222
Escherichia coli (K12/C600)	4.64	50.79071	7.471329	8.275116	0.622473	0.195168	0.902867
Lactobacillus ruminis (DSM 20403)	2.06	43.46954	5.989091	5.817943	0.991808	0.372782	1.029417
Bacteroides dorei (DSM 17855)	5.49	41.98595	4.044603	4.656586	0.522493	0.168339	0.868577
Enterococcus faecalis (DSM 20478)	3.36	37.38115	3.46643	4.303907	0.686198	0.30042	0.805415
Staphylococcus epidermidis (DSM 20044)	2.56	32.04594	2.19827	2.927855	0.533301	0.202762	0.750812
Lactobacillus salivarius (DSM 20555)	2.13	33.04047	1.863285	3.221878	0.444618	0.156142	0.578323
Campylobacter jejuni (DSM 4688)	1.64	30.54857	1.439128	4.071752	0.245023	0.095889	0.353442
Clostridioides difficile (DSM 1296)	4.30	29.05841	0.722267	2.459301	0.173704	0.059794	0.293688
Clostridium perfringens (DSM 756)	3.26	28.37915	0.481871	2.425567	0.095895	0.032702	0.198663
Fusobacterium nucleatum (DSM 15643)	2.17	27.15159	0.295839	2.2413	0.087561	0.034399	0.131994

3.1.2 DNA Pre-Processing and Fragment Length Standardization

Bacteria were grown and harvested (**Methods 2.1.2-2.1.5 and 2.1.7**) and the DNA extracted via Zymo Quick-DNA Fungal and Bacterial Miniprep Kit. As experiments had shown variable TLR9 activation due to DNA fragment length on human PBMCs (**Results 3.6.1**), an attempt was made to obtain standard fragment lengths of 50-200 bp using enzymatic fragmentation with NEB's dsDNA Fragmentase (**Methods 2.4.1**). Following recommendations from company[119] and taking account of an earlier test run with *Escherichia coli* and *Campylobacter jejuni* genomic DNA showing variation of fragmentation time (**Supplementary Figure S6**), a time course was performed on all bacterial genomes. Twelve genomes were selected for test fragmentation, ten of which were shown below. While 30 min was sufficient for most genomes to obtain the target fragment length of 50-200 bp, several genomes (both *Bifidobacterium* species, *Clostridium perfringens* and *Micrococcus luteus*) required 60 min, and one genome (*Clostridioides difficile*) required 120 min (**Figure 1**). The varied fragmentation times needed for different bacterial genomes did not depend on genomic G+C and CpG concentrations, but could potentially be due to other sequence patterns[120] or presence of inhibitors[121]. The lack of an optimal fragmentation time for all microbial genomic DNA to one standard target fragment length could implicate the unsuitability of the NEB's dsDNA Fragmentase in generating length-standardized DNA fragments from metagenomic DNA samples.
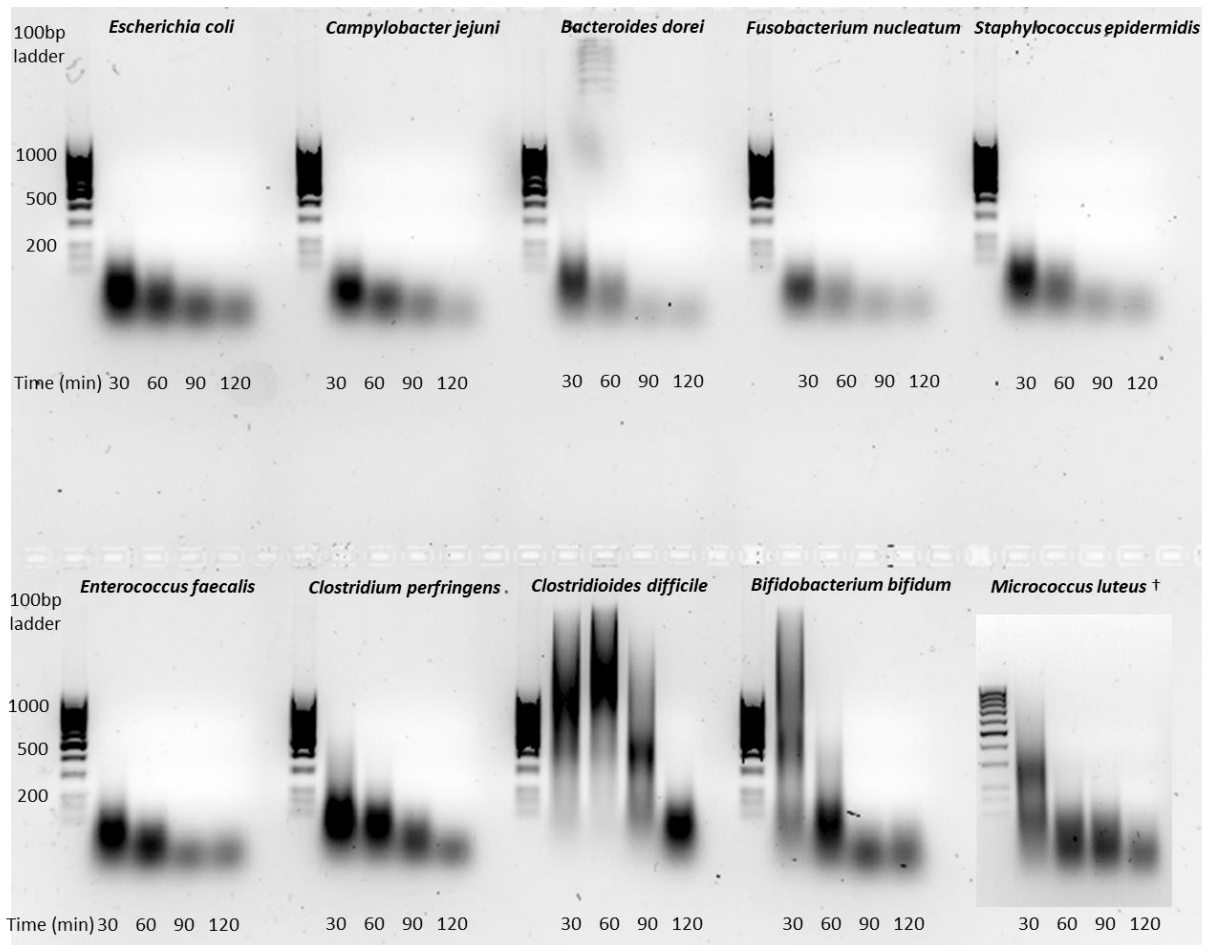
Figure 1. Time course of Bacterial DNA Treated with NEB dsDNA Fragmentase. 3 µg of bacterial genomic DNA were treated with NEB dsDNA Fragmentase for 30, 60, 90, and 120 min, and loaded onto an 1.5% agarose gel together with the 100-bp ladder. The desired fragment length of 50-200 bp was obtained at different time points depending on species origin. 30 min of fragmentation time was used in subsequent experiments for *Escherichia coli*, *Campylobacter jejuni*, *Bacteroides dorei*, *Fusobacterium nucleatum*, *Staphylococcus epidermidis* and *Enterococcus faecalis*. 60 min of fragmentation time was used for *Bifidobacterium bifidum*, *Bifidobacterium breve, Clostridium perfringens* and *Micrococcus luteus*. 120 min of fragmentation time was used for *Clostridioides difficile*. († *Micrococcus luteus* fragmentation and gel electrophoresis was performed at a later date under identical conditions.)

3.1.3 Strong Positive Correlation was Found Between TLR9 Activation and Bacterial Genomic CpG Concentration Across Nine Bacterial Genomes

Nine fragment length-standardized (**Results 3.1.2**) bacterial genomic DNA spanning a

range of high (such as *Bifidobacterium bifidum* and *Bifidobacterium breve*), moderate (such as

Bacteroides dorei and *Enterococcus faecalis*), and low (such as *Clostridium perfringens* and

Fusobacterium nucleatum) CpG concentration and relevant as members of the human microbiota

(**Results 3.1.1**) were chosen to stimulate human TLR9. Genomic DNA was pre-complexed with

10 µg/mL of DOTAP and added to HEK-Dual TLR9 cells, a line of HEK (human embryonic

kidney) cells engineered with secreted embryonic alkaline phosphatase (SEAP) reporter protein

downstream of NF-kB and Lucia Luciferase reporter gene transcribed by the human IL-8

promotor[105]. HEK cells with engineered TLR9 and downstream reporters had previously been

used in studies of TLR9 activation by ODNs and bacterial genomic DNA[50,63] and the engineered

TLR9 responded predictably to both ODNs and microbial genomic DNA[50,122], thus the HEK-

Dual TLR9 cell line represents a simplified cell model where the engineered response can

exclusively be inferred as a result of TLR9 activity compared to the more complex response of

natural immune cells possessing multiple innate immune receptors[50]. The cells were plated at 1 x

10^5 cells per well, at the final bacterial DNA concentrations of 5 and 2.5 µg/mL, in duplicates.

Figure 2 shows NF-kB activity as SEAP, a signal of TLR9 activation in HEK-Dual TLR9 cells,

compared to the positive control of cells treated with 0.5 µg/mL of ODN-2006, 20 h later. The

genomes are listed in decreasing order of CpG concentration in **Figure 2 a)** and **b)**, and the

background signal was measured in wells treated with DOTAP alone in the same plate, while **c)**

and **d)** correlates NF-kB activity with genomic TLR9 concentration.

Results from four independent experiments (two for *Bifidobacterium breve*) indicated

robust positive correlation of TLR9 activation with genomic CpG concentration ($\rho = 0.825$ and

0.850, and $p = 1.93 \times 10^{-10}$ and 1.49×10^{-11}, respectively, for 5 and 2.5 µg/mL of DNA,

correlation and significance determined via Spearman test). *Bifidobacterium bifidum* DNA,

which has the highest CpG concentration of 12.4% in the experiments, induced high TLR9

activation at 65.1% and 37.5% of ODN-2006 standard using 5 and 2.5 µg/mL of DNA.

Conversely, the low-CpG genomes of *Clostridium perfringens* and *Fusobacterium nucleatum* induced little TLR9 activation (respectively 13.2% and 10.4% of ODN-2006 for 5 µg/mL of DNA, and 7.67% and 8.66% of ODN-2006 for 2.5 µg/mL of DNA, the latter of which showed no statistically significant difference to untreated wells).
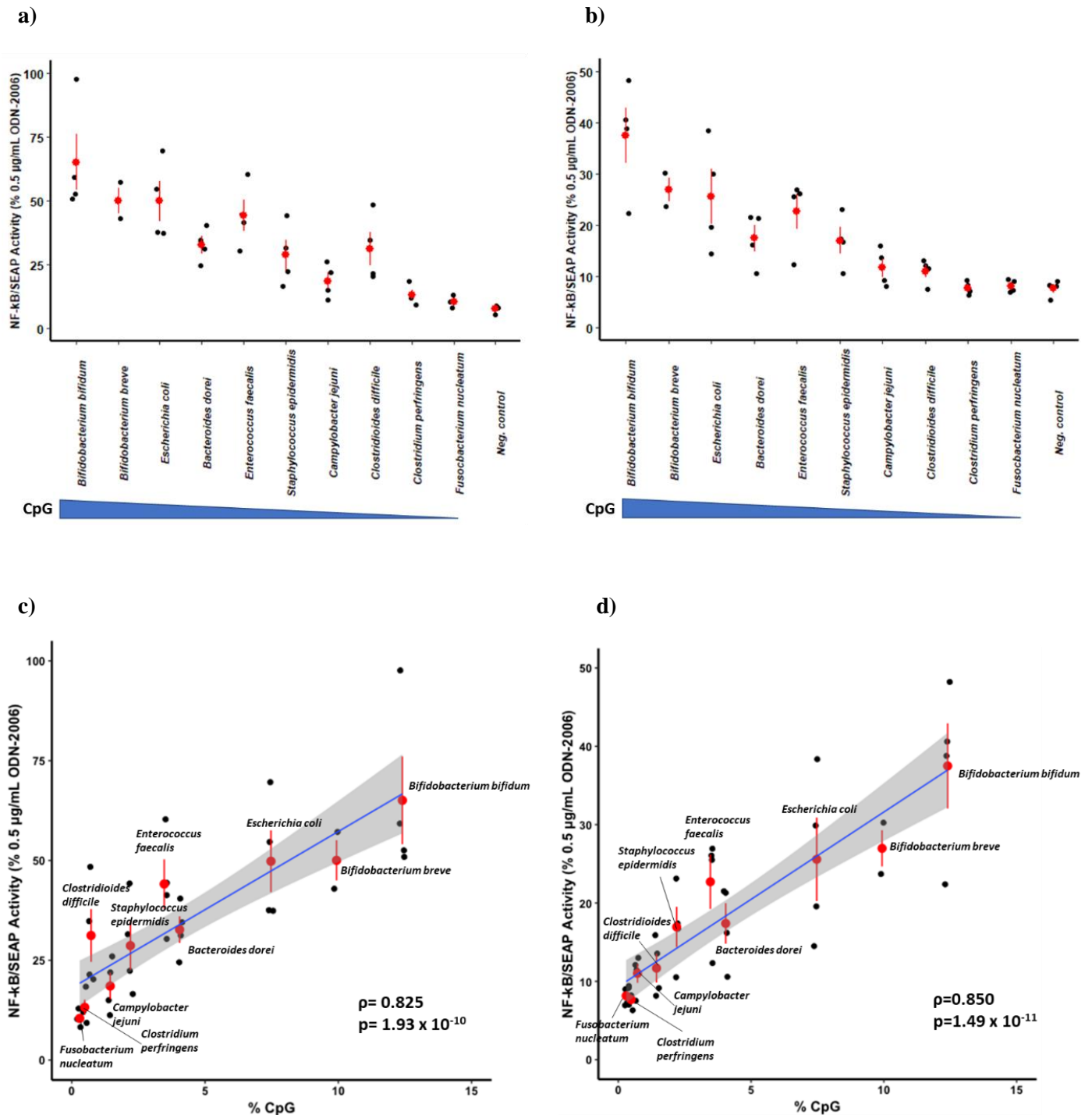
Figure 2. TLR9 Activation of HEK-Dual TLR9 Cells Stimulated with Bacterial Genomic DNA. HEK-Dual TLR9 cells were stimulated with bacterial genomic DNA fragments 50-200 bp of length. Comparison of ten different bacterial genomes was made for **a)** 5 µg/mL and **b)** 2.5 µg/mL of input DNA (final concentration in wells). Correlations with genomic CpG concentration for the two input concentrations were shown in **c)** and **d)**, respectively. TLR9 activation as Secreted Embryonic Alkaline Phosphatase (SEAP) in the supernatant was compared to plate positive control of 0.5 µg/mL ODN-2006. Results from four independent experiments (two experiments for *Bifidobacterium* breve) were shown, with each dot representing results of one experiment and bars representing ± s.e.m. Correlation coefficient ρ and significance value p were determined via Spearman's rank correlation test.

3.1.4 Experimental Variations to CpG/TLR9 Activation Correlations Suggested Additional Influences on Human TLR9 Activation Besides CpG Concentration

We found the correlation of TLR9 activation to genomic CpG concentration to be highly

significant, but evidence of experimental variations or other factors beyond CpG affecting human

TLR9 activation was found, as the activation signal of several genomes did not conform to the

overall positive correlation. For example, the genomic DNA of *Enterococcus faecalis*, whose

CpG concentration is at a moderate 3.47%, activated human TLR9 at a similar range to

Bifidobacterium breve (9.93% CpG) and *Escherichia coli* (7.47% CpG), and was significantly

more activating than the genomic DNA of *Bacteroides dorei* (4.04% CpG) at both 5 and 2.5

µg/mL. The genomic DNA of *Clostridioides difficile*, at 0.72% of CpG, was more activating of

TLR9 than expected from the correlation at 5 µg/mL, but not at 2.5 µg/mL.

The genomic DNA of *Micrococcus luteus*, which has a CpG concentration of 14.97%,

was another genome that did not conform to the above correlation (**Supplementary Figure S7**),

but in the opposite direction. Result of one experiment, which included *Micrococcus luteus* DNA

processed in the same manner, indicated that the genomic DNA of *Micrococcus luteus* was less

activating than that of both *Bifidobacterium bifidum* and *Escherichia coli* for both concentrations,

and less activating than *Enterococcus faecalis* DNA at 5 µg/mL. The exceptional status of the

Micrococcus luteus DNA suggested that genomic CpG concentration alone might not be the only genomic pattern the human TLR9 responds to.

3.2 Correlation of TLR9 Activation to Genomic CpG Concentration Was Confirmed on Human Peripheral Blood Mononuclear Cells (PBMCs)

Four bacterial genomic DNA, *Bifidobacterium bifidum*, *Escherichia coli*, *Lactobacillus salivarius* and *Campylobacter jejuni* were used to stimulate human PBMCs. *Bifidobacterium bifidum* and *Lactobacillus salivarius* were chosen as representatives of high (12.40%) and low (1.86%) CpG members of the healthy human gut microbiota, particularly in infants, while *Escherichia coli* and *Campylobacter jejuni* are members of the human microbiota that can become pathogenic[116,123] with moderately high (7.47%) and low (1.44%) genomic CpG concentration, respectively. The use of human PBMCs represented TLR9 activation in a more natural and physiologically relevant context compared to engineered HEK cells[124]. Extracted bacterial genomic DNA was enzymatically fragmented to 1000 bp of length, gel-extracted, and added to human PBMCs extracted from fresh blood sample from a young adult in three different concentrations: 2.5, 0.5 and 0.1 µg/mL (pre-complexed with 12.5 µg/mL of DOTAP per well). TLR9 activation was measured via ELISA as release of human Type I interferon-beta (IFN-β) into the supernatant (**Figure 3**). IFN-β was determined to be specific to TLR9 activation via control wells treated with TLR9-stimulatory ODNs versus 10 ng/mL of *Escherichia coli* LPS, which induced TLR4 activation and release of IL-6 but not IFN-β (**Supplementary Figure S5**).

The strongest correlation of IFN-β release by human PBMCs to bacterial CpG concentration occurred at the input DNA concentration of 0.5 µg/mL. A plot of IFN-β concentration in the supernatant versus genomic CpG concentration for this DNA concentration is shown in **Figure 3 b)**. At 2.5 µg/mL, however, *Bifidobacterium bifidum* DNA induced lower-

than-expected IFN-β release in the supernatant. Cellular IFN-β release from treatment with

Lactobacillus salivarius DNA was comparable to cells treated with *Escherichia coli* DNA

(**Figure 3 a**)). No difference was found between the genomes at 0.1 µg/mL of input bacterial

DNA. Our results suggested that the correlation of human TLR9 activation to Genomic CpG

concentration was consistent across different cell types. The patterns of TLR9 activation in

engineered HEK-Dual TLR9 cells could therefore be confirmed in a more natural,

physiologically relevant model.
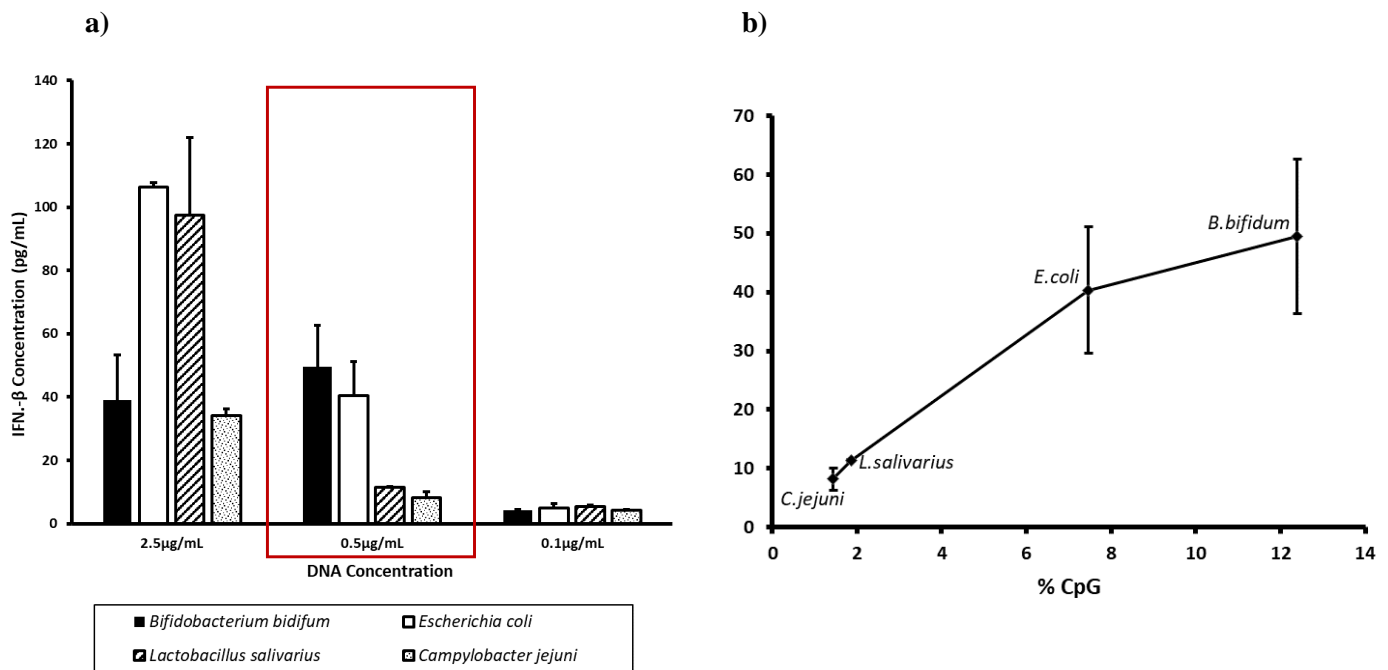
Figure 3. TLR9 Activation of Human Peripheral Blood Mononuclear Cells (PBMCs) Stimulated with Bacterial Genomic DNA.

a) **IFN-β Production by PBMCs Treated with 1000 bp Long Bacterial DNA Fragments in Three Different Concentrations.** PBMCs were treated with bacterial DNA at 1000 bp of average fragment length. DNA from four bacterial species were added to cells to the final concentrations of 2.5, 0.5 and 0.1 µg/mL together with DOTAP. IFN-β release into the supernatant was measured via ELISA. Result of 1 experiment with bars representing ± s.d. of duplicates.

b) **Correlation of IFN-β production by PBMCs and Bacterial CpG Concentration.** IFN-β production was plotted against % CpG of different bacterial species added to cells at 0.5

µg/mL final concentration (indicated by a red box in **Figure 3 a)**). Bars represent ± s.d of duplicates.

3.3 Concentration of CpG-Containing K-Mer Showed Additional Effects to Human TLR9 Activation

3.3.1 Higher CpG Concentration Did Not Necessarily Induce Higher Human TLR9 Response: Results from Stimulation of HEK-Dual TLR9 Cells by Five Bacterial Genomes Fragmented to 3000 Bp

The genomes of five bacterial species, *Micrococcus luteus*, *Bifidobacterium bifidum*, *Escherichia coli*, *Lactobacillus ruminis* and *Lactobacillus salivarius* were fragmented to 3000 bp of length using the Covaris S220 Ultrasonicator with miniTUBES (**Methods 2.4.3**) and left-hand size selection with SPRI Beads (**Methods 2.5.2**). The five genomes were chosen due to their unusual CpG-containing 4-mer TCGT and 5-mer TCGTT concentrations, which did not directly correspond to their genomic CpG concentration. Genomic DNA was complexed with 20 µg/mL DOTAP and added to HEK-Dual TLR9 cells to three different final concentrations: 5, 2.5 and 1.25 µg/mL. TLR9 activation was determined as SEAP production and Lucia Luciferase, an engineered Luminescence reporter under the control of human IL-8 promoter signaled by NF-kB and AP-1.

Results from SEAP using these five genomes exhibited no positive correlation of TLR9 activity to genomic CpG concentration ($\rho = -0.1637$, $p = 0.56$, see **Figure 4a)**). Similar lack of correlation was observed when Lucia Luciferase was used as an alternative reporter for TLR9 activity (**Supplementary Figure S8**). The results suggested that unusual CpG-containing k-mer concentrations could become a source of noise interfering with direct correlation of TLR9

activation with bacterial CpG concentration observed in **Results 3.1.3.**, and additional genomic patterns could play a role in human TLR9 activation.

3.3.2 Human TLR9 Activation Positively Correlated to CpG-Containing 5-mer TCGTT

As the correlation of TLR9 activation against genomic CpG concentration was found to be weak using the aforementioned five genomes (**Results 3.3.1**), additional genomic patterns, that of CpG-containing k-mers, were examined. The CpG-ODN used for human TLR9 activation, ODN-2006, is a 24-nucleotide sequence containing four TCGT 4-mers i.e., the CpG dimers are flanked by Tyrosine bases on both ends. Three of the motives are also part of the TCGTT 5-mer. What is significant concerning *Micrococcus luteus*, the genome previously found to be an exception in TLR9 activation in spite of its high genomic CpG concentration (**Results 3.1.4**), is its low concentration of TCGTT 5-mer, which at 0.0800% is the lowest of the five genomes tested. The genome of *Lactobacillus ruminis*, on the other hand, had exceptionally high TCGTT concentration at 0.373% (**Table 2**).

Figure 4 c) and d) correlated TLR9 activation as SEAP production with genomic TCGT 4-mer and TCGTT 5-mer concentration. Positive correlations were found relative to TCGTT 5-mer concentration (ρ=0.535, p = 0.04). Similar correlation was noted using the Lucia Luciferase reporter assay (**Supplementary Figure S8**). The high TCGTT 5-mer concentration of 0.300% could also explain the higher TLR9 activation of *Enterococcus faecalis* genomic DNA relative to its CpG concentration of 3.47% (**Results 3.1.4**). The results suggested that the flanking Tyrosine bases around the CpG dimer aided in the human TLR9 response, and could be an explanation for the exceptional status of several genomes such as *Micrococcus luteus* and *Enterococcus faecalis* in activating human TLR9 (**Results 3.1.4**) compared to expected via their genomic CpG concentration.
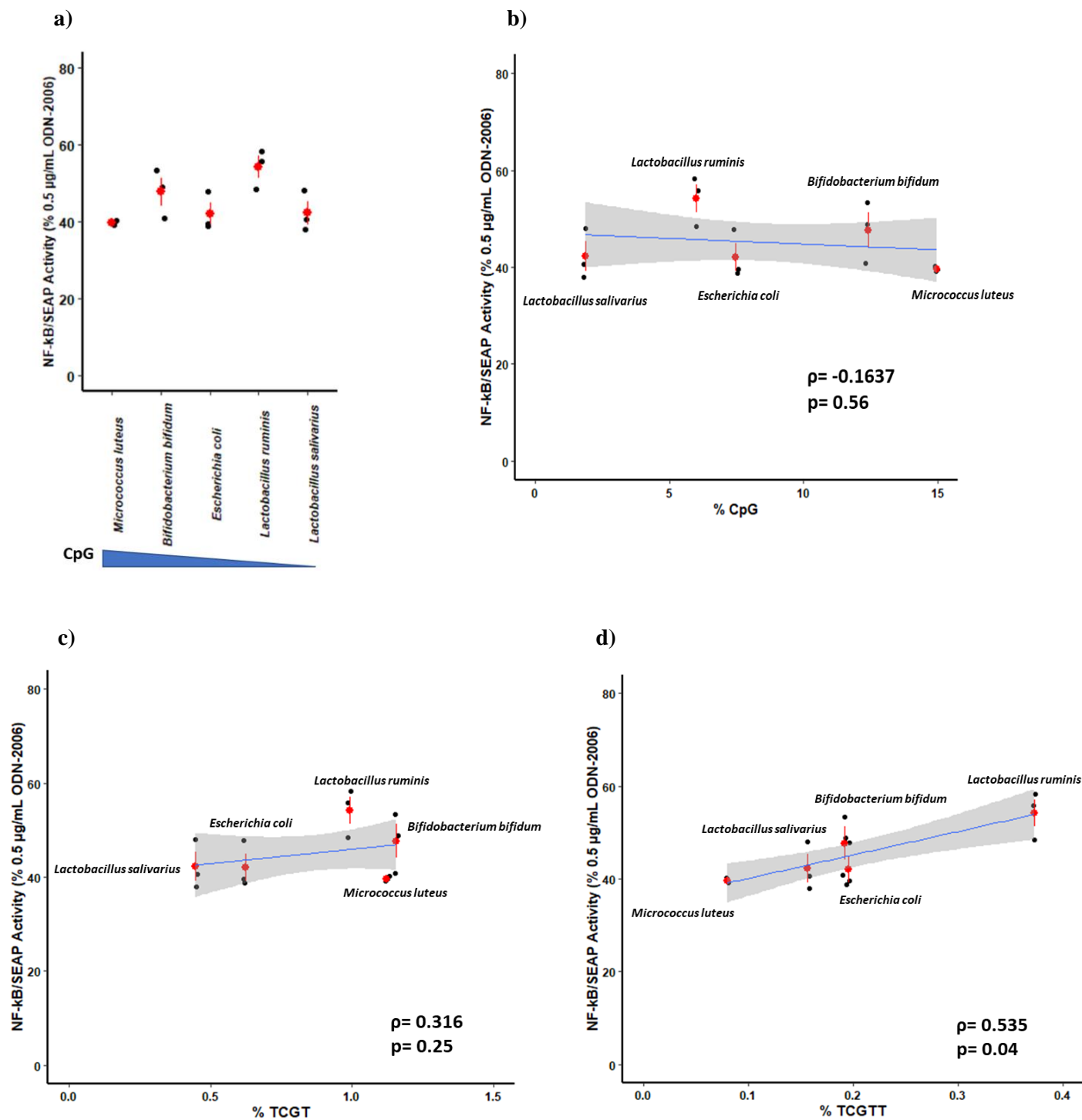
Figure 4. Correlation of TLR9 Activation of HEK-Dual TLR9 Cells to Genomic K-mer Concentrations. HEK-Dual TLR9 cells were stimulated with 5 µg/mL (final concentration) of bacterial DNA from five bacterial species of diverse CpG and CpG-containing k-mer concentration sheared to 3000 bp of length with DOTAP. **a)** SEAP production of 5 genomes, arranged in the order of decreasing CpG concentration. **b)** Correlation of SEAP production with

genomic CpG concentration, **c)** genomic TCGT 4-mer concentration and **d)** genomic TCGTT 5-mer concentration. Results were normalized to plate positive control of 0.5 µg/mL ODN-2006. Results from three independent experiments, with each dot representing results of one experiment and bars representing ± s.e.m. Correlation coefficient ρ and significance p were determined via Spearman's rank correlation test.

3.4 Short Synthetic Oligonucleotide TCGTT Was a Human TLR9 Co-Activator with Different DNA Types

3.4.1 TCGTT Co-Activated Human TLR9 with Synthetic Oligonucleotides ODN-2006 and ODN-2006GC

Pohar et al. (2017) showed short CpG-containing synthetic oligonucleotides less than 10 bp long to be non-activating of human TLR9 on their own, but co-activated human TLR9 of cells treated together with CpG-containing ODNs at low concentrations[63]. We verified TLR9 co-activation by 20 and 100 µg/mL of TCGTT using Ramos Blue cells (an engineered human B cell line) plated at 2 x 10^5 cells/well and co-treated with a low concentration of 2.5 µg/mL of ODN-2006. No DOTAP was used in experiments with Ramos Blue cells due to observed toxicity (**Supplementary Figure S9**).

TLR9 activation was measured as SEAP under the control of NF-kB, a transcription factor known to respond to TLR9 activation in B cells[42], 20 h later, compared to a standard plate positive control of 25 µg/mL ODN-2006. We additionally observed co-activation of TLR9 by TCGTT together with the non-activating GpC control of ODN-2006, ODN-2006GC, and a lack of TLR9 co-activation by TTTTT with either ODN-2006 or ODN-2006GC (**Figure 5**).

a)
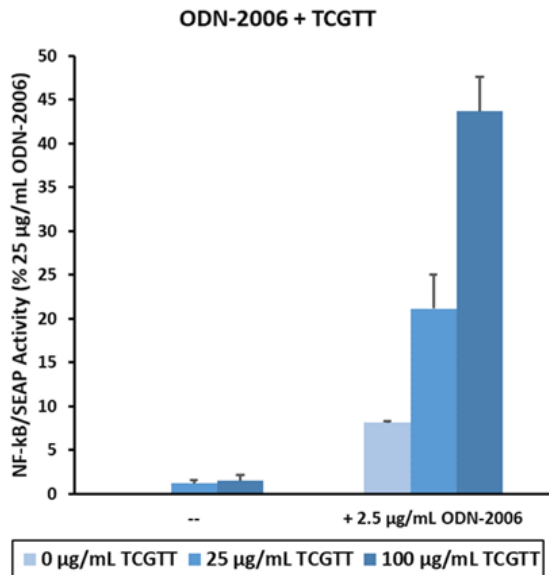
b)
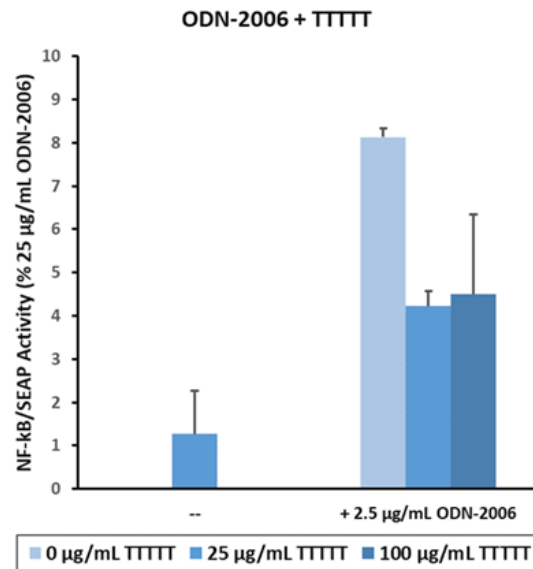
c)
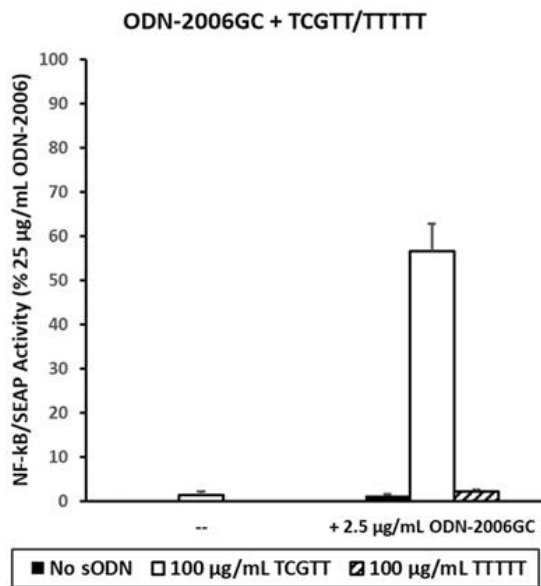
Figure 5. Effect of TCGTT and TTTTT on ODN-2006 and ODN-2006GC-Induced Human TLR9 Response. 25 and 100 µg/mL of **a)** TCGTT or **b)** TTTTT was added to 2.5 µg/mL of ODN-2006. **c)** 100 µg/mL of TCGTT or TTTTT was added to 2.5 µg/mL of ODN-2006GC. TCGTT or TTTTT was also added alone as controls. All experiments were performed on Ramos Blue cells and TLR9 response measured as NF-kB activity relative to plate positive control of 25 µg/mL ODN-2006 (Mean from two independent experiments, bars ± s.e.m.).

3.4.2 TCGTT Co-Activated Human TLR9 with Caco-2 DNA Fragmented to Different Lengths

Co-activation was found when 100 µg/mL of TCGTT was combined with 20 µg/mL of human DNA extracted from Caco-2 cells. DNA extracted from intact Caco-2 cells without additional fragmentation (>10,000 bp long), enzymatically fragmented (200-6000 bp or 50-200 bp) via NEB dsDNA Fragmentase (**Methods 2.4.1**), as well as DNA extracted from apoptotic bodies and blebs from Staurosporine-treated Caco-2 Cells (**Methods 2.2.4, 2.2.5, and 2.3.2**), were used, performed on Ramos Blue cells plated at the density of 2×10^5 cells/well and pre-starved for 16 hours in 0.2% Bovine Serum Albumen (BSA). The two target fragment lengths of human DNA via enzymatic fragmentation were chosen to approximate the typical fragment lengths of apoptotic body and apoptotic bleb DNA and confirmed via PAGE gel and Agilent Bioanalyzer (**Methods 2.6.2 and 2.6.3, Supplementary figure S2**). Results are shown as SEAP production relative to plate positive control of 5 µg/mL ODN-2006. Co-activation of human TLR9 by TCGTT with Caco-2 DNA was found to be lower than that with either ODN-2006 or ODN-2006GC, and no difference was observed between the different fragment lengths and types of Caco-2 DNA (**Figure 6**). The result suggested that TCGTT co-activated TLR9 with human DNA of varying qualities and without regard to whether the human DNA originated from intact or apoptotic cells.
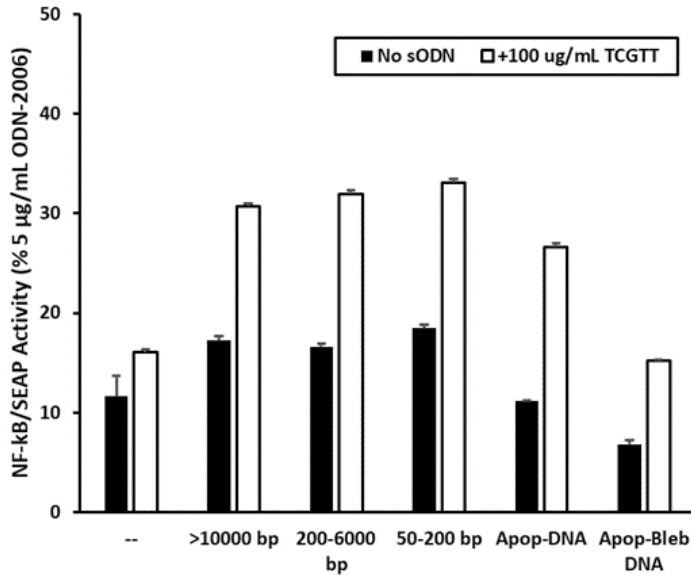
Figure 6. Co-Activation of human TLR9 Using TCGTT with Human DNA Fragments. 100 µg/mL of co-activating short DNA TCGTT was used to co-activate human TLR9 together with Caco-2 DNA fragments of the following lengths: >10,000 bp, 200-6000 bp, 50-200 bp, and DNA derived from apoptotic Caco-2 cell body and blebs, all at 20 µg/mL of concentration, on overnight-starved Ramos Blue cells plated at 2 x 10^5 cells/well. TLR9 activation was measured as SEAP activity normalized to plate positive control of 5 µg/mL ODN-2006 (results from one experiment, bars ± s.d. of duplicates).

3.5 DNase I-Digested Bacterial DNA Co-Activated Human TLR9 With Low Concentrations of ODN-2006 in Ramos Blue Cells

3.5.1 DNase I-Digested Bacterial DNA Enhanced TLR9 Activation by Low Concentrations of ODN-2006 In Ramos Blue Cells

The genomic DNA of six bacterial species: *Micrococcus luteus*, *Bifidobacterium bifidum*, *Escherichia coli*, *Bacteroides dorei*, *Lactobacillus salivarius* and *Campylobacter jejuni*, were selected to represent members of the common human microbiota spanning a wide range of CpG concentration from 14.97% (*Micrococcus* luteus) to 1.44% (*Campylobacter jejuni*). The extracted DNA were digested with DNase I to the target fragment length of less than 15 bp. Complete

69

digestion was confirmed via PAGE gel (**Supplementary Figure S10**), and the digested DNA was added to the final concentration of 25 µg/mL to Ramos Blue cells plated at a density of 5 x 10^4 cells/well, with or without 2.5 µg/mL of ODN-2006.

As shown in **Figure 7 a)** and **b)**, addition of 25 µg/mL of DNase I-digested microbial genomic DNA fragments did not activate TLR9 in Ramos Blue cells, while they co-activated human TLR9 with 2.5 µg/mL of ODN-2006, i.e., additional activation was observed in cells co-treated with ODN-2006 and DNase I-digested bacterial DNA, compare to 2.5 µg/mL of ODN-2006 alone. Co-activation ranged from 7.26% for *Campylobacter jejuni* to 26.34% for *Micrococcus luteus* over treatment with 2.5 µg/mL of ODN-2006 alone (as % of 25 µg/mL of ODN-2006) ($n = 3 \sim 5$, $p < 0.05$, Wilcoxon test with Benjamini-Hochberg correction to 5% FDR). The experiment was repeated twice using a higher Ramos Blue cell density of 2 x 10^5 cells per well and showed similar patterns (**Supplementary Figure S11**). Under the higher plating density, 25 µg/mL of concentration of fragmented *Escherichia coli* DNA co-activated human TLR9 with low concentrations of ODN-2006, at similar intensity as 25 µg/mL TCGTT, while *Micrococcus luteus* gDNA fragment, with its high CpG concentration, was a more potent co-activator (**Supplementary Figure S12**). The results indicated that small DNA fragments of microbial origins, which contains the CpG dimer, could similarly enhance human TLR9 activation by stimulatory ODNs, similar to TCGTT. Small microbial DNA fragments can therefore play a similar role as synthetic sODNs as co-activators of human TLR9.

3.5.2 Small Microbial DNA Fragments as Human TLR9 Co-Activators: Correlation to Genomic CpG Concentration

Correlation between TLR9 co-activation of 2.5 µg/mL ODN-2006 by fragmented bacterial genomic DNA and CpG concentration is shown in **Figure 7 b)**. The correlation was

found to be highly significant ($\rho = 0.603$, $p = 8.72 \times 10^{-4}$). However, the correlation was likely

due to the genomic DNA of *Micrococcus luteus* alone, which possesses the highest CpG

concentration and also co-activated human TLR9 with ODN-2006 more than the small genomic

DNA fragments from other species. Intergenomic comparison indicated significantly higher co-

activation of human TLR9 by *Micrococcus luteus* genomic DNA fragments compared to other

genomes, while the remaining five genomes did not significantly differ from each other (Pairwise

Student's t-Test with Benjamini-Hochberg Correction to 5% FDR).
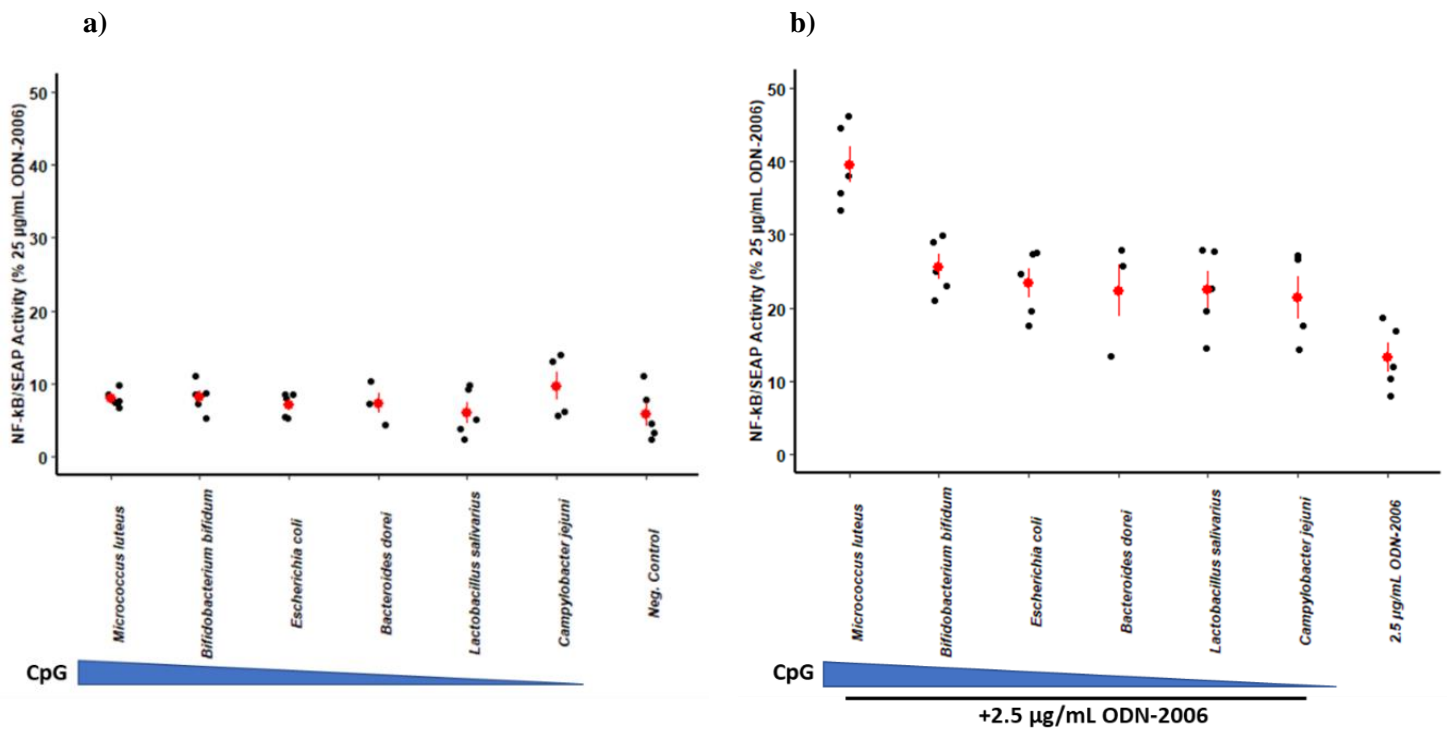
c)
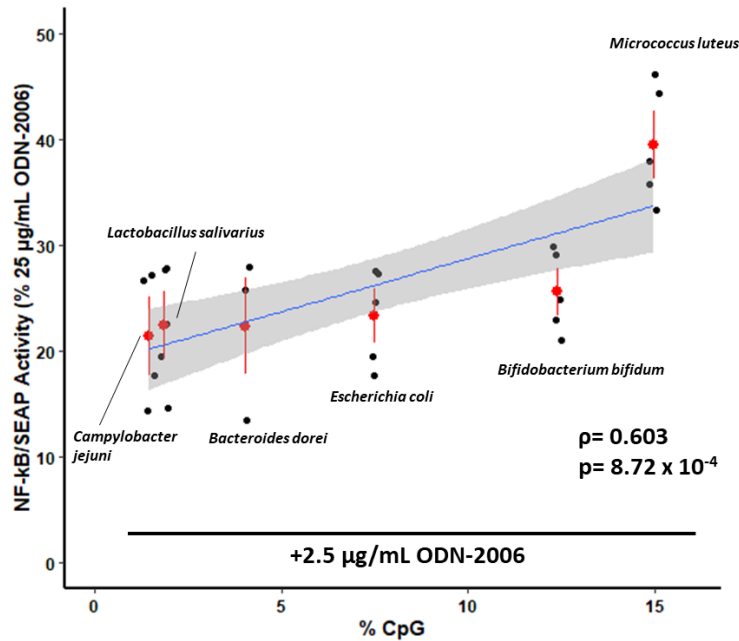
Figure 7. Direct Activation versus Co-Activation of Human TLR9 With Low Concentrations of ODN-2006 by DNase I-Digested Bacterial DNA on Ramos Blue Cells. a) Direct activation of human TLR9: DNase I-digested bacterial DNA fragments (< 15 bp long) were added to the final concentration of 25 µg/mL. **b)** Co-activation of human TLR9 with ODN-2006: DNase I-digested bacterial DNA fragments were added to the final concentration of 25 µg/mL together with 2.5 µg/mL of ODN-2006. **c)** Correlation of TLR9 co-activation with ODN-2006 to bacterial CpG concentration. Stimulation was performed on Ramos Blue cells plated at 5 x 10^4 cells/well in antibiotics-free IMDM medium. TLR9 activation was measured as SEAP activity normalized to plate positive control of 25 µg/mL ODN-2006. Results from at least three and up to five independent experiments, with each dot representing results of one experiment. Bars represented ± s.e.m. Correlation coefficient ρ and significance p were determined via Spearman's rank correlation test.

3.6 TLR9 Activation by Microbial and Human DNA of Different Lengths

Several optimization experiments using *Escherichia coli* and *Campylobacter jejuni* genomic DNA of different fragment lengths on human PBMCs early on alerted us on the influence of DNA fragment length on human TLR9 activation. The outputs of the optimization experiments were summarized in **Results 3.6.1**, and the impact of DNA fragment lengths on TLR9 activation were confirmed on HEK-Dual TLR9 cells (**Results 3.6.2**). Co-activation of

human TLR9 with ODN-2006 and suppression of ODN-2006-induced TLR9 activation were further tested on Ramos Blue cells (**Results 3.7.1-3.7.3).** Therefore, we standardized and specified DNA fragment lengths prior to all TLR9 stimulation experiments.

3.6.1 Activation by *Escherichia coli* and *Campylobacter jejuni* DNA of Different Fragment Lengths on PBMCs

Results from human PBMC stimulation indicated a noticeable influence of fragment length of input bacterial DNA on human TLR9 activation. We used *Escherichia coli* and *Campylobacter jejuni* mechanically fragmented via Covaris g-tube or enzymatically fragmented via NEB Fragmentase to different lengths (**Methods 2.4.1 and 2.4.2**). Fragmented DNA was complexed with DOTAP at the ratio of 1:5 DNA: DOTAP and used to stimulate human PBMCs. TLR9 activation was measured as IFN-β release via ELISA, 24 hours after stimulation. All experiments were performed in duplicates with bars representing ± s.e.m.

Supplementary Figure S13 a) compared DNA extracted from *Escherichia coli*, with or without fragment length standardization to 20,000 bp of length, at input DNA concentrations of 5, 1 and 0.2 μg/mL. For all three concentrations, DNA fragmented to 20,000 bp induced greater IFN-β release from PBMCs compared to unprocessed DNA. **Supplementary Figure S13 b)** compared treatment with *Escherichia coli* and *Campylobacter jejuni* DNA fragmented to 20,000 bp, 6000 bp and 200 bp of length at 1 μg/mL of DNA. Again, differences were observed between DNA fragmented to different lengths for both bacterial species. DNA fragmented to 200 bp induced lower IFN-β release from PBMCs compared to that fragmented to 20,000 bp and 6000 bp, while *Escherichia coli* DNA was more activating than *Campylobacter jejuni* DNA at comparable fragment lengths. **Supplementary Figure S13 c)** compared *Escherichia coli* and *Campylobacter jejuni* fragmented to the lengths of 6000 bp and 1000 bp, added to PBMCs to the

final DNA concentrations of 2.5, 0.5 and 0.1 µg/mL. In this case, no difference was found between the two fragment lengths regardless of species or input DNA concentration. *Escherichia coli* DNA consistently induced greater IFN-β release than that of *Campylobacter jejuni* for all concentrations.

3.6.2 Activation by *Escherichia coli* and Caco-2 DNA of Different Lengths on HEK-Dual TLR9 Cells

Extracted *Escherichia coli* genomic DNA was used to stimulate HEK-Dual TLR9 cells at the concentration of 5 µg/mL with or without fragmentation to 50-200 bp of length. DNA from intact and apoptotic Caco-2 cells, as well as human genomic DNA fragmented to 50-200 bp was used at the concentration of 5 and 20 µg/mL. For both conditions, DOTAP was complexed with DNA at the concentration of 10 µg/mL before stimulation.

No significant difference was observed in TLR9 activation of HEK-Dual TLR9 cells treated with *Escherichia coli* DNA of different lengths (**Supplementary Figure S14**). Extracted *Escherichia coli* DNA, confirmed by gel electrophoresis to be over 10,000 bp long, resulted in comparable activation of TLR9 to DNA fragmented to 50-200 bp, the length previously used in multi-species stimulations (**Results 3.1.3.**).

On the other hand, human DNA fragmented to 50-200 bp long was found to be non-activating, unlike DNA extracted from intact and apoptotic Caco-2 Cells. Intact Caco-2 DNA at the concentration of 5 µg/mL of final input concentration was less activating than *Escherichia coli* DNA (17.7% versus *Escherichia coli*'s 49.3% of 0.5 µg/mL ODN-2006), consistent with prediction from the low genomic CpG concentration of human-origin DNA. 5 µg/mL of apoptotic Caco-2 DNA induced 19.9% of SEAP activity relative to ODN-2006 (refer to **Results 3.8.2** and **Figure 10**).

3.7 TLR9 Activity of Ramos Blue Cells Stimulated by Microbial and Human DNA of Different Lengths, and Activity Changes in Combination with ODN-2006

3.7.1 TLR9 Activation by Microbial and Caco-2 DNA of Different Fragment Lengths on Ramos Blue Cells

Direct activation of TLR9 by *Escherichia coli* and Caco-2 DNA was tested on pre-starved Ramos Blue cells plated at 2 x 10^5 cells/well. TLR9 activity was measured as SEAP relative to plate positive control of 5 µg/mL ODN-2006, 20 h after stimulation. Activation by *Escherichia coli* DNA at 40 µg/mL was observed at the fragment length of 50-200 bp, but not from DNase I-digested fragments at less than 15bp (**Supplementary Figure S15 a) and b)**). No activation was observed for extracted DNA at greater than 10,000 bp long. For Caco-2 DNA, activation was observed at DNA fragmented to 200-60,000 bp and 50-200 bp long, but not from extracted DNA at greater than 10,000 bp long or DNase I-digested DNA fragments at less than 15 bp long, though the signal was rather weak and no statistical analyses could be performed from the low number of experiments conducted (**Supplementary Figure S16**). An unexpected result was noted when comparing DNA at 50-200 bp long: at 40 µg/mL of concentration, Caco-2 DNA was more activating than *Escherichia coli* DNA, inverse to our prediction from their genomic CpG concentration (**Supplementary Figure S15 a)**).

3.7.2 Co-Activation of Human TLR9 with ODN-2006 by Microbial and Caco-2 DNA Fragments of Different Lengths

DNA from *Escherichia coli* and Caco-2 were fragmented to either 50-200 bp or less than 15 bp long, then used to stimulate pre-starved Ramos Blue cells at 40 µg/mL final concentration in combination with 0, 1 and 2 µg/mL of ODN-2006. **Supplementary Figure S15 a) and b)**

detailed co-activation of TLR9 with ODN-2006 of two fragment lengths for both types of DNA, respectively. Co-activation was weak for both genomes at fragment lengths of less than 15 bp, though *Escherichia coli* DNA was slightly more co-activating of TLR9 with low concentrations of ODN-2006 compared to Caco-2 DNA. Greater co-activation was also found for DNA fragmented to 50-200 bp long regardless of origin compared to those less 15 bp long. Contrary to expectation, it was Caco-2 DNA that co-activated human TLR9 with 1 µg/mL of ODN-2006 more, compared to *Escherichia coli* DNA. No co-activation was observed on top of direct activation by 2 µg/mL of ODN-2006, suggesting that a critical concentration of ODN-2006 was needed for co-activation to occur.

3.7.3 Long Caco-2 DNA Fragments Repressed Human TLR9 Activation by ODN-2006 and TCGTT

Caco-2 DNA of four different fragment lengths, greater than 10,000 bp, 200-6000 bp, 50-200 bp and less than 15 bp, were used at 40 µg/mL of final concentration on top of 5 µg/mL of ODN-2006, or 0.5 µg/mL of ODN-2006 plus 100 µg/mL of TCGTT on pre-starved Ramos Blue cells. The concentrations of ODN-2006 and TCGTT used were representative of optimal concentration for TLR9 activation by ODN-2006 or TCGTT co-activation as determined in calibration experiments.

Results showed that Caco-2 DNA at greater than 10,000 bp of length reduced TLR9 activation by ODN-2006: TLR9 activation dropped to 45% of 5 µg/mL ODN-2006 when 40 µg/mL of DNA was added (n = 4, $p < 0.05$, Student's T-Test) (**Figure 8**). Repression lessened with the three shorter fragment lengths, where no significance was found. Results from two experiments also showed Caco-2 DNA of greater than 10,000 bp long to reduce TLR9 activity of cells treated with 0.5 µg/mL ODN-2006 plus 100 µg/mL of TCGTT (**Supplementary Figure**

S17), but no statistical analyses could be performed due to the small number of independent experiments performed.
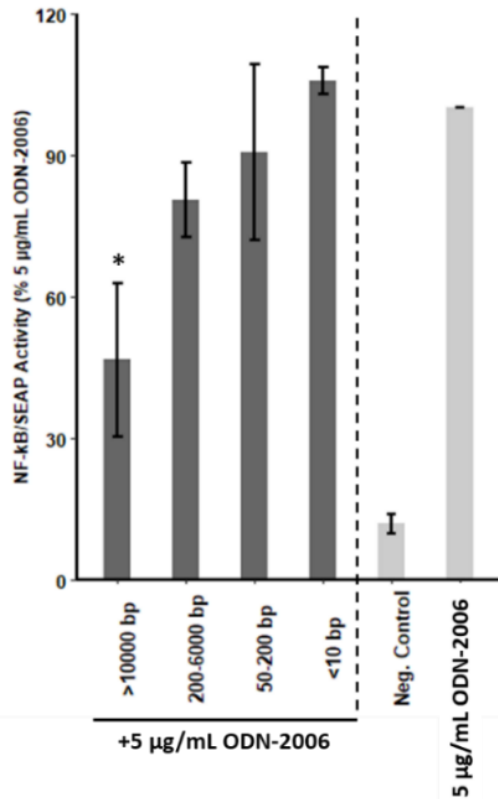
Figure 8. Long Human DNA Fragments Repressed TLR9 Activation by ODN-2006 on Ramos Blue cells. 40 µg/mL (final concentration) of Caco-2 DNA sheared to different fragment lengths were added on top of 5 µg/mL ODN-2006, compared to 5 µg/mL ODN-2006 alone. Results showed long (>10,000 bp) Caco-2 DNA fragments significantly reduced TLR9 activation in Ramos Blue cells. Ramos Blue cells overnight-starved and plated at 2×10^5 cells/well were used. TLR9 activation was measured as SEAP activity normalized to plate positive control of 5 µg/mL ODN-2006. Results from four independent experiments. Bars represented ± s.e.m. (*: $p < 0.05$, Student's t-test)

3.8. Human DNA at 50-200 Bp of Length Co-Activated Human TLR9 With Low Concentrations of ODN-2006 and Select Bacterial DNA on HEK-Dual TLR9 Cells

3.8.1 Observed Co-activation of Human TLR9 by Caco-2 DNA Fragments With ODN-2006 and *Escherichia coli* Genomic DNA

Caco-2 DNA was enzymatically fragmented to the length of 50-200 bp and used to co-activate human TLR9 together with 0.05 µg/mL ODN-2006 or 2 µg/mL DNA of three different bacterial species of contrasting CpG concentration: *Bifidobacterium bifidum, Escherichia coli* and *Fusobacterium nucleatum*, on HEK-Dual TLR9 cells. Human DNA was added at the concentration of 20 µg/mL, and human DNA alone was used in all experiments as an internal

negative control. A standard concentration of 10 µg/mL DOTAP was used to pre-complex with the DNA mixture for all conditions, identical to the multispecies experiment in **Results 3.1.3**. Results were measured as SEAP compared to the positive control of 0.5 µg/mL ODN-2006.

As shown in **Figure 9 b)**, human DNA fragments were powerful co-activators of human TLR9 with low concentrations of ODN-2006, and 20 µg/mL human DNA on top of 0.05 µg/mL ODN-2006 resulted in SEAP signals on par with 0.5 µg/mL ODN-2006, and the results were highly significant ($p < 0.01$, Wilcoxon test with Benjamini-Hochberg correction to 5% FDR). On the other hand, no direct activation was observed when 20 µg/mL human DNA was used on its own (**Figure 9 a)**).

Human DNA could also co-activate human TLR9 with low concentrations of DNA from *Escherichia coli*, but not that from *Bifidobacterium bifidum* or *Fusobacterium nucleatum* ($p < 0.05$, Wilcoxon test with Benjamini-Hochberg correction to 5% FDR). Co-activation with *Escherichia coli* DNA was weaker than that with ODN-2006, at 10.5% of 0.5 µg/mL ODN-2006 over *Escherichia coli* DNA alone. Human DNA did not induce additional TLR9 activation with *Bifidobacterium bifidum* DNA, which was activating on its own at the concentration of 2 µg/mL with SEAP activity at 49.0% of 0.5 µg/mL ODN-2006. No co-activation was observed with the addition of human DNA on top of *Fusobacterium nucleatum* DNA, which was also non-activating on its own (**Figure 9 c) to e)**). The results indicated that there could be an optimal baseline TLR9 activation for co-activation with high concentrations of human DNA to occur, which was achievable with 0.05 µg/mL ODN-2006 or 2 µg/mL *Escherichia coli* DNA.

a) Caco-2 DNA Only
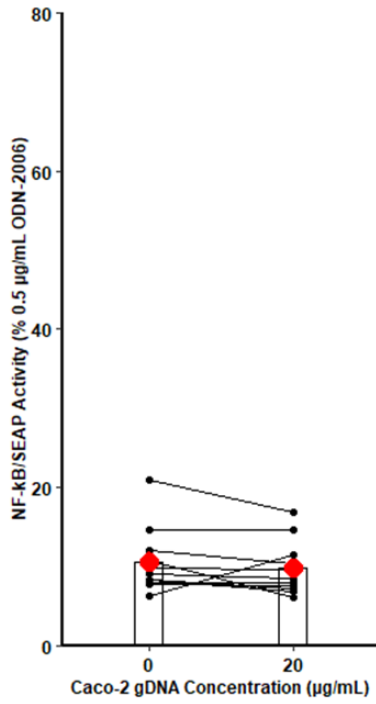
b) +0.05 µg/mL ODN-2006
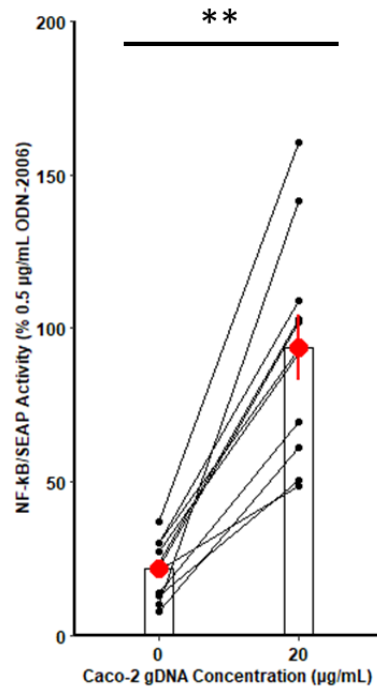
c) +2 µg/mL *B. bifidum* DNA
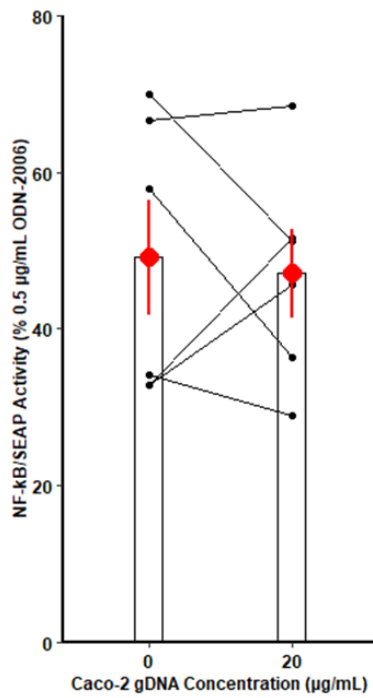
d) +2 µg/mL *E. coli* DNA
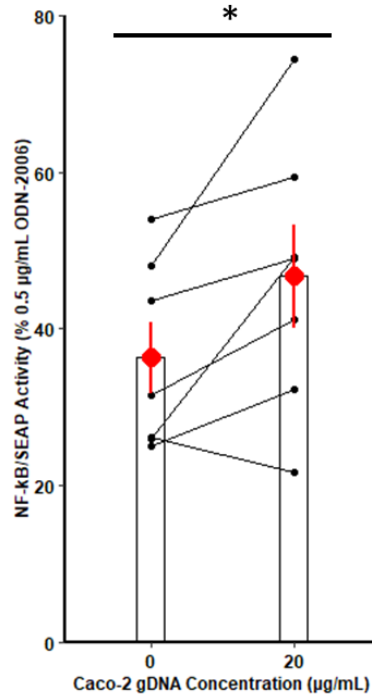
e) +2 µg/mL *F. nucleatum* DNA
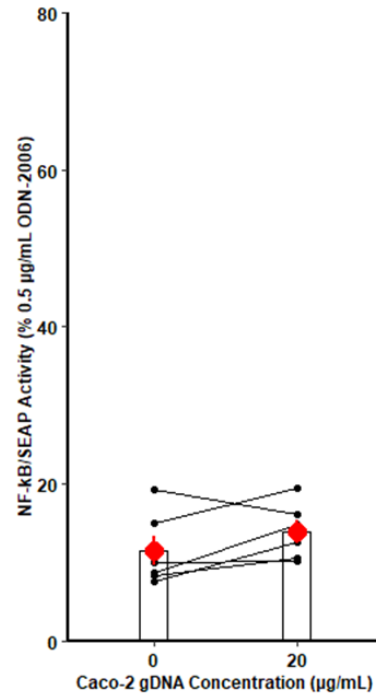
Figure 9. Human DNA Fragments at 50-200 Bp Co-Activated TLR9 with Low Concentrations of ODN-2006 and *Escherichia coli* Genomic DNA on HEK-Dual TLR9 Cells. Caco-2 DNA fragmented to 50-200 bp of length were used to co-activate different types of DNA. 20 µg/mL Caco-2 DNA fragments were added to **a)** directly activate TLR9, or co-activate **b)** 0.05 µg/mL of ODN-2006, **c)** 2 µg/mL of *Bifidobacterium bifidum* DNA, **d)** 2 µg/mL of *Escherichia coli* DNA, or **e)** 2 µg/mL of *Fusobacterium nucleatum* DNA. A comparison of 0 vs. 20 µg/mL of Caco-2 DNA was shown for each condition. Similar to Caco-2 DNA, all bacterial DNAs were fragmented to 50-200 bp using NEB dsDNA Fragmentase. A standardized DOTAP concentration of 10 µg/mL was used for all wells including controls. TLR9 activation was measured as SEAP activity normalized to plate positive control of 0.5 µg/mL ODN-2006. Results from at least 6 and up to 11 independent experiments are shown, with each dot and connecting line representing results of one experiment. Bars represented ± s.e.m. (*: p<0.05, **: p<0.01, Wilcoxon test with Benjamini-Hochberg correction to 5% FDR)

3.8.2 Length of Human DNA Fragments Determined Whether Direct TLR9 Activation or Co-Activation With ODN-2006 Occurred

DNA was extracted from intact and apoptotic Caco-2 cells, with the fragment lengths at respectively >10,000 bp and 200-6000 bp in a typical step ladder pattern of multiples of 180 bp. Caco-2 DNA was also enzymatically fragmented to 50-200 bp. The three types of DNA: intact, apoptotic, or enzymatically fragmented, were added with or without 0.05 µg/mL of ODN-2006 to HEK-Dual TLR9 cells at the concentrations of 5 and 20 µg/mL with a standard DOTAP concentration of 10 µg/mL. TLR9 activation was measured as NF-kB/SEAP Activity compared to the positive control of 0.5 µg/mL of ODN-2006.

Figure 10 depicts the effect of fragment length of human DNA on TLR9 activation. To summarize, Caco-2 DNA extracted from intact or apoptotic Caco-2 cells directly activated TLR9 in HEK-Dual TLR9 cells. No co-activation was observed with the addition of low concentrations of ODN-2006. On the other hand, Caco-2 DNA fragmented to 50-200 bp was completely non-activating of human TLR9 on its own, but significantly co-activation occurred when low concentrations of ODN-2006 was added. Results suggest that the fragment length of a given DNA ligand can have drastically different effects on the human TLR9 receptor.

a) Intact Caco-2 DNA
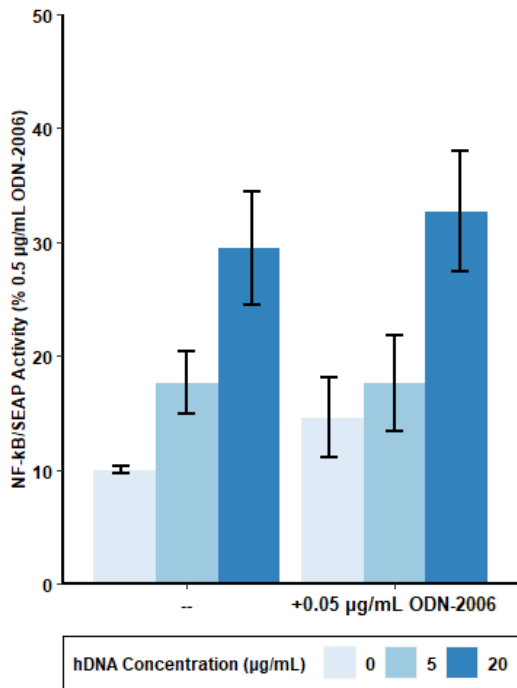
b) Apoptotic Caco-2 DNA
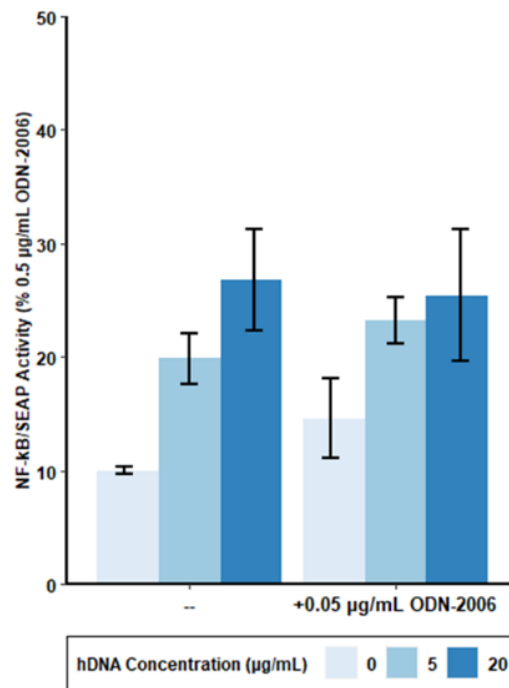
c) Caco-2 DNA (50-200 bp)
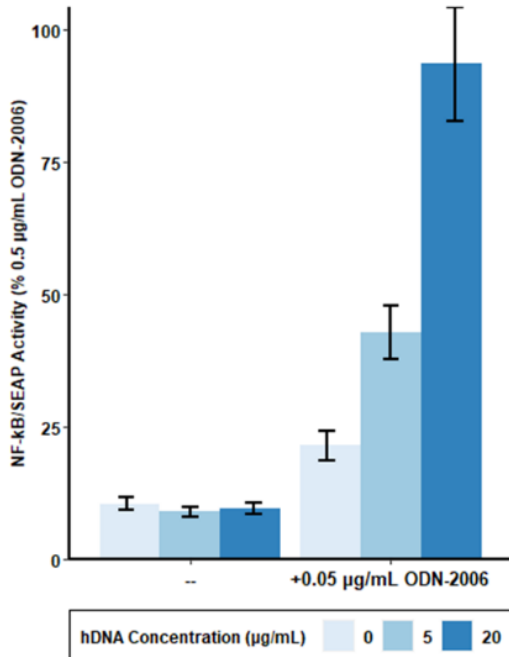
Figure 10. Length of Human DNA Fragments Determined Whether They Directly Activated human TLR9 or Indirectly Co-Activated TLR9 With Low Concentrations of ODN-2006 on HEK-Dual TLR9 Cells. DNA extracted from **a)** intact Caco-2 cells and **b)** apoptotic Caco-2 cells directly activated TLR9. No difference was seen when 0.05 µg/mL ODN-

2006 were added. **c)** Caco-2 DNA fragmented to 50-200 bp long did not directly activate TLR9, but strongly co-activated TLR9 when 0.05 µg/mL of ODN-2006 was added. TLR9 activation was measured as SEAP activity normalized to plate positive control of 0.5 µg/mL ODN-2006. A standardized DOTAP concentration of 10 µg/mL was used for all wells including positive and negative control. Results from at least 3 and up to 11 independent experiments with bars representing ±s.e.m. (*p<0.05, **p<0.01, Student's T-Test with Benjamini-Hochberg correction to <5% FDR)

3.9 Effect of TLR9 Activation by ODN-2006 and Human DNA on *Fusobacterium nucleatum* Invasion of HT-29 Cells

3.9.1 Increased Presence of Intracellular *Fusobacterium nucleatum* upon TLR9 Activation by ODN-2006 and Caco-2 DNA on HT-29 Cells

Results from two independent invasion experiments using confluent HT-29 cells and *Fusobacterium nucleatum* confirmed the invasive ability of *Fusobacterium nucleatum* (**Figure 11a**)) in line with literature[113,114]. The negative control *Escherichia coli* DH5α was non-invasive, with fewer than 0.001% of bacteria surviving antibiotics treatment. *Fusobacterium nucleatum* invasion increased when cells were simultaneously treated with 5 µg/mL of ODN-2006, 20 µg/mL of Caco-2 DNA at 50-200 bp of fragment length, or both (**Figure 11 a**)), compared to wells without DNA treatment, in both experiments. However, the percent of invasive *Fusobacterium nucleatum* with or without ODN-2006 and Caco-2 DNA differed considerably by experiment, resulting in large, overlapping error bars. The growth rate of *Fusobacterium nucleatum* was found to be different between different batches, and could contribute to the observed variation between experiments. Moreover, no statistical analyses were possible due to the small number of biological replicates (n = 2). Therefore, additional experimental repeats are needed to confirm that TLR9 activation consistently promotes *Fusobacterium nucleatum* invasion.

3.9.2 IL-8 Release upon Addition ODN-2006, Caco-2 DNA and Bacteria

Human IL-8 release was measured via ELISA on filtered supernatant frozen for a maximum of 4 days. Wells treated with ODN-2006, Caco-2 DNA, and a combination ODN-2006 with Caco-2 DNA with or without *Fusobacterium nucleatum* were compared, together with three control conditions (wells without bacteria, wells with *Escherichia coli* DH5α, and wells with *Fusobacterium nucleatum* without addition of DNA). An additional experiment was performed on HT-29 cells treated with ODN-2006 and/or Caco-2 DNA alone for 4 hours, identical to the "no bacteria" conditions in invasion assays.

Increase in IL-8 production occurred upon addition of bacteria, with the addition of *Escherichia coli* leading to higher IL-8 production compared to that of *Fusobacterium nucleatum* (**Figure 11 b**)). However, no statistical analyses were performed due to the small number of biological replicates (n=2). Addition of ODN-2006 and Caco-2 DNA to cells infected with *Fusobacterium nucleatum* did not further increase IL-8 production. As *Escherichia coli* DH5α was not invasive, the increase of IL-8 production upon *Escherichia coli* DH5α treatment was likely due not to TLR9 activation but to recognition by surface receptors to other components of the bacteria. There was also no sign that increased *Fusobacterium nucleatum* invasion upon ODN-2006 and/or Caco-2 DNA treatment (**Results 3.9.1**) lead to additional IL-8 release from HT-29 cells. Therefore, our experiments could not confirm TLR9 activation or increased *Fusobacterium nucleatum* presence within the cell as sources of inflammatory cytokine release from HT-29 cells.

a)
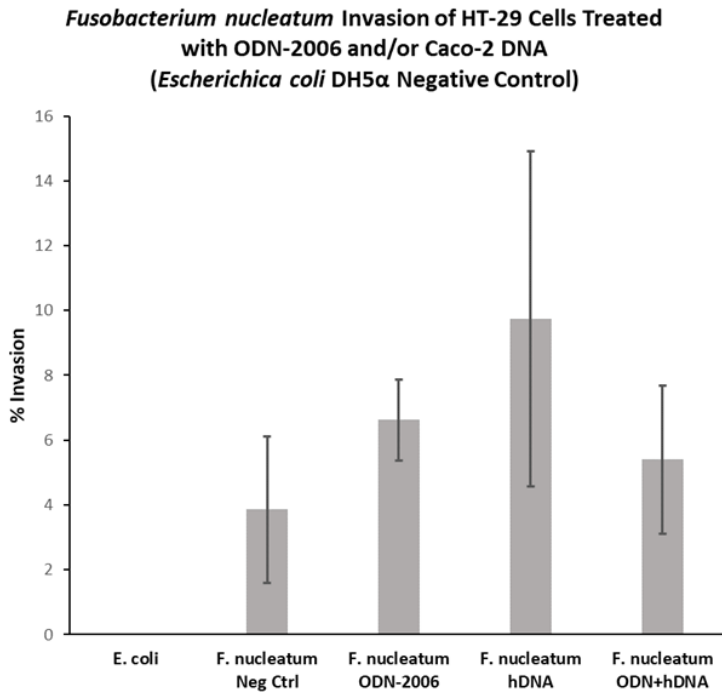
b)
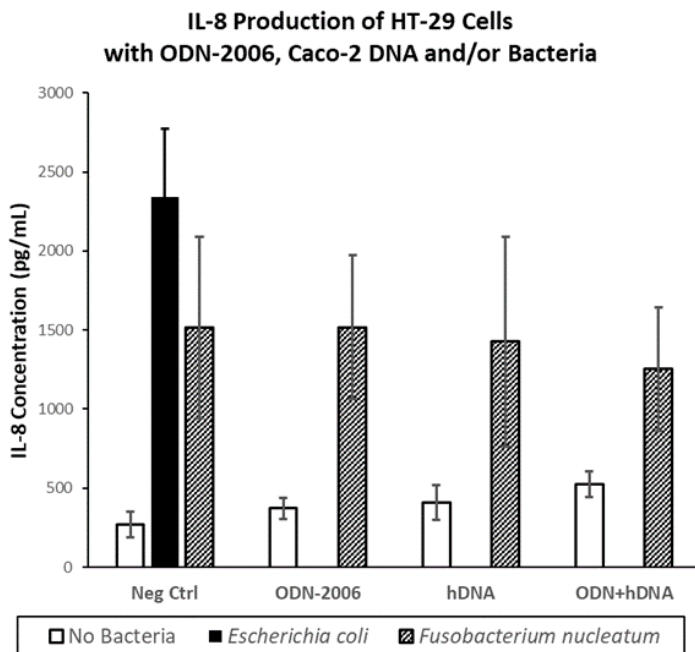
Figure 11. Invasion Experiments To Determine Role of TLR9 Activation by ODN-2006 and human DNA on *Fusobacterium nucleatum* Invasion and Human IL-8 Production.

a) Intracellular *Fusobacterium nucleatum* with Treatment of ODN-2006, Caco-2 DNA, or Both. Confluent HT-29 cells were treated with ODN-2006, human DNA (extracted from Caco-2 cells and enzymatically fragmented to 50-200 bp), or both. *Fusobacterium nucleatum* was added at the Multiplicity of Infection (MOI) of 5:1 bacteria to HT-29 cells. Controls consisted of the same number of *Escherichia coli* DH5α or *Fusobacterium nucleatum* without additional DNA. % Intracellular bacteria were determined as number of colony-forming units after 1 hour of antibiotics treatment and cell lysis with Saponin, divided by the number of input bacteria (results from two independent experiment, bars ± s.e.m.).

b) IL-8 Production by HT-29 Cells Treated with ODN-2006, Human DNA, or Both, With or Without Addition of Bacteria. IL-8 production by HT-29 cells were measured via ELISA from supernatant collected from invasion experiments, as well as that from cells treated with ODN and/or human DNA but without addition of bacteria. Noninvasive *Escherichia coli* DH5α was added only in control wells without DNA treatment. Increase of IL-8 release into the supernatant was found with the addition of both *Escherichia coli* and *Fusobacterium nucleatum*, though no additional statistical analyses were performed due to the low number of experiments conducted (results from two independent experiments with bars representing ± s.e.m.).

4. Discussion

The human microbiota population is vast and outnumbers human somatic cells at the ratio of circa 1.3:1[101]. Most of the microbiota reside in the human large intestine and provide a rich source of non-human DNA. DNA of diverse origins, with different concentrations of CpG, k-mer and other immunostimulatory sequences, are thus highly represented in the intestinal lumen.

Typically, a thick and impermeable mucosal barrier prevents the microbiota from direct contact with the gut epithelium, and the antimicrobial peptides at the bottom layer of the intestinal mucus further destroy any bacteria that manage to penetrate through the upper layer[82]. The bacteria are moved, along with shed gut cells, down the digestive tract, together with the gel matrix of the lose upper mucus layer. When the mucosal layer is intact, neither the intestinal epithelial cells, nor the immune cells present in the intestinal epithelium, are perturbed by the presence of normal human microbiota population in the luminal region[125]. The cells of the human intestinal epithelium, similar to those of other human internal tissues, are largely devoid of bacteria presence and remain sterile[82].

However, situations exist where human cells containing TLR9 can contact the microbiota population and their genomic DNA and thus be activated. Examples include intestinal pathogens that penetrate through the mucosal region and invade intestinal epithelial cells[126], and certain specialized cell types that actively take in and sample bacteria, such as the M cells of Peyer's patch in the small intestine[127,128]. Dendritic cells, which are present in Payer's patch, mesenteric lymph nodes, and throughout the lamina propria layer of the small intestine and colon, similarly sample the intestinal lumen in cooperation with intestinal epithelial cells[129]. Moreover, in

pathological conditions such as IBD and infectious diseases such as *Clostridioides difficile* colitis, the intestinal barrier can be compromised[92,117]. In these situations, the local immune cell populations with PRRs that reside in the lamina propria layer of the intestine can recognize bacterial breach of the intestinal barrier, and mount an innate and adaptive immune response against bacterial intruders, which may include TLR9 responses to microbial DNA[94].

4.1 Toward an *In Vitro* Model of TLR9 Activation Within the Human Gut Environment and Their Applications

While our *in vitro* assays did not fully characterize the complex interaction of microbial DNA and human DNA with human TLR9, we were able to identify trends regarding human TLR9 activation and co-activation relative to CpG concentrations of human microbiota members, DNA fragment lengths, and the presence of human DNA in the gut environment. Our results provide insights into the possible role of TLR9 activation for infant gut development, inflammatory bowel diseases, and bacterial invasion.

4.1.1 TLR9 Activation and Infant Gut Development

A number of studies described microbiota changes in infants from birth to circa 1 year of life[130]. Infants undergoing vaginal birth are likely seeded with maternal stool microbiota, as well as facultative anaerobes from the surrounding environment and the maternal skin[131]. In contrast, infants undergoing Caesarian section birth, who had less direct contact with the maternal stool microbiota, are colonized predominantly with environmental and skin microbiota at birth[131, 132]. Regardless of birth mode, the microbiota of the infant shifts within the first two weeks of birth to

one dominated by *Bifidobacterium* spp., species not only found in maternal stool, but also in breast milk, which is the likely source of the species in infants born by C-section[133].

Bifidobacterium spp., which exist as the majority population in the infant gut but gradually decrease as the infant's gut microbiota evolves, are noted for their higher CpG concentration compared to typical species found in adult guts[52]. *Bifidobacterium bifidum* (12.40% CpG) and *Bifidobacterium breve* (9.93% CpG)[46] were two of the species with the highest CpG concentration used in our multispecies stimulation experiment (**Results 3.1.3**).

The adult microbiota, in contrast, is typically dominated by *Bacteroidetes* and *Firmicutes*[134]. The *Bacteroidetes* species we used in the multispecies experiment, *Bacteroides dorei*, is a typical adult gut microbiota member[135] with a CpG concentration of 4.04%[46], far lower than the two *Bifidobacterium* species tested. And as predicted, **Results 3.1.3** revealed higher TLR9 activation upon exposure to genomic DNA of *Bifidobacterium spp.* compared to that of *Bacteroides dorei*. Therefore, we expected the *Bifidobacterium spp.*-rich environment of the infant gut, where the key residents are high in CpG as predicted by bioinformatics[136], to induce higher local TLR9 activation compared to the adult gut environment. TLR9 activation from CpG-rich microbiota induces IFN-β production from pDCs, which converts naïve CD4 T cells into Th1 cells[137] and suppresses Th2 cytokine expression, provides the needed balance against the Th2-biased T-cell population of newborns[138]. As Th2 overrepresentation is a known source of allergy[139], high TLR9 activation within the infant gut environment can have an impact on immune response to pathogens and autoimmune issues such as allergies later in life[140], thus play a role in early development of the infant innate immune system.

4.1.2 Human DNA in the Gut Environment, and Its Role in Activation and Co-Activation Of Human TLR9

A major characteristic of human intestinal epithelial cells is their rapid turnover. Typically, intestinal epithelial cells are completely replaced in 2 to 6 days, and for humans, approximately 10^{11} intestinal cells are shed by day[85]. The replaced cells are shed into the lumen and soon die of apoptosis[87]. The apoptotic intestinal epithelial cells, along with necrotic cells, are major contributors of human DNA in the mucosal layer of the gut[83]. According to our calculations[(Note 2)], if apoptotic DNA from shed intestinal epithelial cells is evenly distributed throughout the intestinal tract, each cm^3 within the intestinal tract contains 138 µg of human DNA from apoptotic intestinal epithelial cells, a concentration high enough to either directly activate or co-activate human TLR9 with microbial DNA (**Results 3.8.1** and **3.8.2**). Local higher concentration of human DNA is expected to occur near the surface of the intestinal lumen, potentially generating a more intense TLR9 response.

Apoptotic DNA is fragmented and consists of a ladder of different DNA lengths which could potentially be directly stimulatory or co-stimulatory of human TLR9 depending on their actual length (**Results 3.8.2**), and can also interact with microbial DNA in multiple ways depending on the CpG concentration and fragment lengths of the microbial DNA. In addition, apoptotic DNA, which are present within apoptotic bodies and blebs, can be absorbed by cells containing the TLR9 receptor such as human macrophages[141,142] and pDCs[143].

In pathologies such as IBD, intestinal epithelial cells are replaced at a higher rate, implying greater apoptosis and thus higher amount of apoptotic DNA[142,144]. Cells can also die of necrosis, a pathological version of cell death that involves cell bursting and spewing of intracellular materials, necrotic DNA included, into the surrounding environment[91]. While

apoptotic DNA shows an ordered "ladder" pattern consisting of fragment lengths of multiples of 180 bp[145], necrotic DNA is not fragmented and only degrades upon exposure to DNase within the gut environment[146]. The higher concentration of human DNA upon increased cell apoptosis and necrosis in gut pathologies can co-present with a compromised gut barrier, increasing contact of TLR9-containing immune cells with DNA from dead cells, resulting in potential autoimmune reactions[94]. All of this is consistent with our findings in **Results 3.8.1** that even with a small amount of microbial DNA, self-DNA from shed intestinal epithelial cells can result in human TLR9 activation when the cells are not efficiently cleared and removed, instead accumulate in the intestinal environment.

Note 2: The amount of human DNA in the gut can be calculated as such: a diploid human cell contains approximately 6.4×10^{-12} g of DNA[147], thus 10^{11} shed intestinal epithelial cells yield 0.64 g of DNA per day. The volume of the human intestine could be calculated as such[148]: 1.25cm (radius of small intestine) * 1.25cm * 3.14 * 333cm (length of small intestine) = 1633 cm^3 (volume of small intestine) and 2.4cm (radius of large intestine) * 2.4cm * 3.14 * 167cm (length of large intestine) = 3020 cm^3 (volume of large intestine), which together adds to 4654 cm^3 of volume. The amount of human DNA per cm^3 of intestinal tract produced per day would be 0.64g / 4654 cm^3 = 1.38×10^{-4} g or 138 µg per cm^3 of human intestinal tract.

4.1.3 TLR9 Activity Can Affect Pathogen Invasion and Survival Inside Host Cells

Preliminary results from our invasion experiments of *Fusobacterium nucleatum* of HT-29 cells, with or without the inclusion of ODN-2006 and human DNA, suggest that TLR9 activation could affect bacterial invasion. We observed increased invasion from *Fusobacterium nucleatum*, a pathogen of very low CpG concentration (0.296%), with the addition of ODN-2006 and/or Caco-2 DNA (**Results 3.9.1**), though the result was yet to be statistically validated.

However, increased intake of bacteria upon TLR9 activation had been previously shown using mouse macrophages and *Escherichia coli*[149]. Macrophages stimulated with ODN-2006 internalized more *Escherichia coli* compared to those without[149], suggesting that TLR9 activation has a positive effect on macrophage phagocytosis. In the same study, several genes

involved in phagocytosis and apoptotic cell clearance were found to be upregulated in TLR9-activated macrophages[149]. Therefore, it is possible that TLR9 activation induces cells to proactively take in pathogens in an attempt to destroy them, leading to increased bacterial presence inside the cells.

     Fusobacterium nucleatum is a species associated with colorectal cancer[102], and if TLR9 activation by ODNs or human DNA increases their presence inside intestinal epithelial cells, it hints at a possible oncogenic role of TLR9, which is upregulated in several tumor cell lineages including gastric cancer, breast cancer, prostate cancer, and lung cancer cells[150]. Human carcinoma cell lines MCF-7, A549, and Caco-2 exhibited greater motility, indicating pro-metastasis phenotypes, upon stimulation of their TLR9 via ODNs[151] or bacterial DNA fragments[152]. *Fusobacterium nucleatum* invasion also increased cell motility and permeability through the Transwell membrane in co-cultivation experiments, likely due to increased production of inflammatory cytokines IL-8 and CXCL1 by invaded cells[153], identifying the bacterium's role in oncogenesis and metastasis. Thus, increased TLR9 activity of the gut could potentially escalate colorectal cancer development, and be used as a possible diagnostic marker for cancer. When the gut barrier is compromised, and frequency of contact between the microbiota and TLR9-expressing cells is increased, TLR9 activation in the gut can be a possible link between diseases such as IBD and a greater risk of cancer development.

4.2 Potential Confounders of TLR9 Activation Experiments

4.2.1 Fragment Length Needs to be Standardized for Experiments, Posing Technical Limitations to Metagenomic DNA Stimulation Experiments

Our experiments showed major influence of DNA fragment length for both microbial and human DNA on TLR9 activation (**Results 3.6.1, 3.6.2 and 3.8.2**), TLR9 co-activation with ODN-2006 (**Results 3.7.2, Results 3.8.2**), or repression of TLR9 activation by ODN-2006 (**Results 3.7.3**). Additionally, highly fragmented bacterial DNA at 15 bp or less of length co-activated TLR9 with low concentrations of ODN-2006 without being activating on their own (**Results 3.5.1**).

Therefore, to obtain consistent multi-species TLR9 activation experiments and correlate TLR9 activation with genomic CpG or k-mer concentrations, DNA fragments of similar length are needed. Moreover, short DNA fragments must be removed to preclude co-activation. Standardized fragment lengths also ensure cells are exposed to similar numbers of DNA fragments at a given DNA concentration.

One major difficulty in our attempts to standardize fragment length of microbial DNA were the large differences of fragmentation times from one genome to the next. Depending on the genome, 30 to 120 min were needed to obtain DNA fragments of 50-200 bp using the NEB dsDNA fragmentation enzyme (**Results 3.1.2**). The reason behind the differences was unknown, but speculated to be the presence of enzyme inhibitors in extracted DNA. No correlation between fragmentation time and genomic G+C or CpG concentration was found.

By extension, the varied fragmentation times between bacterial species would be a major limitation if NEB's dsDNA Fragmentase was used to fragment metagenomic DNA. If the DNA

of one species fragmented more quickly than another, the DNA fragments of the fast-fragmenting species would be shorter at any point of time compared to those of the slow-fragmenting species, leading to inconsistent fragment lengths between species. The use of clean-up kits at later steps could also filter out the fast-fragmenting DNA and reduce their representation within a sample, and skew the overall species contribution of the metagenomic DNA used for stimulation.

The Covaris S220 sonication system, which used mechanical fragmentation, could standardize DNA fragments more effectively, though even for this system, differences between genomes existed. **Supplementary Figure S4** showed that while all genomes were standardized to 3000 bp after 600 s of sonication, DNA from *Escherichia coli* smeared across a larger range of fragment lengths compared to DNA from the other four genomes which were more uniform. In addition, Covaris microTUBES would be needed to mechanically shear DNA to shorter lengths, and they are expensive and have low capacity compared to enzymatic fragmentation.

In a parallel study of infant microbiota with samples collected from newborn and 3-week-old infants, as well as adult stool from mothers, we found that the DNA extracted from human stool presented in various levels of intactness, from greater than 10kb to less than 100 bp of length (**Supplementary Figure S18**). While storage and delivery conditions were possible reasons behind the varied intactness of extracted DNA, technical issues were likely not the only contributor to the differences. Other reasons for fragment length inconsistency of extracted DNA could include whether the bacteria were alive or dead at time of extraction, whether the stool was firm or loose, or the condition of the gut microenvironment at time of sample collection. The use of antibiotics, for instance, is likely to cause massive death of the gut microbiota population and a more fragmented microbial DNA profile[154]. DNA fragments of different lengths could also co-

exist within a microbiota sample, and potentially some species of bacteria could have genomic DNA more fragmentable than others, similar to how varied fragmentation time was needed to achieve target lengths via the NEB dsDNA Fragmentase enzyme for the various species. It is also possible that DNA from different microbial species presents in different fragment lengths within the gut environment, thus contributes to human TLR9 activation in the gut differently, with some playing a greater role than others independent of genomic CpG concentration. Additionally, plasmid DNA could be expressed in different lengths compared to chromosomal DNA from bacteria[155], potentially complicating the overall sequence contribution to TLR9 activation in the gut.

4.2.2 Actual DNA Fragment Length in Contact with Human TLR9

Human TLR9 is located in the endosomal region, which is separate from the surface of the cell. Microbial DNA fragments, which are enclosed inside the cell wall of a bacterium or the protein coat of a virus, must be introduced to the endosome through phagocytosis of the microbial cell or viral unit, active invasion of the microbe, or as free DNA released upon the death of the microbe taken in by the human cells. In all cases, microbial DNA must travel from the surface of the cell to the endosome to come into contact with human TLR9[156].

Through FITC labeling of calf thymus genomic DNA and PAGE gel electrophoresis, Pohar et al. showed that the DNA within the endosome of the cell was extremely fragmented, at an average length of less than 20 nucleotides. No DNA fragments of over 15bp could be detected, suggesting that DNA was degraded prior to arrival in the endosome, and presented as short fragments, most of which lower than the fragment length threshold of 10bp for direct human TLR9 activation to occur[63]. Our experiments using synthetic short ODN TCGTT and

DNase I digested microbial DNA fragments was able to confirm the lack of direct human TLR9 activation from small DNA fragments of 15bp or less (**Results 3.5.1**). Pohar et al. also showed that serum, a prominent component of blood and a growth additive in culture medium, was able to degrade DNA[157]. Serum degradation was a reason behind the adoption of overnight cell starvation in several of our stimulation and co-stimulation experiments with Ramos Blue cells. However, the use of serum-free medium posed a question of how realistic a situation was regarding TLR9 activation in blood. While cell-free DNA fragments circulate in blood, their low concentration (median at 7.9 ng/mL for normal subjects and 13.8 ng/mL for patients with colorectal cancer[158]) implied that they are probably insufficient to activate human TLR9.

4.2.3 Co-Presence of DNA of Microbial, Human and Synthetic Origins Made TLR9 Activation Patterns Rather Complex

In our study, we combined DNA of different types—ODN, sODN, and DNA of bacterial and human origin and fragment lengths—to identify the varied dynamics of TLR9 activation by complex DNA combinations that can present in the human gut environment. High concentrations of small DNA fragments, for instance, could result in a characteristic co-activation dynamic, where small DNA fragments are in themselves non-activating but co-activate TLR9 with other longer, though less concentrated DNA fragments, analogous to the way short ODNs co-activated TLR9 with ODN-2006 and human DNA.

Co-activation by small DNA fragments negated the specificity of the CpG dimer within a given DNA sequence as the activator of human TLR9. For example, **Results 3.4.1** showed TCGTT not only co-activated ODN-2006, but also ODN-2006GC, the negative control ODN to ODN-2006 with swapped Cytosine and Guanine bases. The intensity by which TCGTT co-

activated TLR9 with ODN-2006GC was as high as that with ODN-2006, suggesting that presence of high amounts of small CpG-containing DNA fragments could relax the sequence specificity of larger DNA fragments co-present and subsequently induce powerful human TLR9 responses. Pohar et al. also showed TCGTT to co-activate with methylated DNA fragments, thus negating the effect of methylation in vertebrate genomic DNA to avoid triggering TLR9[63]. Co-activation could thus drastically alter TLR9 response to microbial versus self-DNA and be a possible cause for TLR9-related autoimmunity. High presence of undegraded human DNA fragments can result from cells undergoing apoptosis or necrosis which are not efficiently cleared, as in the case of SLE[159], as well as NET (neutrophil extracellular trap) consisting a mixture of human and microbial DNA complexed to proteins[77,160], which can co-present with degraded microbial DNA fragments in the endosome. Co-activation of small CpG-containing DNA fragment of microbial origin together with larger DNA fragments with low CpG concentration characteristic of human DNA could therefore occur and induce autoimmune reactions, as modeled in **Results 3.4.2** via TCGTT and DNA fragments from apoptotic human cells.

We also identified the ability of human DNA of certain fragment lengths (50-200 bp, analogous to the smallest DNA fragment from apoptotic ladder at 180 bp) to co-activate TLR9 with ODN-2006 and *Escherichia coli* DNA in HEK-Dual TLR9 cells (**Results 3.8.1**). The result suggested that when large amounts of apoptotic DNA are present together with bacterial DNA, co-activation occurred, which could be relevant for the human gut where massive numbers of shed intestinal epithelial cells co-exist with smaller numbers of bacteria i.e., based on our findings, when an ideal ratio of human to microbial DNA concentration for TLR9 co-activation was achieved. Similarly, microbial DNA fragmented to less than 15 bp was non-activating on its

own on Ramos Blue cells, but co-activated TLR9 with low concentrations of ODN-2006 (**Results 3.5.1**), while *Escherichia coli* and Caco-2 DNA of 50-200 bp of length was directly activating (**Results 3.7.1**). Therefore, smaller DNA fragments generally acted more like a co-activator than a direct activator, compared to larger fragments. Smaller DNA fragments can originate from apoptotic cells that undergo secondary necrosis, or from naked DNA fragments of both microbial and human origin exposed to DNases present in the serum or the intestinal fluid[146]. Therefore, disease conditions that cause increased cell apoptosis and necrosis, or inadequate removal of dead cells, will likely increase local presence of small DNA fragments. In cancer, for example, local increase of small DNA fragments from necrotic tumor cell occurs[161], and increased TLR9 co-activation with microbial DNA could exacerbate local inflammation and subsequent tumor development and metastasis[162].

The complex interaction of DNA fragments of various lengths suggested that if one simply used metagenomic DNA extracted from human stool samples to stimulate human TLR9, the results will be confounded by the varied quality of the extracted microbial DNA. Firstly, a typical DNA extraction kit already selected for fragment length. The columns of Zymo Quick-DNA Fecal/Soil Microbe Kit, for instance, only retain microbial DNA fragments greater than 50 bp, and highly fragmented DNA is lost during extraction. Secondly, even for DNA retained by the columns, varied fragment lengths could affect the ease by which microbial DNA fragments enter the endosome and reach TLR9 and the number of DNA fragments present in the local cell environment. As differences in TLR9 activation were observed when cells were stimulated by microbial and human DNA from the range of 50-200 bp versus that of over 10,000 bp (**Results 3.6.1 and 3.8.2**), metagenomic DNA stimulation of human TLR9 by DNA of varying quality and

fragment lengths extracted from the stool of different individuals is likely to generate unreliable results.

Furthermore, the gut microenvironments of an adult and a newborn, as well as healthy versus diseased individuals, can differ significantly from each other, which not only shapes the representative microbiota profile of the individuals in question, but also the quality of DNA present within the gut. The stool of an individual suffering from diarrhea, for instance, contains large amounts of water, which could both lyse bacteria[163] and induce DNA degradation[164]. Stool from newborn infants is likewise watery[165]. Therefore, it would be fruitful to study stool water content and its impact on DNA intactness, and also examine whether health conditions impacting stool consistency, such as diarrhea and constipation, can result in different degrees of DNA intactness upon extraction.

4.2.4 Additional Influence of CpG-Containing K-mers

Results 3.1.4 showed that while the human TLR9 response correlated with genomic CpG concentration of input DNA, significant variations exist, which may be linked to unusual CpG-containing k-mer concentrations such as the TCGTT-rich *Lactobacillus ruminis* and *Enterococcus faecalis*, or conversely *Micrococcus luteus* with its high-CpG but low TCGTT 5-mer concentration. Therefore, it is likely that human TLR9 does not respond to the CpG dimer alone, but CpG flanked by specific sequences. In contrast to human TLR9, mouse TLR9 is more responsive to ACGTT 5-mer. The mouse-specific ODN-1826 had previously been shown not to activate HEK cells engineered with human TLR9, but strongly activated mouse macrophages[166].

In addition, it was shown that the human TLR9 receptor responds optimally to CpG dimer and GTCGTT 6-mers that are separated by a distance of 4 to 8 bases[166]. Oligonucleotides

with a single CpG dimer were insufficient to activate human TLR9 unless co-activated by short CpG-containing sequences[63,166]. These additional k-mer patterns suggest that while CpG concentration is generally predictive of human TLR9 activation, the correlation is unlikely to be one-to-one, and other sequence patterns can cause a given bacterial genome to be more or less activating compared to predicted via their CpG concentration alone. The complex relationship of how various CpG-containing k-mers affect human TLR9 activation has been studied using self-designed synthetic oligonucleotides by Pohar et al. (2015)[167]. Therefore, while our results revealed robust correlation of human TLR9 activation with genomic CpG concentration (**Results 3.1.3**) and in specific cases, TCGTT 5-mer concentrations (**Results 3.3.2**), additional sequence patterns might explain the observed experimental variations not accounted for via correlation with single genomic CpG-containing k-mer concentrations alone.

4.2.5 Complex Formation of DNA with Peptides and Lipids

Proteins and peptides in human cells are able to complex with DNA and prevent DNA degradation in the bloodstream[75]. By forming complexes with other molecules, DNA fragments could retain lengths that directly activate human TLR9 where smaller, more degraded DNA fragments cannot (**Results 3.8.2**). Two peptides, HMGB2 and the antimicrobial peptide LL-38, for instance, were studied in conjunction with diseases such as SLE, where peptides complexed with apoptotic and necrotic DNA generate autoimmune conditions associated with TLR9 activation[78,79]. The same peptides complexed with parasite-origin DNA also enhanced TLR9 response[168]. DNA-peptide complexes are more resistant to degradation[169]. The peptides also facilitate DNA entrance into the cell, thus increasing opportunity for them to become in contact with the endosomal TLR9 present in immune cells[79].

DOTAP, a lipid-based transfection reagent, was used in our experiments to promote entrance of bacterial and human DNA into HEK-Dual TLR9 cells, Without DOTAP, the DNA concentrations we used to stimulate human TLR9 in HEK-Dual TLR9 cells was insufficient in inducing a measurable response, and very high input microbial DNA concentrations of 50 or 100 µg/mL were used to induce significant human TLR9 activation in previous studies[50,124]. The use of DOTAP posed the question of whether there exist similar lipid vesicles in real-life situations that can complex with DNA, promoting increased DNA presence inside the cells. A possible candidate is that of apoptotic blebs, which contains fragmented DNA surrounded by a lipid membrane[170]. While apoptotic blebs induced high SEAP release from Ramos Blue cells (Personal communication, Jelka Pohar and results from preliminary experiments), we were unable to reproduce the results in HEK-Dual TLR9 cells, and the question of whether blebs facilitate DNA entrance into cells and enhance TLR9 activation remains unclear.

On the other hand, DOTAP had not shown in our experiments to increase TLR9 response in Ramos Blue cells, suggesting that cells of immunological origins may be more efficient in DNA fragment uptake without assistance of a lipid. Similarly, ODN-2006 can enter both HEK-Dual TLR9 cells and Ramos Blue cells and activate TLR9 without DOTAP, though TLR9 activation by ODNs is more efficient in HEK-Dual TLR9 cells with DOTAP. Additionally, DOTAP at high concentrations could have a toxic side effect on cells (**Supplementary Figure S10** and microscopic observation)**.** Thus, DOTAP should only be used, sparingly and at a low, non-toxic concentration in the condition where it was clearly shown to assist in DNA uptake on non-specialist cell lines.

4.2.6 TLR9 of Cells from Different Origins Responded Differently

Our TLR9 activation experiments with microbial and human DNA were performed on four different cells types: HEK-Dual TLR9 cells, Ramos Blue cells, human PBMCs, and HT-29 cells. Of the four cell types, Ramos Blue cells and PBMCs are immune cells, while the TLR9 receptor was engineered onto human embryonic kidney (HEK) cells by InvivoGen to generate the HEK-Dual TLR9 reporter cells[105]. HT-29 cells were derived from human colon carcinoma[109].

We observed differences between the cell types in reaction to immunostimulatory sequences from CpG ODNs and microbial DNAs of diverse species origin. While PBMCs responded strongly and consistently to TLR9-activating sequences, IFN-β production of PBMCs extracted from different human sources vary considerably, at the difference of 5-fold, making general statistical analyses difficult (**Supplementary Figure S19**). A possible workaround for interindividual differences in TLR9 responses would be to utilize PBMCs from a single individual, which would however limit the number of experiments performed due to the amount of blood needed to extract PBMCs for stimulation experiments.

Intestinal epithelial cells, HT-29 cells included, also respond to TLR9 agonists in certain cases. In a seminal study by Lee et al. (2006), intestinal epithelial cell lines were cultivated as a monolayer on 0.4 µm filter inserts and stimulated with TLR9 agonists from the apical or the basolateral side. IL-8 and NF-kB were upregulated only when then the cells were stimulated on the basolateral side, but not on the apical side[171]. TLR9 receptors of intestinal epithelial cells was found to tolerate microbial DNA confined on the apical side through accumulation of ubiquinated IkappaB in the cytoplasm[171]. Microbial DNA at the basolateral side, however, was recognized as microbial breach of the intestinal wall and induced activation of NF-kB and subsequent inflammatory response pathways. In summary, this suggests that the presence of

101

TLR9-activating DNA within the gut lumen can be well-tolerated by TLR9-containing intestinal epithelial cells in a normal physiological state where the intestinal wall barrier is intact, and experiments using Transwall membrane setups with bacteria and TLR9 agonists could be a possible future direction in testing microbial tolerance via intact intestinal barrier alone, which could include tolerance of microbial DNA.

4.2.7 Additional Confounders

Specialized classes of TLR9-activating synthetic oligonucleotides for both human and mouse cells have been previously identified[62]. The human Class A ODN, ODN-2216, has a stem loop structure due to presence of palindromic sequences that must first be separated by DNase II before TLR9 activation can occur[172], while the Class B ODN, ODN-2006, is linear. Class A and Class B ODNs also selectively activate certain cell types more than others[62].

That different cell types respond better to either class A or class B ODNs suggests that there exist higher-order sequence patterns that influence human TLR9 activation. Differences in the most effective stimulatory ODN additionally implied that either TLR9 present with structural differences in different cell types, or that different ODN classes were pre-processed differently in different cell types prior to contact with TLR9, or both.

The signaling cascades downstream of TLR9 activation also differs significantly depending on cell type[42]. While TLR9 activation of pDCs results in activation of the interferon signaling pathway, inflammatory signaling cascades are activated in B cells, resulting in production of IL-8 and cell proliferation[24]. TLR9 receptors of hepatocytes respond by activation of IL-1β inflammasome upon exposure to ODNs and apoptotic DNA[69]. The differences in downstream signaling introduce major obstacles in generating a standardized *in vitro* model of

102

human TLR9 activation across cell lines. Instead, the specific cytokine reporter to be used that reflects TLR9 activation per cell line must be individually determined.

Finally, other pattern recognition receptors involved with DNA sensing have been identified, such as cytosolic sensors that recognize viral DNA[123]. Specifically, cytosolic viral DNA induces production of cyclic AMP-GMP (cGAMP), which binds to the adaptive protein STING (Stimulator of Interferons Gene)[173]. A member of the nucleotidyltransferases named cGAMP synthase (cGAS) catalyzes cGAMP synthesis upon sensing cytosolic dsDNA[174]. Upon binding to cGAMP, STING activates TBK1, which phosphorylates the transcription factor IRF3, which then then signal the transcription of IFN-ß and inflammatory genes in a manner similar to that downstream of TLR9 activation, and STAT6, which induces the production of chemokines[175]. That human pDCs produce IFN-β subsequent to the activation of either TLR9 or cGAS activation upon exposure to DNA poses question of whether IFN-β production is solely due to TLR9 activation. Therefore, we cannot fully rule out potential responses of cytosolic sensors to bacterial DNA in our examination of TLR9 activation in PBMCs by choosing IFN-β as our cytokine reporter.

4.3 Generalizing *in vitro* TLR9 Activation Assays to Real-Life Scenarios: Insights and Limitations

4.3.1 Large Amount of DNA is Needed to Activate Human TLR9

Published studies on TLR9 activation using microbial-origin DNA typically used a large amount of bacterial DNA. In Dalpke et al (2006), for instance, 100 µg/mL of microbial DNA was used[50]. The use of DOTAP, which facilitated cell intake of microbial DNA, reduced the

input DNA concentration to 30 µg/mL[50]. The high concentration of microbial DNA needed for TLR9 activation raised the question of whether there exist realistic situations by which the human cells can come into contact with such large amounts of DNA. A possible location for the very high DNA concentration is that of the intestinal lumen. Certain cells, such as professional phagocytes, could actively take in bacteria from the intestine[143], accumulating bacterial DNA in sufficient amount inside the cells for TLR9 activation to occur.

HEK-Dual TLR9 cells, where TLR9 was engineered to HEK cell lines, is a more sensitive cell line that responds to lower amounts of input DNA. The TLR9 response of this cell line, according to the source company InvivoGen, is said to be exaggerated (personal communication with company). On the other hand, Ramos Blue cells, a cell line of B lymphocytes with a knocked-in SEAP reporter downstream of NF-kB, possesses natural TLR9 receptors.

We identified major sensitivity differences in the TLR9 response between HEK-Dual TLR9 cells and Ramos Blue cells. For positive controls of our experiment, we used 0.5 µg/mL of ODN-2006 for HEK-Dual TLR9 cells, but 25 µg/mL for Ramos Blue cells, a difference of 50-fold. Moreover, even at the high concentration of 25 µg/mL, bacterial DNA extracted from *Escherichia coli* was insufficient to activate TLR9 in Ramos Blue cells without pre-starvation. The results suggested that the results from artificial systems based on engineered cell lines might not be fully applicable to natural human TLR9 response to microbial DNA in real-life contexts.

PBMCs, which were extracted directly from human blood samples and several types of TLR9 receptor-containing cells, were generally more sensitive to ODN-2006 and bacterial DNA compared to Ramos Blue B cells. Yet, they also showed varied sensitivity to both ODN-2006 and extracted bacterial DNA depending on the blood source, which became a major source of

technical difficulties (**Supplementary Figure S19**). On the other hand, the same result could

offer insight into variation of innate immune response in different individuals and potentially be

correlated to other human phenotypes such as allergy proneness[176].

4.3.2 Possible Scenarios of TLR9 Activation in the Human Gut in Health and Disease

    A realistic location for TLR9 activation by microbial DNA is the human colon where the

vast microbial population takes residence. The mucus, a powerful gel-like matrix secreted by

Paneth cells, normally deters direct contact between the microbiota and the intestinal

epithelium[82]. The human microbiota is strictly confined in the intestinal lumen. The mucosal

barrier and the barrier of intestinal epithelial cells form two impenetrable walls, preventing the

microbiota from crossing over and becoming a problematic source of immune activation[125].

Contact between TLR9-receptor containing immune cells and the microbiota is limited, and only

in select regions of the intestines, such as the Peyer's Patch in the human small intestine, are

bacteria funneled through specialized cells such as M cells and which could then contact local

immune cells populations, providing small-scale recognition of bacteria by innate and adaptive

immune cells[127,128].

    Several diseases are believed to increase contact between the microbiota and immune

cells. When the intestinal barrier is breached, the microbiota can more easily penetrate through

the intestinal epithelial cell barrier[92]. Diseases known to compromise the tight junctions of the

intestinal wall include IBD and celiac disease. In IBD, chronic inflammatory conditions in the

intestine lead to the loss of intestinal wall integrity[92], while in celiac disease, gluten activates

zonulin signaling, increasing permeability of the intestinal wall to macromolecules[177]. Increased

intestinal permeability were also implicated in diabetes[177], rheumatoid arthritis[178], and liver

cirrhosis[179], suggesting the role of increased contact between bacteria and human somatic cells in inflammatory conditions and disease development. In these conditions, contacts of multiple PRRs in somatic and immune cells and bacterial components could potentially be involved, among which TLR9. Our assays suggested the potential role of increased cell death, which occurs in the context of IBD[87], to result in possible combinatorial effect in TLR9 activation together with bacterial DNA. In infectious diseases, invasive pathogens are also known to either invade intestinal epithelial cells directly or across the intestinal wall by exploiting tight junctions or specialized cell types[82,127,128]. Some pathogens could also induce gut cell apoptosis and necrosis, thus weaken the intestinal tight junction and increase gut permeability[180,181]. Our *in vitro* experiments combining human and microbial DNA of various lengths thus provided a possible explanation of the interaction between bacterial DNA, human DNA from apoptotic and necrotic cells, and human TLR9 in the gut environment in several disease contexts.

Colorectal cancer is another condition where TLR9 activation could be involved. The tumor microenvironment offers a rich opportunity for interaction between tumor cells, the microbiota, and immune cells[182]. Increased blood flow into tumor growth, a consequence of angiogenesis, transports a variety of immune cells into the tumor growth[183], where they can interact with the microbiota, as well as apoptotic and necrotic cells within the tumor[184,185]. The microbiota, along with apoptotic and necrotic cells, are both rich sources of DNA leading to TLR9 activation and co-activation[63] which is likely to induce an inflammatory response that can aggravate tissue damage and tumor development[150]. Colorectal tumors can also become key locales of intestinal barrier compromise, increasing local trafficking of both bacteria and immune cells[186]. Therefore, TLR9 activation from increased exposure of the receptor to bacterial and human DNA can be a common underlying cause of both IBD and colorectal cancer[187], and play a

role in the progression of colorectal cancer from its *in situ* to metastatic stages[188]. More studies are need to precisely identify how TLR9 activation could occur in and around tumor sites.

4.4 Future Experimental Directions

Results of preliminary experiments using HT-29 cells and *Fusobacterium nucleatum* indicated that TLR9 activation via ODN-2006 and/or human-origin DNA increased *Fusobacterium nucleatum* invasion. As identical experiments were performed only twice, no statistical analysis was available. Interestingly, increased IL-8 production was observed with cells treated with non-invasive *Escherichia coli* DH5α and invasive *Fusobacterium nucleatum* with or without addition of ODN or human DNA, suggesting that HT-29 cells have additional surface receptors that recognize bacterial presence in the environment, resulting in release of inflammatory cytokines regardless of bacterial invasion[189].

Thus, the most urgent future task is to carry out additional invasion experiments with TLR9 activation to confirm our preliminary observations. If TLR9 activation increases *Fusobacterium nucleatum* invasion, the question remains whether it is due to TLR9 activation in particular, or to general activation of the inflammatory signaling cascades independent of TLR9. Pro-inflammatory cytokines such as IL-8 and TNFα could be added HT-29 cells directly and examine if they also increase invasion. TLR9 activation could be one among the many sources that increase cell membrane permeability and/or phagocytosis of bacteria[149] (**Discussion 4.1.3**). Invasion by other known invasive bacteria with different genomic CpG concentration, such as *Shigella spp.*, could also be used to test if prior exposure to TLR9 agonists induces more or less efficient internalization of higher-CpG bacteria, thus demonstrating human TLR9's ability to discriminate and selectively internalize bacteria based on their CpG concentration.

Metagenomic DNA stimulation of human TLR9 could also provide additional insights. Due to the numerous practical difficulties discovered during the project, few metagenomic TLR9 stimulation experiments were conducted. While test stimulations were performed on PBMCs and Ramos Blue cells using adult and infant metagenomic DNA directly extracted from stool samples, the quality of DNA extracted, as indicated by their fragment length, played a major confounding role. Whether infant metagenomic DNA, with its high concentration of high-CpG bacterial species such as *Bifidobacterium spp.*, can predictably cause higher TLR9 activation than adult metagenomic DNA, remains inconclusive.

As DNA of different species origins fragment differently, a possible solution is to use synthetic metagenomic DNA that approximates typical adult or infant metagenomic DNA, combining typical gut species by first cultivating the species separately, extracting their DNA, and fragmenting them to a standard fragment length, then combining them in typical proportions to approximate the more common microbial population of an adult or infant gut. The synthetic method provides a large and unlimited amount of source DNA material, which could compensate for the low volume of infant metagenomic samples and thus generate the large amount of DNA required for typical stimulation experiments. It also permits standardization of fragment length, which was a major technical barrier to accurately characterize metagenomic CpG concentration from infant and adult sources and their ability to activate human TLR9.

Co-cultivation of intestinal epithelial cells and immune cells could be performed as two layers via the Transwell membrane system. Intestinal epithelial cells such as Caco-2 and HT-29 cells can be grown into a confluent monolayer on the membrane, and left to differentiate until they exhibit characteristic small intestine cell morphologies such as the brush border. In addition, a subclass of HT-29 cells, called the HT-29 MTX, are mucus-producing[190]. A fully differentiated

Caco-2 or HT-29 cell monolayer have tight junctions and is impenetrable to microorganisms. Protocols for the generation of differentiated Caco-2 and HT-29 cell monolayer are well-established and used extensively in drug permeability Studies[190,191]. The co-cultivation setup is useful to simulate the effect of an intact physical barrier, such as the intestinal cell monolayer, on TLR9 activation, as characteristic of a homeostatic condition in the human large intestine where high concentration of microbial and human DNA is present within the intestinal lumen.

In co-cultivation, immune cells such as B cells could be cultivated in the lower layer, with a membrane containing fully differentiated intestinal epithelial cell monolayer placed on top. TLR9 ligands, such as ODN-2006, as well as living bacteria, could be added to either side of the Transwell membrane. It was shown that stimulatory ODNs did not cross over the intestinal epithelial cell monolayer, and when introduced on the apical side, did not generate a TLR9-dependent response, but those added at the basolateral side did[171]. A possible experimental setup will be to introduce stimulatory ODN at either side of the intestinal epithelial cell monolayer, thus inducing different levels of TLR9 activation, then add invasive bacteria on top of the monolayer to examine the effect of TLR9 activation on invasion.

Finally, several mouse models that have been developed are particularly suited for TLR9 activation and inflammation studies. TLR9 knockout mouse models, used as controls in studies of liver inflammation[192], mitochondrial DNA activation[68], and parasite invasion[193], was an obvious choice. Additional mouse models include germ-free mice, which develop a defective mucus more penetrable to bacteria due to lack of developmental exposure[84]. Germ-free mice could also be inoculated with single bacterial species to create gnotobiotic mice and thus be used as models of innate immune development in early life. IL-10 knockout mice, which develop spontaneous colitis, could be used to study the role of TLR9 in the context of IBD[194].

5. Conclusion

Activation of human TLR9 was verified experimentally to correlate to genomic CpG concentration of a given microbial genomic DNA, though the flanking sequences around CpG could play an additional role. As the fecal metagenomic DNA of infants were determined via bioinformatics analyses to be higher in the CpG dimer compared to that of adults[44], results suggested that the TLR9-stimulatory aspects of infant metagenomic DNA could be involved with the development of innate immune system in early life.

Additionally, we identified more complex patterns of human TLR9 activation which included direct activation, co-activation, and repression of human TLR9 when DNA of various CpG concentration and fragment lengths were used, which could provide insight relevant to the human gut microenvironment, in which large amounts of DNA of various origins and qualities can co-present. The results of our stimulation experiments can thus be relevant to potential TLR9-dependent inflammatory response and autoimmunity that might characterize a wide range of disease conditions of the human gut such as IBD, infectious diseases, and colorectal cancer.

Further *in vitro* and *in vivo* studies are needed to clarify the impact of TLR9 activation within the context of the human gut environment, as well as specific disease situations that can lead to increased contact between the human gut microbiota and TLR9-expressing cells. Additional topics can include how TLR9 activation affect bacterial invasion and internalization of bacteria by host cells, the role of human-origin DNA released from apoptotic and necrotic intestinal epithelial cells in TLR9 co-activation, as well as TLR9 involvement within the tumor microenvironment in colorectal cancer models.
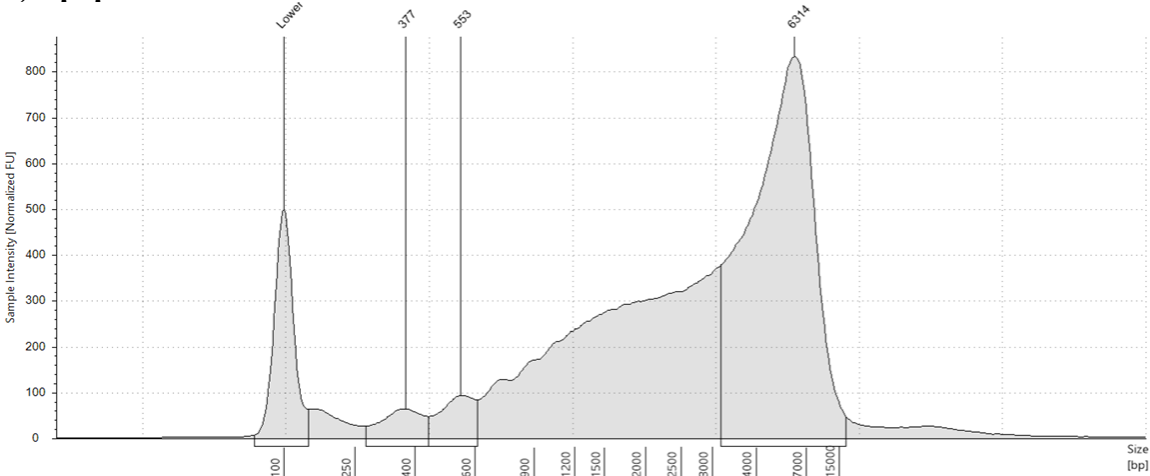
6. Supplementary Figures
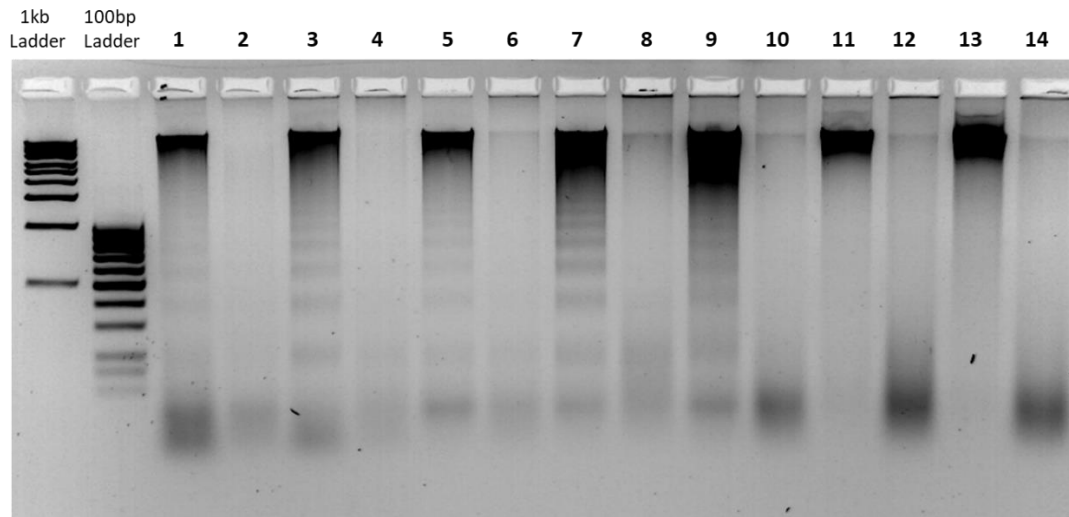
Gel Loading Order:
1. **Caco-2 6-day 2 µM Staurosporine (STS) 8-Hour Treatment (8h) Pellet**
2. **Caco-2 6-day 2 µM STS 8h Bleb**
3. **Caco-2 6-day 8 µM STS 8h Pellet**
4. **Caco-2 6-day 8 µM STS 8h Bleb**
5. **Caco-2 6-day 2 µM STS 24h Pellet**
6. **Caco-2 6-day 2 µM STS 24h Bleb**
7. **Caco-2 6-day 8 µM STS 24h Pellet**
8. **Caco-2 6-day 8 µM STS 24h Bleb**
9. **Caco-2 2 µM STS 96h Pellet**
10. **Caco-2 2 µM STS 96h Bleb**
11. **Caco-2 4 µM STS 96h Pellet**
12. **Caco-2 4 µM STS 96h Bleb**
13. **Caco-2 8 µM STS 96h Pellet**
14. **Caco-2 8 µM STS 96h Bleb**

Supplementary Figure S1. Gel Electrophoresis of Caco-2 DNA Undergoing Staurosporine Treatment. Staurosporine was used to induce apoptosis in Caco-2 cells, and different Staurosporine concentrations and treatment times were attempted to induce the characteristic ladder pattern of apoptotic DNA. Apoptotic DNA ladder was enriched in the flowthrough from the spin column of Qiagen's DNeasy Blood and Tissue Miniprep Kit as opposed to the DNA extracted from columns, where large fragments were found. Therefore, a modified protocol was used (**Methods 2.3.2**), where the column flowthrough was ethanol precipitated and then cleaned with the Zymo Clean-and-Concentrator-25 kit. DNA extracted from apoptotic blebs were small and enriched in the flowthrough, where it was ethanol precipitated and later cleaned with the Zymo Oligo Clean-and-Concentrator Kit. Based on the results of this gel, Caco-2 cells were left to partially differentiate for 6 days and treated with 8 µM Staurosporine for 24 h for all future apoptotic DNA preparations.

a) Apoptotic DNA

b) 5 min
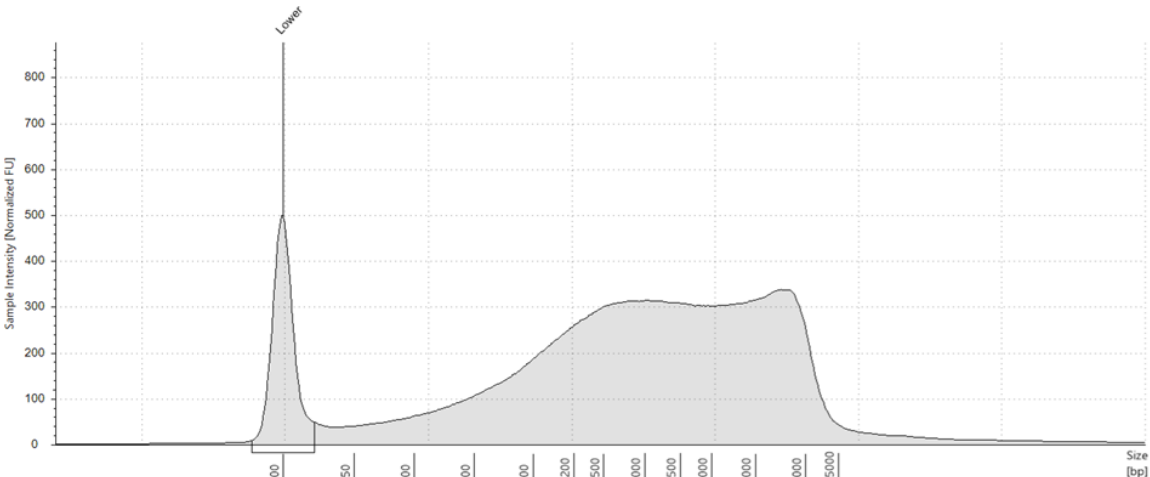
c) 7 min
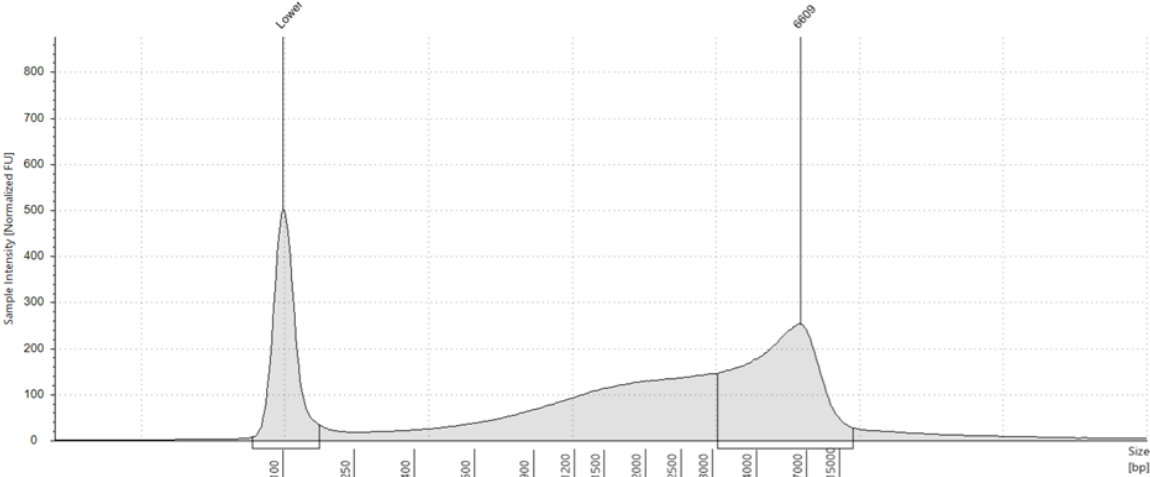
d) 10 min
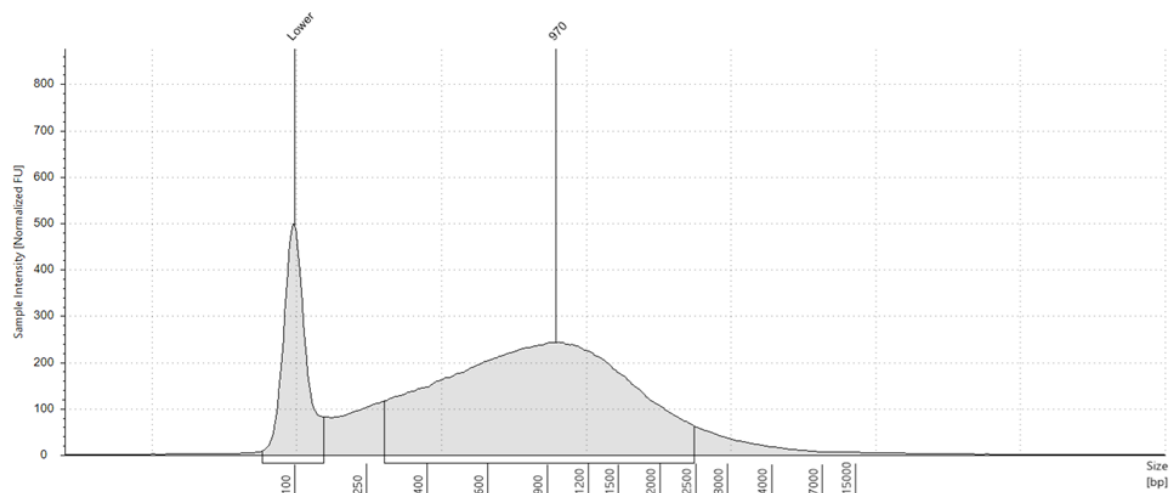
Supplementary Figure S2. Size Distribution of NEB dsDNA Fragmentase-Treated Caco-2 DNA Revealed 7 Min as the Optimal Fragmentation Time. a) Fragment length distribution of DNA extracted from apoptotic bodies of Caco-2 cells according to the modified protocol (**Methods 2.3.2**). **b)** to **d)** Fragment length distribution of Caco-2 DNA after **b)** 5 min **c)** 7 min, and **d)** 10 min of NEB dsDNA Fragmentase Treatment. After DNA was fragmented for 7 min, a peak at 6609 was found with a smooth slope downward from 4000 to 1000 bp, and a smaller number of fragments at the range of 200 to 1000 bp, similar to the size distribution of apoptotic DNA without the ladder peaks. DNA was fragmented without $MgCl_2$ and cleaned up prior to the loading of the tape station. All DNA samples above 100 ng/µL of concentration were pre-diluted to 100 ng/µL.
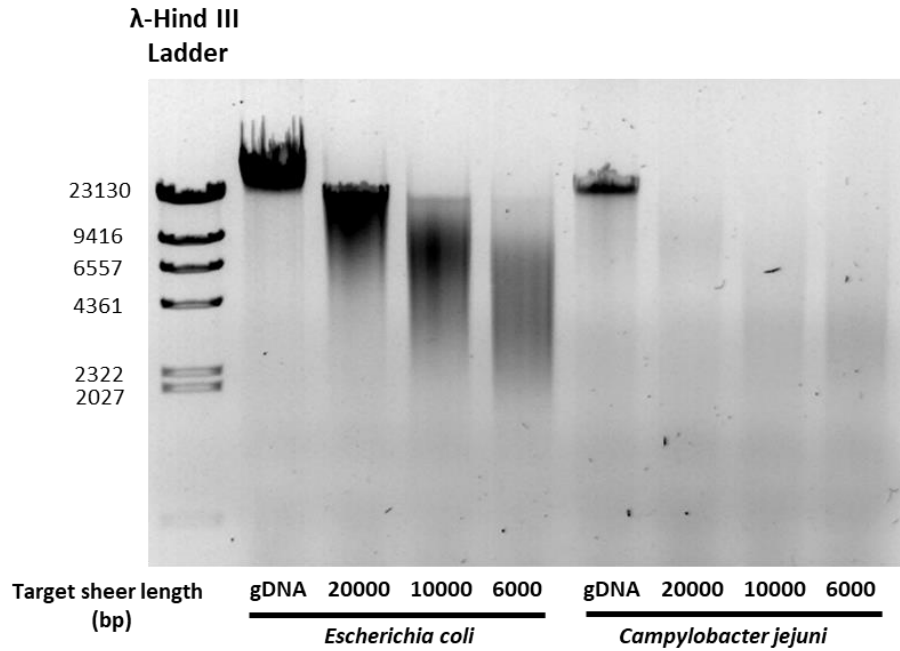
Supplementary Figure S3. Fragment Length Distribution of *Escherichia coli* and *Campylobacter jejuni* DNA sheared the target lengths of 20,000, 10,000 and 6000 bp via the Covaris g-tube. Sheared DNA was loaded onto a 0.7% Agarose gel with the Lambda Hind III ladder. The numbers at the left indicated the fragment length in bp for the Lamba Hind III Ladder used to characterize large DNA fragments.

a)

b)

c)
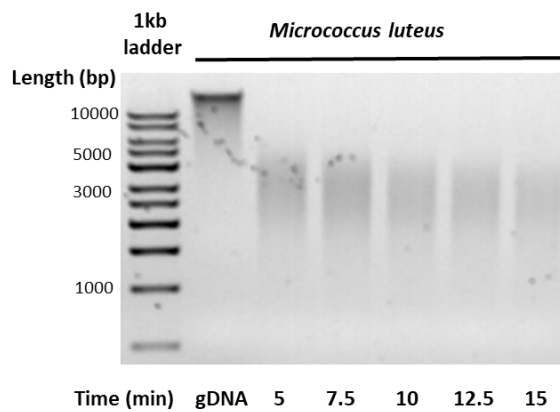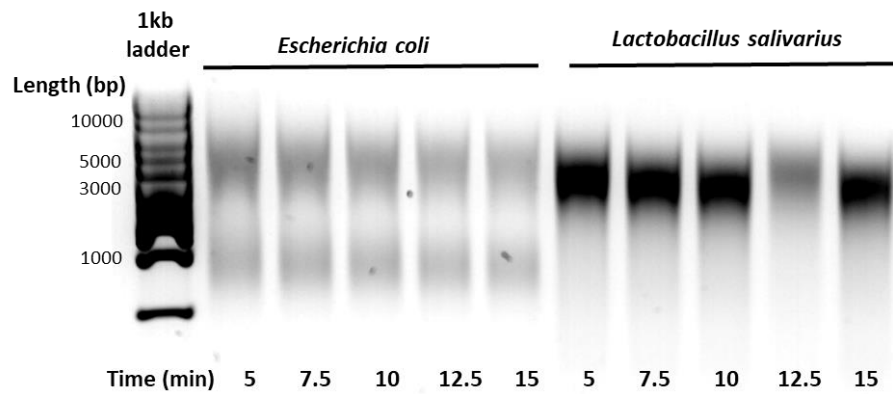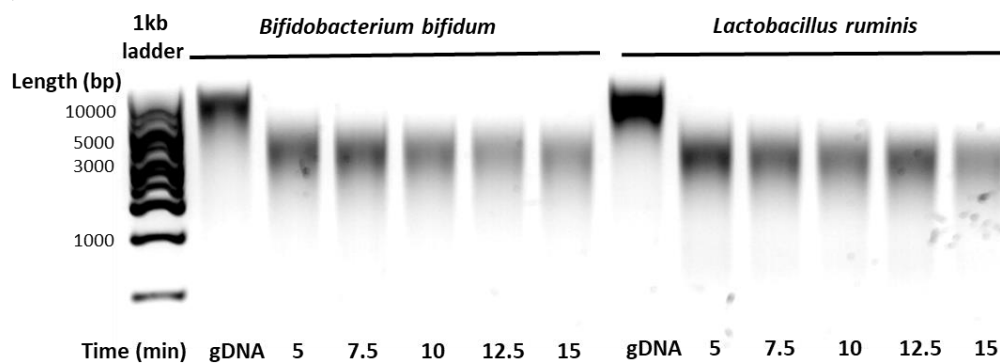
Supplementary Figure S4. Time Course of Five genomes Sheared by Sonication via the Covaris S220 Instrument with 3kb miniTUBE (Covaris Inc.) Showed Fragmentation to 3kb of Length after 10 min of Sonication. Fragment length distribution of time course experiment using 5, 7.5, 10, 12.5 and 15 min of sonication. The sonication time specified by Covaris Inc., 600 seconds (=10 minutes), was suitable for all five genomes. Representative gel (0.7% Agarose with 1kb Ladder) of **a)** *Micrococcus luteus*, **b)** *Escherichia coli* and *Lactobacillus salivarius*, and **c)** *Bifidobacterium bifidum* and *Lactobacillus ruminis*.

a) IFN-β
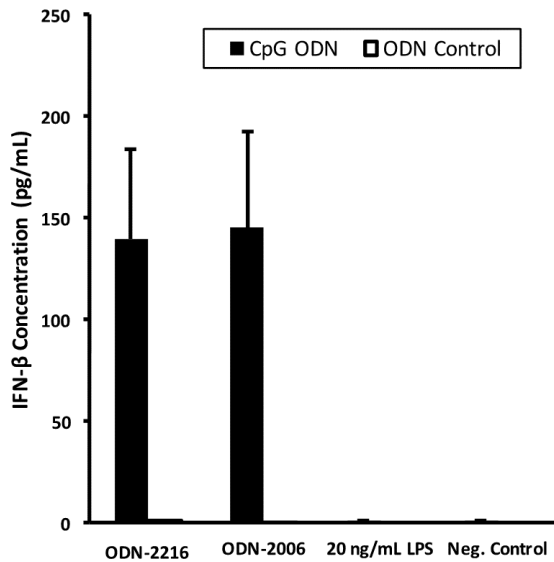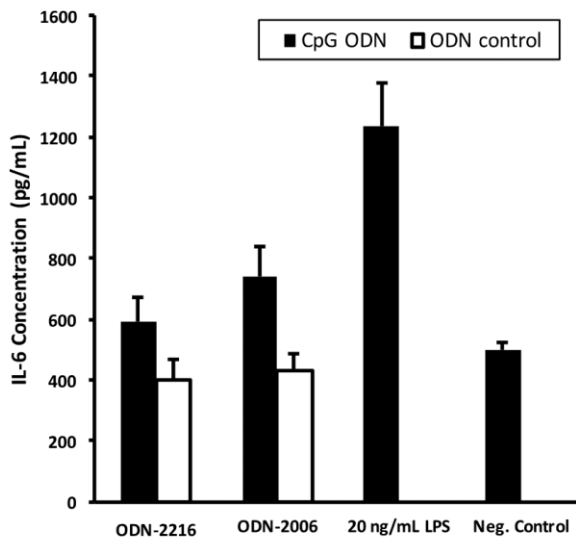
b) IL-6

Supplementary Figure S5. IFN-β was Produced Only when TLR9 Agonists were Used.
Supernatant concentration of **a)** IFN- β and **b)** IL-6 of PBMCs treated with 1 µM of CpG-ODN (TLR9-stimulatory oligonucleotides), ODN control, 20 ng/mL of *Escherichia coli* lipopolysaccharides (LPS), or serum-free medium (as negative control), as measured by ELISA (bars ± s.d. of duplicates)
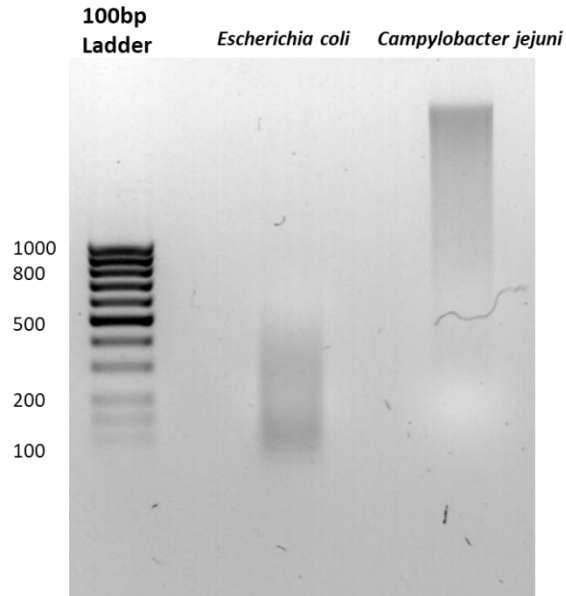
Supplementary Figure S6. Fragment Lengths of *Escherichia coli* and *Campylobacter jejuni* Genomic DNA After 8 Min Digestion by NEB's dsDNA Fragmentase Showed Major Differences. Extracted and fragmented DNA were loaded onto 1.5% agarose gel with 100 bp ladder (Carl Roth). The target fragment length of 200bp was not reached by *Campylobacter jejuni* DNA after 8 minutes of fragmentation time, suggesting major intergenomic differences in response to enzymatic fragmentation via NEB's dsDNA Fragmentase.
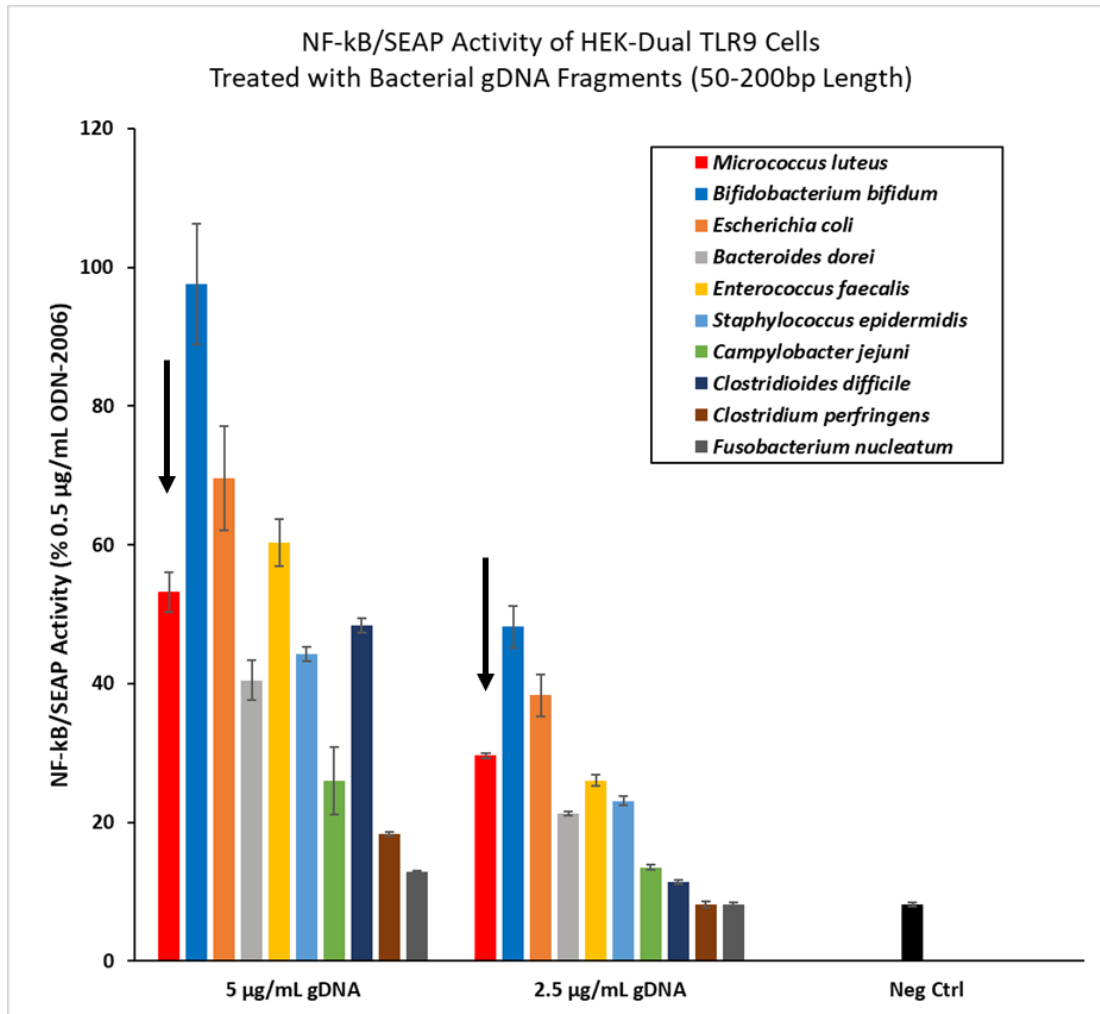
Supplementary Figure S7. *Micrococcus luteus* **DNA Induced Lower TLR9 Activation than Expected from Its Genomic CpG Concentration.** *Micrococcus luteus* DNA (red bars, indicated by arrows), the genome with the highest (14.97 %) CpG concentration, induced lower SEAP production from HEK-Dual TLR9 cells compared to *Bifidobacterium bifidum* and *Escherichia coli* at both 5 and 2.5 µg/mL of concentration (result from one experiment, bars ± s.d. of duplicates).

a) CpG
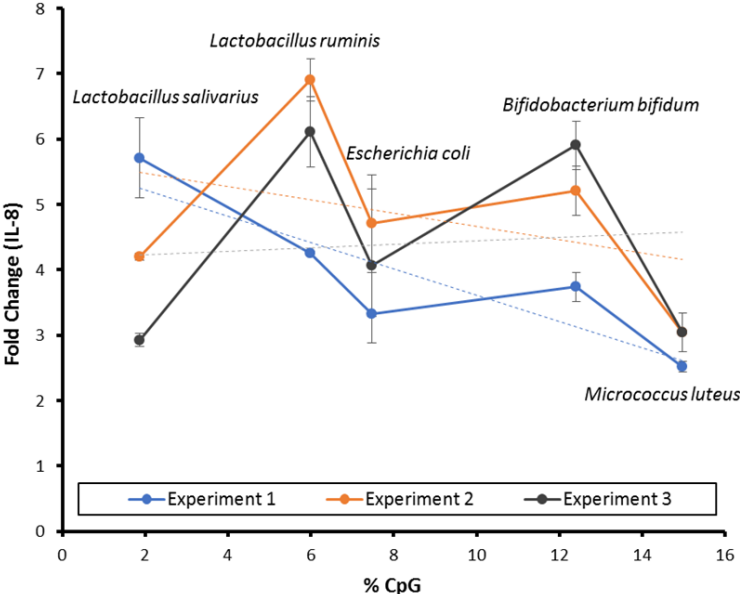
b) TCGT 4-Mer
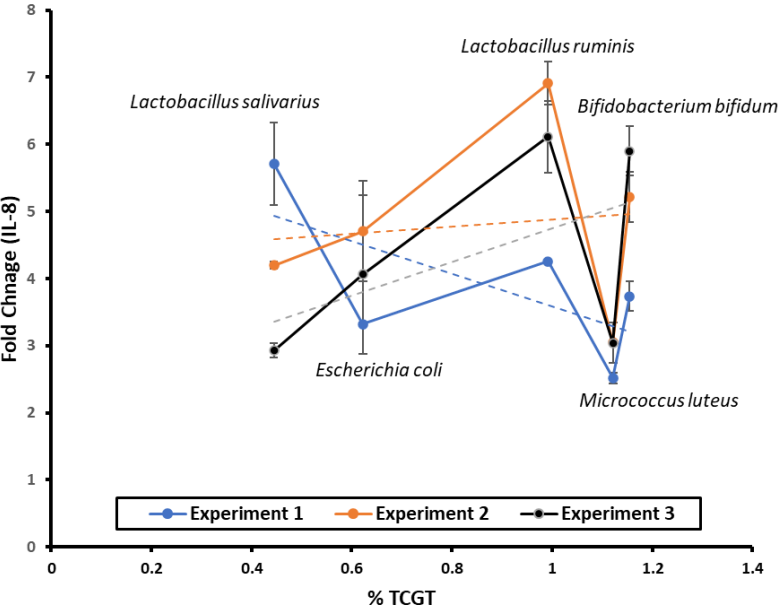
c) TCGTT 5-Mer
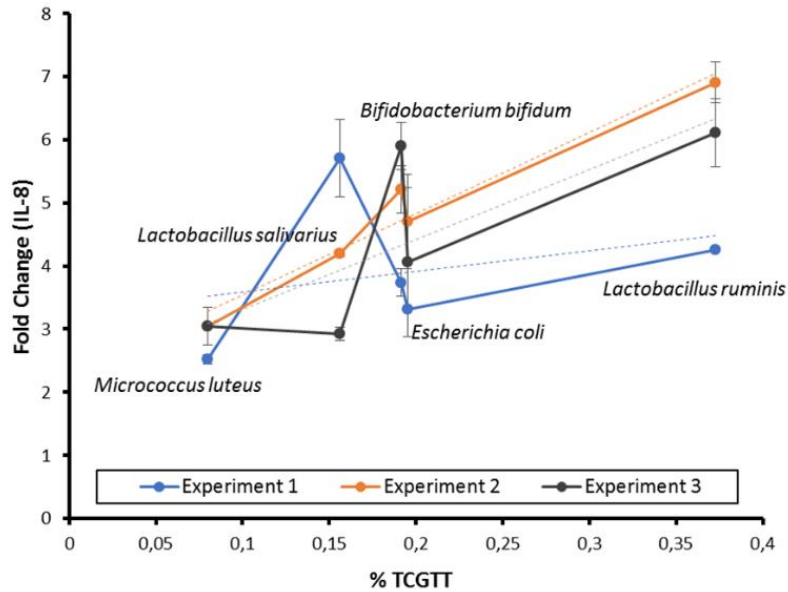
Supplementary Figure S8. TLR9 Activation of Five Genomes Relative to Bacterial CpG, TCGT 4-Mer and TCGTT 5-Mer Concentration Using Lucia Luciferase Under the Control of Human IL-8 Promoter. HEK-Dual hTLR9 cells were stimulated with 5 µg/mL of bacterial genomic DNA from *Micrococcus luteus*, *Bifidobacterium bifidum*, *Escherichia coli*, *Lactobacillus ruminis* and *Lactobacillus salivarius* fragmented to 3kb via Covaris AFA sonication and size-selected with SPRI beads. IL-8 response was calculated by dividing the luminescence value of gDNA-stimulated cells by negative control. IL-8 fold change was plotted against **a)** genomic CpG, **b)** TCGT 4-mer, and **c)** TCGTT 5-mer concentration of the five species tested (results from three independent experiments, bars ± s.e.m.).

Titration Curve of ODN-2006 on Ramos Blue Cells ±DOTAP
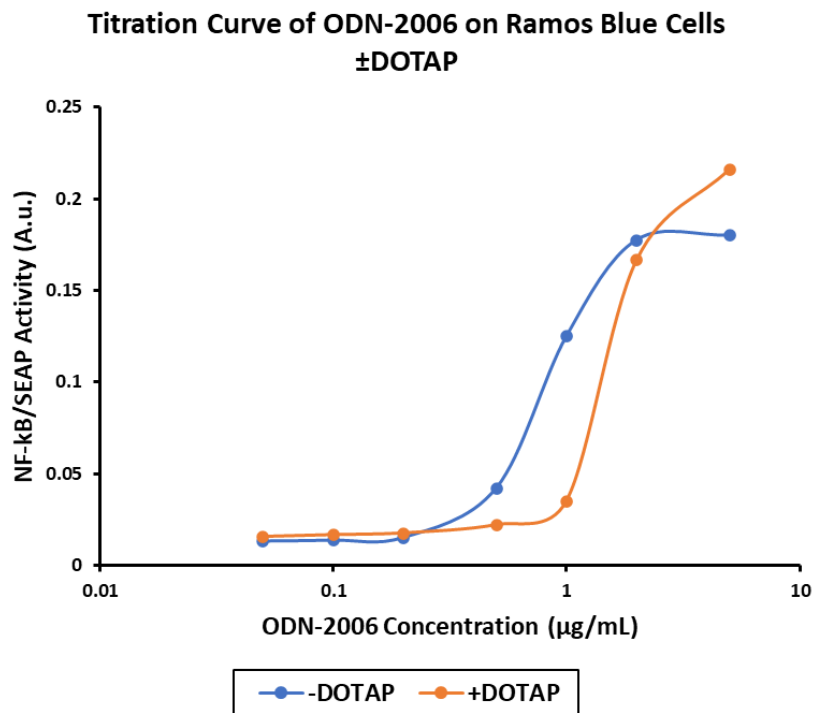
Supplementary Figure S9. Titration Curve of ODN-2006 on Ramos Blue Cells With or Without DOTAP. A titration experiment using ODN-2006 on Ramos Blue cells showed that DOTAP did not improve cellular TLR9 response (shown as SEAP activity under the control of NF-kB). Instead, addition of DOTAP reduced SEAP production, suggesting possible toxicity. Therefore, DOTAP was not used in subsequent experiments using Ramos Blue cells.

Gel image with ULR Ladder and lanes 1–6. Ladder markings at 300, 100, 50, 20, 10 bp. Length (bp).

1. *Micrococcus luteus*
2. *Bifidobacterium bifidum*
3. *Escherichia coli*
4. *Bacteroides dorei*
5. *Lactobacillus salivarius*
6. *Campylobacter jejuni*
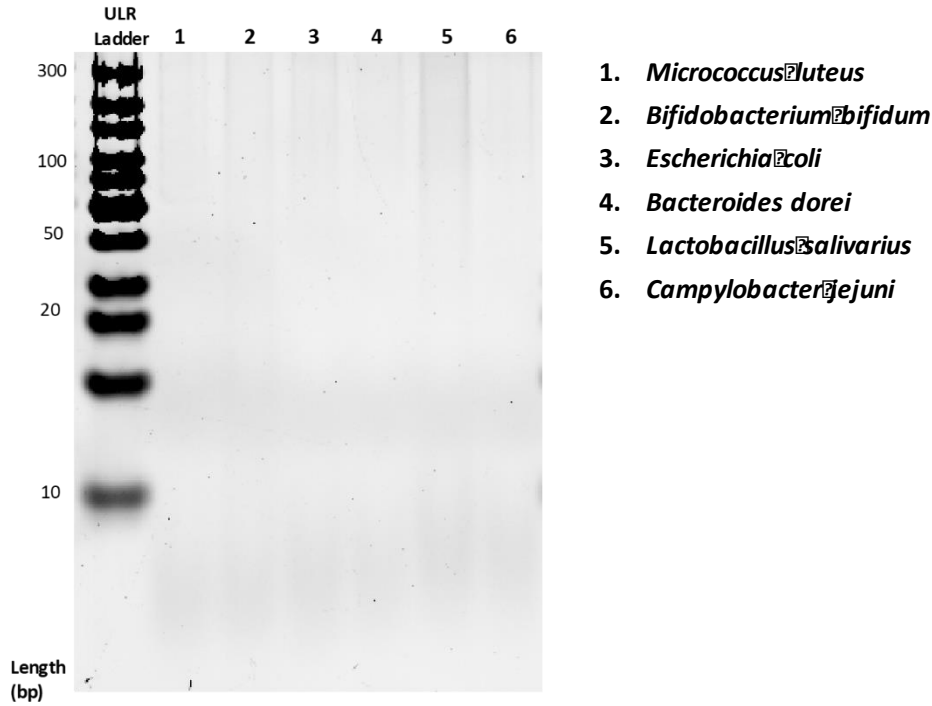
Supplementary Figure S10. Fragment Lengths of DNase I-Digested Genomic DNA. 2 µg of DNase I-digested genomic DNA fragments from all six genomes were loaded onto a PAGE gel to confirm complete digestion. DNA was confirmed to be fragmented to less than 15 bp of length. Gel with 20% Acrylamide:Bisacrylamide was used and DNA fragment length was compared to Ultra Low Range Ladder (Thermo Scientific).

a)

Scatter plot of NF-κB/SEAP Activity (% 25 µg/mL ODN-2006) versus species (Micrococcus luteus, Bifidobacterium bifidum, Escherichia coli, Bacteroides dorei, Lactobacillus salivarius, Campylobacter jejuni, Neg. Control). Plating Density (Cells/Well): 50000, 200000.

b)
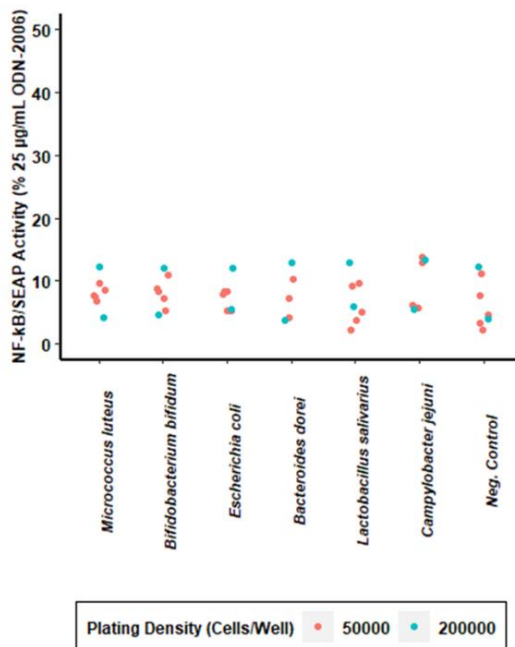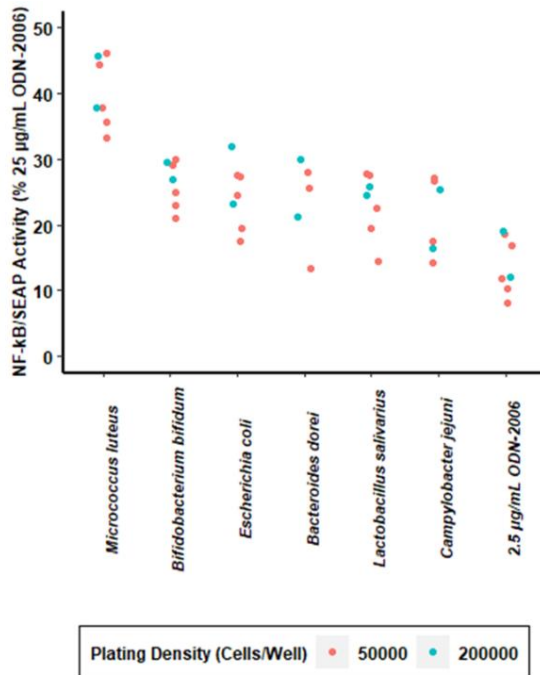
Supplementary Figure S11. TLR9 Activation and Co-Activation Normalized to Plate Positive Control was Not Affected by Plating Density. TLR9 co-activation experiments by 25 µg/mL of DNase I-digested bacterial DNA **a)** without or **b)** with 2.5 µg/mL of ODN-2006 were performed on Ramos Blue cells plated at 50000 or 200000 cells/well in 96-well plates. Results suggested no difference in TLR9 activation normalized to plate positive control of 25 µg/mL ODN-2006 for all conditions. Each dot represented results of one independent experiment (five experiments for 50000 cells/well and two experiments for 200000 cells/well, sorted by color).
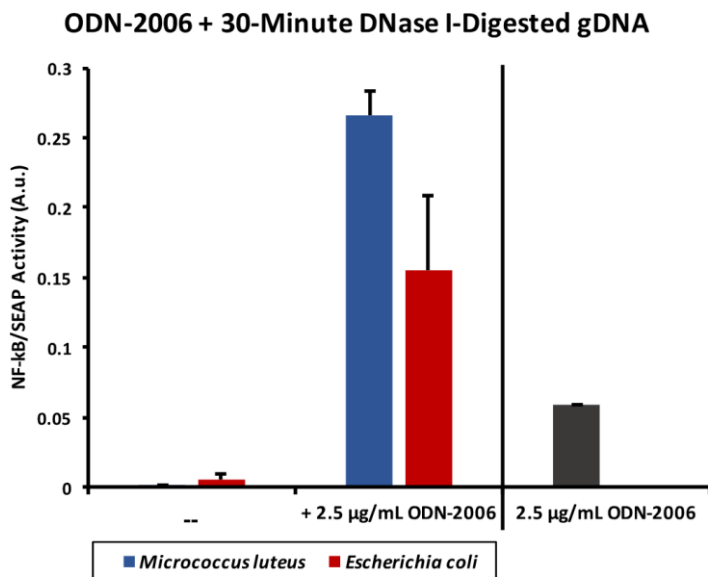
Supplementary Figure S12. Human TLR9 Activation of DNase I-Digested Bacterial gDNA With or Without ODN-2006. Bacterial genomic DNA from *Micrococcuccus luteus* and *Escherichia coli* were digested for 30 minutes with DNase I and added to Ramos Blue cells at the concentration of 25 µg/mL together with 2.5 µg/mL of ODN-2006. 25 µg/mL bacterial gDNA or 2.5 µg/mL ODN-2006 were added alone as comparison (mean from two independent experiments, bars ± s.e.m.).

a)

b)

c)
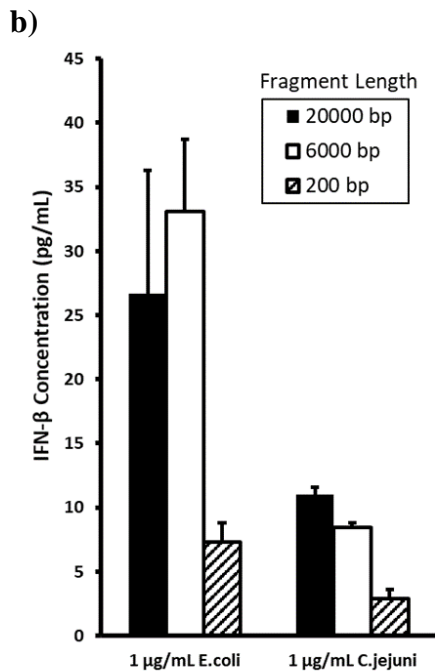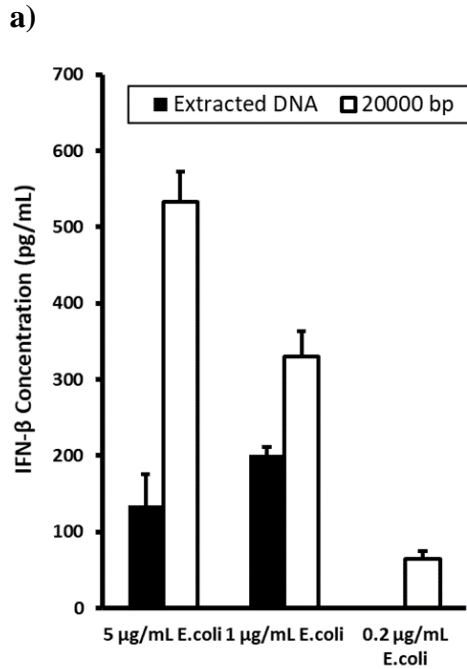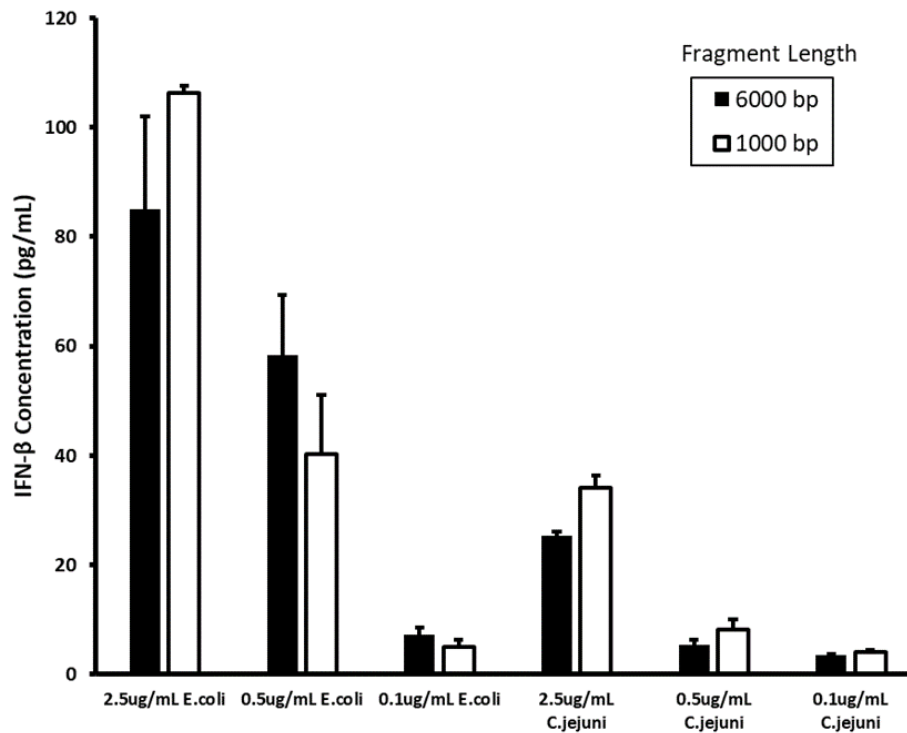
Supplementary Figure S13. Bacterial DNA Fragment Length Had Some Effect on TLR9 Activation in Human PBMC. a) Extracted *Escherichia coli* DNA was less activating of human TLR9 on PBMCs compared to DNA sheared to 20,000 bp of length at final concentrations of 5, 1, and 0.2 µg/mL. **b)** *Escherichia coli* and *Campylobacter jejuni* DNA were fragmented to 20,000, 6000, and 200 bp and used to activate human TLR9. Short DNA fragments of 200 bp from both genomes activated TLR9 less compared to longer fragments. **c)** No difference in TLR9 activation between 6000 and 1000 bp of length was seen for *Escherichia coli* and *Campylobacter jejuni* DNA at three different final concentrations. TLR9 activation was measured as IFN-β concentration released in the supernatant via ELISA. Results from one experiment with bars representing ± s.d. of duplicates.
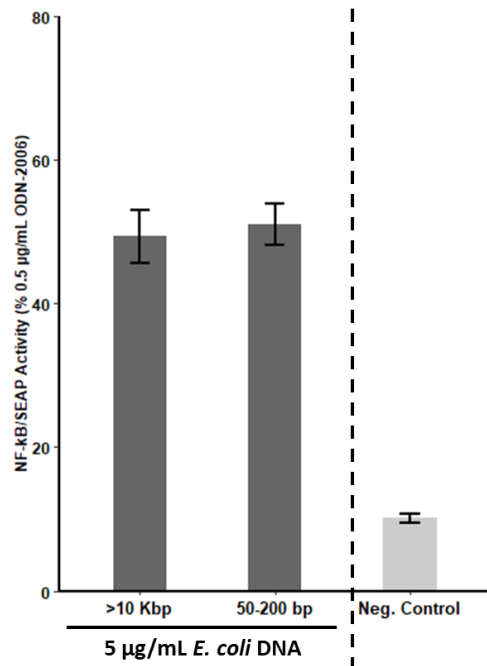
Supplementary Figure S14. No Difference was Found in TLR9 Activation by Bacterial DNA at >10,000 bp and 50-200 bp of Length on HEK-Dual TLR9 Cells. Extracted *Escherichia coli* DNA (>10,000 bp fragment length) showed similar abilities to activate TLR9 in HEK-Dual TLR9 cells compared to *E. coli* DNA fragmented to 50-200 bp. A standardized DOTAP concentration of 10 µg/mL was used for all wells including positive and negative control. TLR9 activation was measured as SEAP activity normalized to plate positive control of 0.5 µg/mL ODN-2006 (results from two independent experiment, bars ±s.e.m.).

a) 50-200 bp
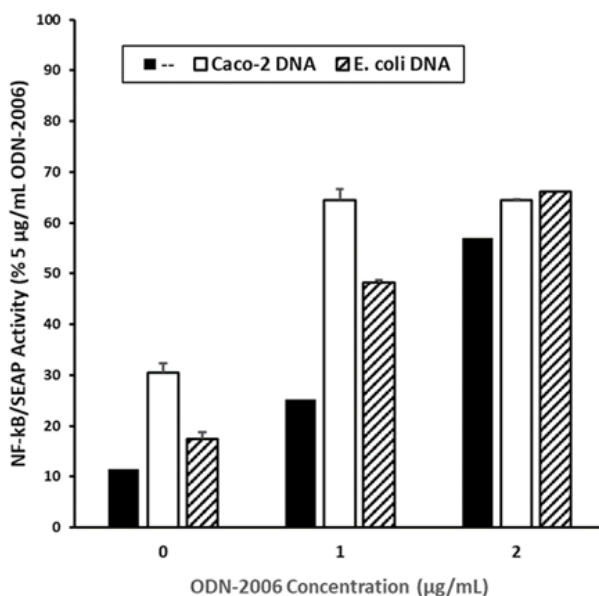
b) < 15 bp
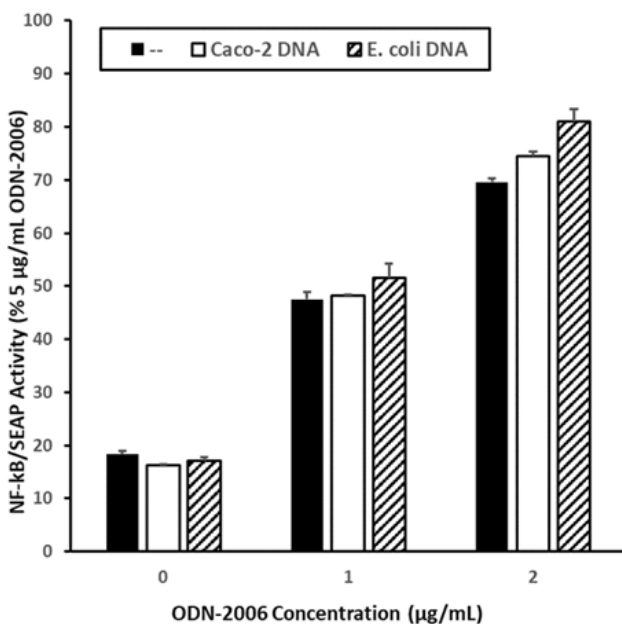
Supplementary Figure S15. Co-Activation of ODN-2006 by Bacterial and Human DNA at Two Different Fragment Lengths on Ramos Blue Cells. DNA from Caco-2 and *Escherichia* were fragmented to **a)** 50-200 bp or **b)** less than 15bp of length, and used to co-activate 0, 1, or 2 µg/mL ODN-2006 at the final concentration of 40 µg/mL. Ramos Blue cells overnight-starved and plated at 200,000 cells/well in starvation medium were used. TLR9 activation was measured as SEAP activity normalized to plate positive control of 5 µg/mL ODN-2006 (results from one experiment with bars representing ± s.d. of duplicates).
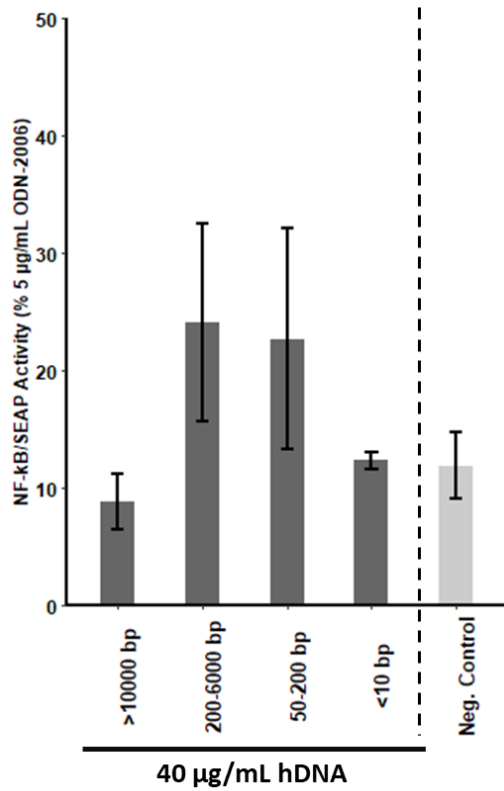
Supplementary Figure S16. Direct TLR9 Activation of Ramos Blue Cells by Human DNA of Various Fragment Lengths. 40 µg/mL human-origin (Caco-2) DNA of different fragment lengths were added to Ramos Blue cells. Cells were overnight-starved and plated at 200,000 cells/well. TLR9 activation was measured as SEAP activity normalized to plate positive control of 5 µg/mL ODN-2006 (mean from 2 independent experiments, bars ± s.e.m.).
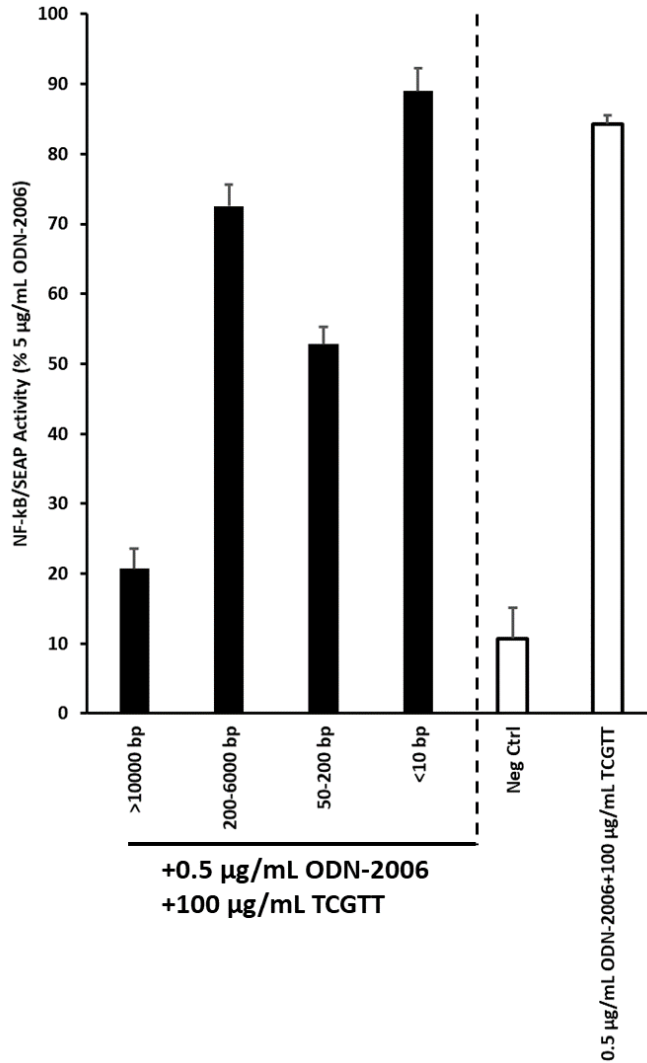
Supplementary Figure S17. Long Human DNA Fragments Repress Low Concentrations of ODN-2006 Co-Activated by TCGTT. 0.5 µg/mL TLR9 activation by ODN-2006 was co-activated with 100 µg/mL short DNA and 40 µg/mL Caco-2 DNA of different fragment lengths were added. TLR9 activation of TCGTT-co-activated Ramos Blue cells was repressed by long (> 10,000 bp) fragments of Caco-2 DNA. Results shown were from 1 experiment performed on overnight-starved Ramos Blue cells plated at 200,000 cells/well. TLR9 activation was measured as SEAP activity normalized to plate positive control of 5 µg/mL ODN-2006. Results of one experiment with bars representing ± s.d. of duplicates.
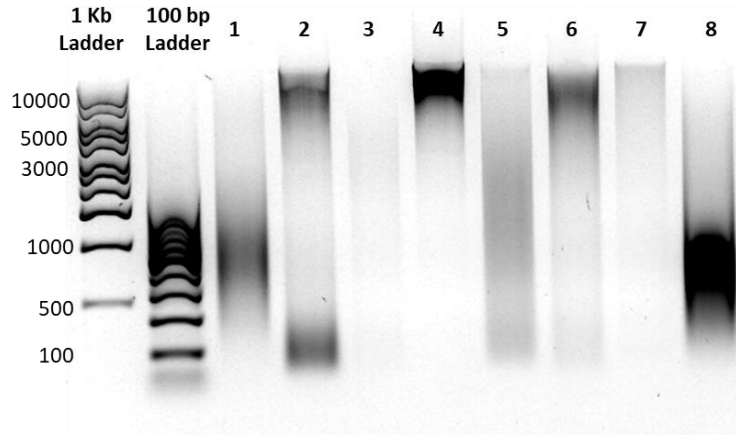
Supplementary Figure S18. Metagenomic DNA Extracted from Infant Stool Samples from Olga Hospital. Quality of DNA extracted from infant stool samples showed large variation, ranging from over 10,000 bp to less than 100 bp of average fragment length. All DNA were extracted via Zymo Quick-DNA Fecal/Soil Microbe Miniprep Kit and loaded onto a 0.8% gel.
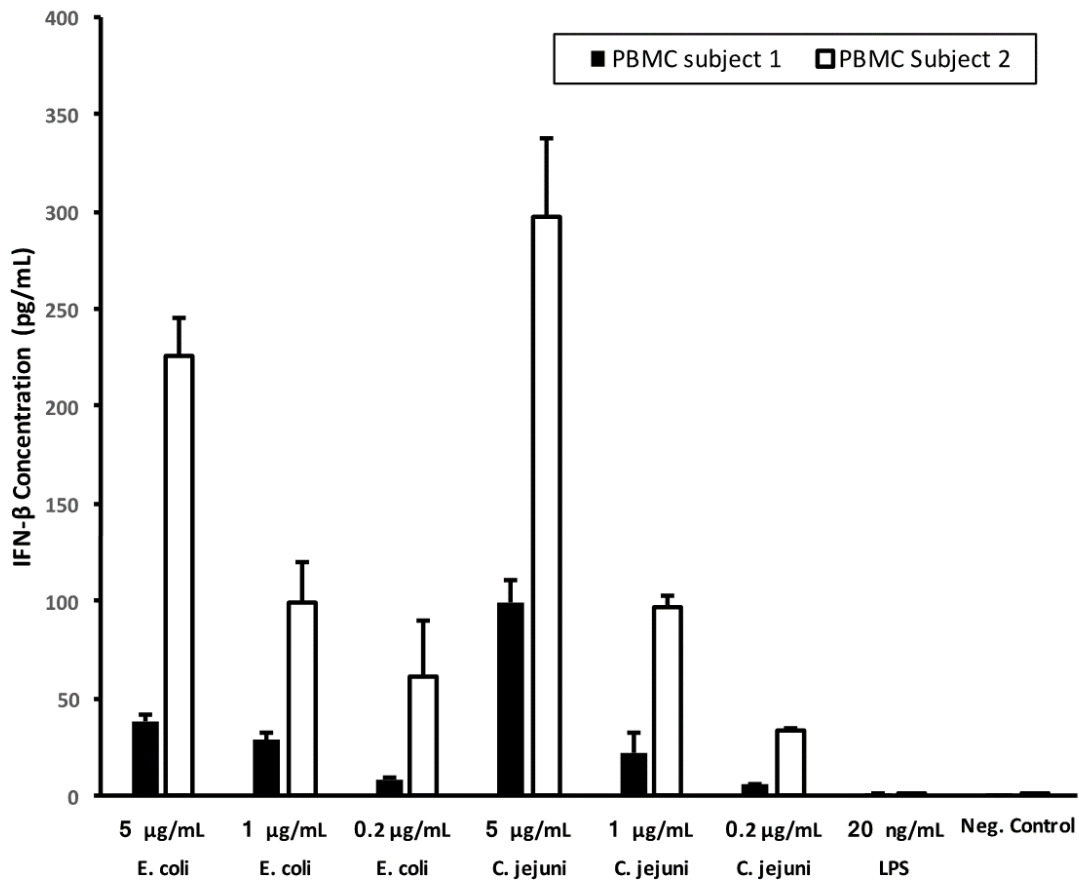
Supplementary Figure S19. IFN-β Production of PBMCs from Two Different Individuals Stimulated by Bacterial DNA. Human PBMCs extracted from two healthy young adults were stimulated with DNA fragments from *Escherichia coli* and *Campylobacter jejuni* in three different concentrations. While results are consistent by species and input concentration, the PBMC extracted from subject 2 produced circa 5 times as much IFN-β compared to those from subject 1.

7. Appendix

7.1 Materials

Table 3. Equipment

Equipment	Manufacturer
DG250 Anaerobic Workstation	Don Whitley Scientific, Bingley, West Yorkshire, UK
Herasafe 2030i Laminar Flow Hood (Class II, Type A2 Biosafety Cabinet)	Thermo Fisher Scientific, Walthem, MA, USA
Herasafe 2025 Laminer Flow Hood (Class II, Type A2 Biosafety Cabinet)	Thermo Fisher Scientific, Walthem, MA, USA
Galaxy® 170S CO_2 Incubator	Eppendorf, Hamburg, Germany
Heracell™ 150i CO_2 Incubator, 150L	Thermo Fisher Scientific, Walthem, MA, USA
INCU-line IL 159R Incubator	VWR, Bruchsal, Germany
HT Ecotron Shaking Incubator	INFORS, Bottmingen, Switzerland
GFL Hydra H20 S Water Bath	LAUDA, Lauda-Königshofen, Germany
AccuBlock™ Digital Dry Baths	Labnet International, Edison, NJ, USA
TSX-600V Ultra Low Temperature Freezer	Thermo Fisher Scientific, Walthem, MA, USA
Oxoid™ AnaeroJar 2.5L	Thermo Fisher Scientific, Walthem, MA, USA
GENBox Anaerobic Box 2.5L	Biomerieux, Marcy l'Etoile, France
Panthera E2 Binocular Microscope	Motic, Barcelona, Spain
Megafuge™ 16R Centrifuge	Thermo Fisher Scientific, Walthem, MA, USA
TX-200 Swinging Bucket Rotor	Thermo Fisher Scientific, Walthem, MA, USA
5430R Refrigerated Centrifuge	Eppendorf, Hamburg, Germany
Rotor FA-23-48-11 Rotor with Aerosol-Tight Caps, to Hold 48 x 1.5/2.0 mL Tubes	Eppendorf, Hamburg, Germany
Rotor FA-45-16-17 Rotor with Aerosol-Tight Caps, to Hold 16 x 5.0 mL Tubes	Eppendorf, Hamburg, Germany
A-2-MTP 2-Plate Swing Bucket Rotor	Eppendorf, Hamburg, Germany
5810R Refrigerated Centrifuge	Eppendorf, Hamburg, Germany
A-2-DWP-AT Rotor with Aerosol-Tight Caps for Plates	Eppendorf, Hamburg, Germany
Sprout® Plus Minicentrifuge	Heathrow Scientific, Vernon Hills, IL, USA
Prism™ Air-Cooled Microcentrifuge	Labnet International, Edison, NJ, USA
Vortex-Genie 2	Scientific Industries, Inc., Bohemia, NY, USA
Horizontal-(24) Microtube Holder (SI-H524)	Scientific Industries, Inc., Bohemia, NY, USA
PTR-35 360° Vertical Multi-function Rotator	Grant-bio, Amsterdam, Netherlands
Enspire® Multimode Plate Reader	PerkinElmer Inc., Walthem, MA, USA
Owl™ EasyCast™ B2 Gel Chamber	Thermo Fisher Scientific, Walthem, MA, USA
Peqlab PerfectBlue™ Horizontal Midi Gel Systems	VWR, Bruchsal, Germany
Power Supply Standard	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Mini-PROTEAN® Tetra Handcast Systems (PAGE Gel Chamber with Supplies)	Bio-Rad, Hercules, CA, USA
Life Technologies E-Gel Imager	Thermo Fisher Scientific, Walthem, MA, USA

Bio-Rad Chemi-Doc XRS+ Gel Imaging System	Bio-Rad, Hercules, CA, USA
Agilent 2200 TapeStation System with Genomic DNA ScreenTape (Part Number 3067-5365) and Genomic DNA Reagents (Part Number 5067-5366)	Agilent Technologies Inc., Santa Clara, CA, USA
S220 Focused-Ultrasonicator	Covaris Inc., Woburn, MA, USA
Holder miniTUBE	Covaris Inc., Woburn, MA, USA
g-TUBE	Covaris Inc., Woburn, MA, USA
MagneSphere® Technology Magnetic Separation Stand for 1.5mL Tubes	Promega, Madison, WI, USA
NanoDrop™ One/One^C Microvolume UV/VIS Spectrophotometer with WLAN	Thermo Fisher Scientific, Walthem, MA, USA
FiveEasy Plus pH Meter FP20	Mettler Toledo, Columbus, OH, USA
Analytical Balance ABS-N_ABJ-NM_ACS_ACJ	Kern & Sohn GMBH, Stuttgart, Germany
C-MAG HS 7 Digital Magnetic Stirrer	IKA, Staufen im Breisgau, Germany

Table 4. Bacteria and Human Cell Culture

Cell Type	Source
Micrococcus luteus (DSM 1790)	Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Culture GMBH (DSMZ), Braunschweig, Germany
Bifidobacterium bifidum (DSM 20239)	DSMZ, Braunschweig, Germany
Bifidobacterium breve (DSM 20213)	DSMZ, Braunschweig, Germany
Lactobacillus ruminis (DSM 20403)	DSMZ, Braunschweig, Germany
Bacteroides dorei (DSM 17855)	DSMZ, Braunschweig, Germany
Enterococcus faecalis (DSM 20478)	DSMZ, Braunschweig, Germany
Staphylococcus epidermidis (DSM 20044)	DSMZ, Braunschweig, Germany
Lactobacillus salivatius (DSM 20555)	DSMZ, Braunschweig, Germany
Clostridioides difficile (DSM 1296)	DSMZ, Braunschweig, Germany
Clostridium perfringens (DSM 756)	DSMZ, Braunschweig, Germany
Fusobacterium nucleatum (DSM 15643)	DSMZ, Braunschweig, Germany
Escherichia coli K12/C600	Department of Food Biotechnology, University of Hohenheim
Campylobacter jejuni (DSM 4688)	Department of Food Biotechnology, University of Hohenheim
Escherichia coli DH5α	Department of Immunology, University of Hohenheim
HEK-Dual™ Human TLR9 (NF-kB-SEAP/KI-[IL8]Lucia) Dual-Reporter HEK293 Cells	InvivoGen Europe, Toulouse, France
HEK-Dual™ Null Dual NF-kB and IL-8 Reporter HEK293 Cells	InvivoGen Europe, Toulouse, France
Ramos-Blue™ NF-kB/AP-1 Reporter B Lymphocytes	InvivoGen Europe, Toulouse, France
Caco-2 Cells (Colon Carcinoma Cells)	Department of Biochemistry, University of Hohenheim
HT-29 Cells (Colon Carcinoma Cells)	Department of Immunology, University of Hohenheim
Peripheral Blood Mononuclear Cells (PBMCs)	Healthy Young Adults

Table 5. Bacteria and Human Cell Culture Reagents

Material	Manufacturer
MRS Broth	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
LB Medium	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Tryptone/Casein Peptone	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Yeast Extract	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Meat Extract	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Agar-Agar, Bacteriological, for Bacterial Culture	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Columbia CNA Agar with 5% Sheep Blood	Biomerieux, Marcy l'Etoile, France
Brain-Heart Infusion Broth	Merck KGaA, Darmstadt, Germany
Vitamin K_1	Merck KGaA, Darmstadt, Germany
Glycerol	Merck KGaA, Darmstadt, Germany
Difco Columbia Broth	Benton, Dickinson and Company, Franklin Lakes, NJ, USA
L-Cysteine-HCl	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Calcium chloride dihydrate ($CaCl_2 \cdot 2H_2O$)	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Magnesium sulfate heptahydrate ($MgSO_4 \cdot 7H_2O$)	Merck KGaA, Darmstadt, Germany
di-Potassium phosphate (K_2HPO_4)	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Potassium dihydrogen phosphate (KH_2PO_4)	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Sodium hydrogen carbonate ($NaHCO_2$)	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
LLC Hemin	MP Biomedicals, Santa Ana, CA, USA
Skim Milk (for Western Blot)	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Oxoid Anaerogen Gas Pack for 2.5L Jar	Thermo Fisher Scientific, Walthem, MA, USA
Oxoid Campygen Gas Pack for 2.5L Jar	Thermo Fisher Scientific, Walthem, MA, USA
Anaerobic Gas Pack for 2.5L Box	Biomerieux, Marcy l'Etoile, France
Microaerophilic Gas Pack for 2.5L Box	Biomerieux, Marcy l'Etoile, France
Sodium Chloride (NaCl)	Labochem International, Heidelberg, Germany
Gibco Delbucco's Modified Eagle Medium 1x with 4.5f/L D-Glucose and L-Glutamine (DMEM)	Thermo Fisher Scientific, Walthem, MA, USA
Gibco Iscove's Modified Delbucco Medium 1x with L-Glutamine and 25mM HEPES (IMDM)	Thermo Fisher Scientific, Walthem, MA, USA
Gibco RPMI 1640 Medium with L-Glutamine	Thermo Fisher Scientific, Walthem, MA, USA
Fecal Bovine Serum (FCS)	Thermo Fisher Scientific, Walthem, MA, USA
Bovine Serum Albumin (BSA)	Merck KGaA, Darmstadt, Germany
Penicillin/Streptomycin	Biochrom Ltd., Cambridge, UK
Normocin	InvivoGen Europe, Toulouse, France
Zeocin	InvivoGen Europe, Toulouse, France
Hyglomycin B Gold	InvivoGen Europe, Toulouse, France
Staurosporine	Merck KGaA, Darmstadt, Germany
Gentamicin (50mg/mL)	Merck KGaA, Darmstadt, Germany
Metronidazole	Alfa Aesar (Thermo Fisher Scientific, Walthem, MA, USA)
Cefoxitin	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Saponin	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Delbucco's Phosphate Buffer Saline (DPBS) (1x)	Thermo Fisher Scientific, Walthem, MA, USA
Gibco 0.05% Trypsin/EDTA (1x)	Thermo Fisher Scientific, Walthem, MA, USA
Gibco Non-Essential Amino Acids	Thermo Fisher Scientific, Walthem, MA, USA
Histopaque-1077	Merck KGaA, Darmstadt, Germany
Dimethyl Sulfoxide (DMSO)	Carl Roth GMBH & Co. KG, Karlsruhe, Germany

Table 6. Materials for DNA Extraction, Fragmentation, Clean-up, and Gel Electrophoresis

Material	Manufacturer
Quick-DNA™ Fungal/Bacterial Miniprep Kit	Zymo Research Corporation, Irvine, CA, USA
Quick-DNA™ Fecal/Soil Microbe Miniprep Kit	Zymo Research Corporation, Irvine, CA, USA
DNeasy Blood & Tissue Kit	QIAGEN, Hilden, Germany
Blood & Cell Culture DNA Maxi Kit	QIAGEN, Hilden, Germany
MiraCLEAN Endotoxin Removal Kit	Mirus Bio LLC, Madison, WI, USA
NEBNext® dsDNA Fragmentase®	New England Biolabs, Ipswich, MA, USA
Blue miniTUBE for 3Kb, used with Covaris s220 Sonicator and miniTUBE holder	Covaris Inc., Woburn, MA, USA
Invitrogen™ DNA-*free* DNA Removal Kit	Thermo Fisher Scientific, Walthem, MA, USA
DNA Clean & Concentrator-25® Kit	Zymo Research Corporation, Irvine, CA, USA
Oligo Clean & Concentrator™ Kit	Zymo Research Corporation, Irvine, CA, USA
Select-a-Size DNA Clean & Concentrator™ Magbead Kit	Zymo Research Corporation, Irvine, CA, USA
SPRIselect Magnetic Beads	Beckman Coulter Inc., Brea, CA, USA
GeneJET Gel Extraction Kit	Thermo Fisher Scientific, Walthem, MA, USA
Ribonuclease A (RNase A)	QIAGEN, Hilden, Germany
Proteinase K	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Sodium Acetate	Amnesco LLC, Solon, OH, USA
Nuclease-Free Water	Thermo Fisher Scientific, Walthem, MA, USA
Agarose Broad Range	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Acrylamide: Bisacrylamide	Merck KGaA, Darmstadt, Germany
Tris Base	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Boric Acid	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Sodium dodecyl sulfate (SDS)	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Ammonium persulfate (APS)	Merck KGaA, Darmstadt, Germany
Tetramethylethylenediamine (TEMED)	Merck KGaA, Darmstadt, Germany
50x TAE	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
ROTI-Stain	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
GelStar Nucleic Acid Gel Stain	Lonza Bioscience, Basal, Switzerland
6x Sample Buffer	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Lambda Hind III Ladder	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
1 Kb Ladder	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
200 bp Ladder	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
100 bp Ladder	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
Ultra-Low-Range Ladder	Thermo Fisher Scientific, Walthem, MA, USA

Table 7. Synthetic Oligonucleotides

Oligonucleotide name	Sequence†	Manufacturer
ODN-2006 (Class B ODN)	5'-tcgtcgttttgtcgttttgtcgtt-3'	InvivoGen Europe, Toulouse, France
ODN-2006GC (Negative Control of ODN-2006)	5'-tgctgcttttgtgcttttgtgctt-3'	InvivoGen Europe, Toulouse, France
ODN-2216 (Class A ODN)	5'-ggGG<u>GACGA:TCGTC</u>ggggggg-3'	InvivoGen Europe, Toulouse, France
ODN-2216GC (Negative Control of ODN-2216)	5'-ggGG<u>GAGCA:TGCTC</u>ggggggg-3'	InvivoGen Europe, Toulouse, France

ODN INH-18 (Inhibitory ODN)	5'-cctgaatgggaacttaccgctgca-3'	InvivoGen Europe, Toulouse, France
TCGTT	5'-TCGTT-3'	Eurofins Genomics, Ebersberg, Germany
TTTTT	5'-TTTTT-3'	Eurofins Genomics, Ebersberg, Germany

† Bases with phosphodiester backbone are denoted by capital letters and bases with phosphonothioate backbone by small letters. Palindromic sequences are indicated with an underline with colon indicating the location of the stem loop.

Table 8. Other Reagents for TLR9 Activation Experiments

Material	Manufacturer
DOTAP	Carl Roth GMBH & Co. KG, Karlsruhe, Germany
LPS from *Escherichia Coli* K12	InvivoGen Europe, Toulouse, France
Quanti-Blue Solution	InvivoGen Europe, Toulouse, France
Quanti-Luc	InvivoGen Europe, Toulouse, France
Human Interferon-beta ELISA Kit	R & D Systems, Minneapolis, MI, USA
Human Interleukin-8 (IL-8) ELISA Kit	R & D Systems, Minneapolis, MI, USA
Human Interleukin-6 (IL-6) ELISA Kit	R & D Systems, Minneapolis, MI, USA

Table 9. Hardware for Data Analysis

Hardware Type	Description
Apple Mac Mini i5	CPU: Intel Core i5 2.6 GHz Memory: 4GB 1067 MHz DDR3 Model: A1347 (EMC 2840) Operating System: OS X El Capitan Version: 10.11.5 (15F34)
Acer Aspire A317-51G	CPU: Intel® Core™ i7-10510U CPU@ 1.80 GHz 2.30 GHz Memory: 8GB System Type: 64-Bit Operating System: Windows 11 Home Version: 21H2

Table 10. Software for Image Capture and Data Analysis

Software	Origin
E.Gel® Imager Software Package	Thermo Fisher Scientific, Walthem, MA, USA
Bio-Rad Image Lab Software	Bio-Rad, Hercules, CA, USA
Microsoft® Excel for Mac	Microsoft Corperation, Redmond, USA
Agilent 2200 TapeStation Software Revision A.02.02 (SR1)	Agilent Technologies Inc., Santa Clara, CA, USA
R	R Core Development team (2011)
RStudio® Version v2021.09.0+351	RStudio Team (RStudio 2021)
R Packages	
/tidyverse	H. Wickham et al. (2019)
/scales	H. Wickham et al. (2016)

7.2. Additional Protocol Links

Agilent Technologies

Agilent 2200 Tape Station:
https://www.agilent.com/cs/library/catalogs/public/Catalog-bioanalyzer-tapestation-systems-sw-consumables-5994-0249EN_agilent.pdf

Beckman Coulter

SPRIselect Beads:
https://www.beckman.de/reagents/genomic/cleanup-and-size-selection/size-selection/performance

Covaris Inc.

g-Tube:
https://www.covaris.com/wp-content/uploads/pn_010154.pdf
s220 Ultrasonicator:
https://www.covaris.com/wp-content/uploads/pn_010122.pdf

InvivoGen

QUANTI-Blue Solution:
https://www.invivogen.com/sites/default/files/invivogen/products/files/quanti_blue_solution_tds.pdf
QUANTI-Luc:
https://www.invivogen.com/sites/default/files/invivogen/products/files/quanti_luc_tds.pdf

Mirus Bio LLC

MiraCLEAN® Endotoxin Removal Kit:
http://www.mirusbio.com/assets/protocols/ml14_miraclean_endotoxin_removal_kit.pdf

QIAGEN

DNeasy Blood and Tissue Kit:
http://www.qiagen.com/us/products/discovery-and-translational-research/dna-rna-purification/dna-purification/genomic-dna/dneasy-blood-and-tissue-kit/

R+D Systems

Human IFN-beta DuoSet ELISA kit:
https://resources.rndsystems.com/pdfs/datasheets/dy814.pdf?v=20220713&_ga=2.84770393.1774167290.1657729662-614719184.1657729662
Human IL-6 DuoSet ELISA kit:
https://resources.rndsystems.com/pdfs/datasheets/dy206.pdf?v=20220713&_ga=2.83194198.1774167290.1657729662-614719184.1657729662

Human IL-8 DuoSet ELISA kit

https://resources.rndsystems.com/pdfs/datasheets/dy208.pdf?v=20220713&_ga=2.83194198.1774167290.1657729662-614719184.1657729662

Thermo Fisher Scientific

Ambion DNA-Free^{TM} DNA Removal Kit:
https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets%2FLSG%2Fmanuals%2Fcms_055739.pdf
GeneJET Gel Extraction Kit:
https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets%2FLSG%2Fmanuals%2FMAN0012661_GeneJET_Gel_Extraction_UG.pdf
Oxoid AnaeroJar 2.5L with 2.5L AnaeroGen/CampyGen:
http://www.oxoid.com/UK/blue/prod_detail/prod_detail.asp?pr=AG0025&org=154&c=UK&lang=EN

Biomerieux

GENbox Anaerobic Sachets:
https://biomerieuxdirect.com/clinical/Bacteriology/Culture/Ancillary-Products/Others/Gas-generators/GENBOX-ANAER-10-SACHETS/p/96124
GENbox Microaerophilic Sachets:
https://biomerieuxdirect.com/clinical/Bacteriology/Culture/Ancillary-Products/Others/Gas-generators/GENBOX-MICROAER-10-SACHETS/p/96125

Zymo Research

DNA Clean & Concentrator-25® Kit:
http://files.zymoresearch.com/protocols/_d4005_d4006__d4033_d4034_dna_clean_concentrator-25_kit.pdf
Oligo Clean & Concentrator^{TM}:
http://files.zymoresearch.com/protocols/_d4060_d4061_oligo_clean_concentrator.pdf
Quick-DNA^{TM} Fecal/Soil Microbe Miniprep Kit:
http://files.zymoresearch.com/protocols/_d6010_quick-dna_fecalsoil_microbe_miniprep_kit.pdf
Quick-DNA^{TM} Fungal/Bacterial Miniprep Kit:
http://files.zymoresearch.com/protocols/_d6005_quick-dna_fungal-bacterial_miniprep_kit.pdf
Select-a-Size DNA Clean & Concentrator^{TM} MagBead Kit:
http://files.zymoresearch.com/protocols/_d4084_d4085_select-a-size_dna_clean_concentrator_magbead_kit.pdf

8. Bibliography

1. R. Medzhitov, Toll-like receptors and innate immunity. *Nature Rev. Immunol.* **1**, 135-145 (2001)
2. K. A. Fitzgerald, J. C. Kagan, Toll-like receptors and the control of immunity. *Cell* **180**, 1044-1066 (2020)
3. S. Iwanaga, B. L. Lee, Recent advances in the innate immunity of invertebrate animals, *J. Biochem. and Mol. Biol.* **38**(2), 128-150 (2005)
4. K. Buchmann, Evolution of innate immunity: clues from invertebrates via fish to mammals. *Front. Immunol* .**5**: 459 (2014)
5. M. G. Netea, A. Schlitzer, K. Placek, L. A. B. Joosten, J. L. Schultze, Innate and adaptive immune memory: an evolutionary continuum in the host's response to pathogens. *Cell Host and Microbe*, **25**, 13-26 (2019)
6. J. A. Hoffman, J-M. Reichhart, *Drosophila* innate immunity: an evolutionary perspective. *Nat. Immunol.* **3**(2), 121-126 (2002)
7. J. Kasamatsu, Evolution of innate and adaptive immune systems in jawless vertebrates. *Microbiol. Immunol.* **57**(1), 1-12 (2013)
8. K. Kamm, B. Schierwater, R. DeSalle, Innate immunity in the simplest animals—placozoans. *BMC Genomics* **20**(1), 5 (2019)
9. M. R. Romo, D. Pérez-Martínez, C. C. Ferrer, Innate immunity in vertebrates: an overview. *Immunology* **148**(2), 125-139 (2016)
10. J.S. Marshall, R. Warrington, W. Watson, H. L. Kim, An introduction to immunology and immunopathology. *BMC Allergy, Asthma, & Clin. Immunol.* **14**: 49 (2018)
11. A. Le Bon, D. F. Tough, Links between innate and adaptive immunity via type I interferon. *Curr. Opin. in Immunol.* **14**(4), 432-436 (2002)
12. K. Hoebe, E. Janssen, B. Beutler, The interface between innate and adaptive immunity. *Nat. Immunol.* **5** (10), 971-974 (2014)
13. B. C. Braden, R. J. Poljak, Structural features of the reactions between antibodies and protein antigens, *FASEB J.* **9**(1), 9-16 (1995)
14. I. Lerouge, J. Vanderleyden, O-antigen structural variation: mechanisms and possible roles in animal/plant-microbe interactions. *FEMS Microbiol. Rev.*, **26**(1), 17-47 (2002)
15. H. Okude, D. Ori, T. Kawai, Signaling through nucleic acid sensors and their roles in inflammatory diseases. *Front. Immunol.* **11:**625833 (2021)
16. D. Li, M. Wu, Pattern recognition receptors in health and disease. *Signal Transduction and Targeted Therapy* **6:**291 (2021)
17. O. Takeuchi, S. Akira, Pattern recognition receptors and inflammation, *Cell* **140**(6), 805-820 (2010)
18. A. M. Piccinini, K. S. Midwood, DAMPening inflammation by modulating TLR signaling. *Mediators of Inflammation* 672395 (2010)
19. J.W. Coleman, Nitric oxide in immunity and inflammation. *Int. Immunopharmacology* **1**(8), 1397-1406 (2001)
20. D. A. Wink, H. B. Hines, R. Y. S. Cheng, C. H. Switzer, W. Flores-Santana, M. P. Vitek, L. A. Ridnour, C. A. Colton, Nitric oxide and redox mechanisms in the immune response. *J. Leuko. Biol.* **89**(6), 873-891 (2011)
21. R. Suresh, D.M. Mosser, Pattern recognition receptors in innate immunity, host defense, and immunopathology. *Adv. Physiol. Edu.* **37**(4), 284-291 (2013)
22. M.J. Hickey, P. Kubes, Intravascular immunity: the host-pathogen encounter in blood vessels. *Nat. Rev. Immunol.* **9**, 364-375 (2009)
23. F. M. Ausubel, Are innate immune signaling pathways in plants and animals conserved? *Nat. Immunol.* **6**, 973-979 (2005)
24. L. A. J. O'Neill, D. Golenbock, A. G. Bowie, The history of the Toll-like receptors—redefining innate immunity. *Nat. Rev. Immunol.* **13**(6), 453-460 (2013)

25. J. J. Brennan, T. D. Gilmore. Evolutionary origins of Toll-like receptor signaling. *Mol. Biol. Evol.* **35**(7), 1576-1587 (2018)

26. K. V. Andreson, C. Nüsslein-Volhard, Establishment of dorsal-ventral polarity in the Drosophila embryo: the induction of polarity by the Toll gene product. *Cell* **423**, 791–798 (1985)

27. B. Lemaitre, E. Nicolas, L. Michaut, J.-M. Reichhart, J. A. Hoffmann. The dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. *Cell* **86**, 973-983 (1996)

28. R. Medzhitov, P. Preston-Hurlburt, C. A. Janeway Jr. A human homologue of the Drosophila Toll protein signals activation of adaptive immunity. *Nature*, **388**, 394-397 (1997)

29. J. Y. Kang, J.-O. Lee. Structural biology of the Toll-like receptor family. *Annu. Rev. Biochem.* **80**, 917-941 (2011)

30. M. S. Jin, J.-O. Lee, Structures of the Toll-like receptor family and its ligand complexes. *Immunity* **29**(2), 182-191 (2008)

31. U. Ohto, T. Shibata, H. Tanji, H. Ishida, E. Krayukhina, S. Uchiyama, K. Miyake, T. Shimizu. Structural basis of CpG and inhibitory DNA recognition by Toll-like receptor 9. *Nature*, **520**: 7549, 702-205 (2015)

32. J. C. Roach, G. Glusman, L. Rowan, A. Kaur, M. K. Purcell, K. D. Smith, L. E. Hood, A. Aderem, The evolution of vertebrate Toll-like receptors. *PNAS* **102**(27), 9577-9582 (2005)

33. C. A. Thaiss, N. Zmora, M. Levy, E. Elinav, The microbiome and innate immunity, *Nature* **535**, 65-74 (2016)

34. T. Kawasaki, T. Kawai, Toll-like receptor signaling pathways. *Front. Immunol.* **5**: 461 (2014)

35. A. L. Blasius, B. Beutler, Intracellular Toll-like receptors. *Immunity* **32**(3), 305-315 (2010)

36. D. M. Klinman, A. K. Yi, S. L. Beaucage, J. Conover, A. M. Krieg, CpG motifs present in bacteria DNA rapidly induce lymphocytes to secrete interleukin 6, interleukin 12, and interferon gamma. *PNAS* **93**(7), 2879-2883 (1996)

37. C. A. Leifer, M. N. Kennedy, A. Mazzoni, C. Lee, M. J. Kruhlak, D. M. Segal, TLR9 is localized in the endoplasmic reticulum prior to stimulation. *J. Immunol.* **173**, 1179-1183 (2004)

38. B. L. Lee, G. M. Barton, Trafficking of endosomal Toll-like receptors. *Trends Cell Biol.* **24**(6), 360-369 (2014)

39. A. Chockalingam, J. C. Brooks, J. L. Cameron, L. K. Blum, C. A. Leifer, TLR9 traffics through the Golgi complex to localize to endolysosomes and respond to CpG DNA. *Immunol. Cell Biol.* **87**(3), 209-217 (2009)

40. G. M. Barton, J. C. Kagan, R. Medzhitov, Intracellular localization of Toll-like receptor 9 prevents recognition of self DNA but facilitates access to viral DNA. *Nat. Immunol.* **7**(1), 49-56 (2006)

41. M. Schlee, G. Harmann, Discriminating self from non-self in nucleic acid sensing, *Nat. Rev. Immunol.* **16**, 566-580 (2016)

42. T. Kawai, S. Akira, TLR signaling. *Cell Death & Differentiation* **13**, 816-825 (2006)

43. S. Uematsu, S. Akira, Toll-like receptors and Type I interferons, *J. Biol. Chem.* **282**(21), 15319-15323 (2007)

44. R. Hershberg, D. A. Petrov, Evidence that mutation if universally biased toward AT in bacteria. *PLOS Genetics* **6**, e10001115 (2010)

45. W.-H. Chen, G. Lu, P. Bork, S. Hu, M. J. Lercher, Energy efficiency trade-offs drive nucleotide usage in transcribed regions. *Nature Communications* **7**:11334, 1-10 (2016)

46. D. Podlesny, In silico prediction of TLR9-dependent immune modulation by genomic and metagenomic DNA. [master thesis]. University of Hohenheim (2016)

47. K. J. Stacey, G. R. Young, F. Clark, D. P. Sester, T. L. Roberts, S. Naik, M. J. Sweet, D. A. Hume, The molecular basis for the lack of immunostimulatory activity of vertebrate DNA. *J. Immunol.* **170**, 3614-3620 (2003)

48. V. Kamuju, S. Kumar, W. H. Khan, P. Vivekanandan, Hypervirulent *Clostridium difficile* ribotypes are CpG depleted. *Virulence*, **9**(1), 1422-1425 (2018)

49. K. Hoelzer, L. A. Shackelton, C. R. Parrish, Presence and role of cytosine methylation in DNA viruses in animals. *Nucleic Acid Res.* **36**(9), 2825-2837 (2008)

50. A. Dalpke, J. Frank, M. Peter, K. Heeg, Activation of Toll-like receptor 9 by DNA from different bacterial species. *Infection and Immunity* **74**(2), 940-946 (2006)
51. M. Wojciechowski, H. Czapinska, M. Bochtler, CpG underrepresentation and the bacterial CpG-specific DNA methyltransferase M.Mpel. *PNAS* **110(1)**, 105-110 (2013)
52. T. Senorer, Relative abundance profiles of TLR9-activating sequence motifs distinguish the human fecal microbiota of infants, mothers and adults. [bachelor thesis]. University of Hohenheim, 2017
53. S. Saxonov, P. Berg, D. Brutlag, A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters, *PNAS* **103**(5), 1412-1417 (2006)
54. L. D. Moore, T. Le, Guoping Fan, DNA Methylation and Its Basic Function. *Neuropsychopharmacology* **38**, 23-38 (2013)
55. D. E. Kebir, L. József, W. Pan, L. Wang, J. G. Filep, Bacterial DANN activates endothelial cells and promotes neutrophil adherence through TLR9 signaling. *J. Immunol.* **182**(7), 4386-4394 (2009)
56. A. M. Krieg, CpG motifs in bacterial DNA and their immune effects. *Annu. Rev. Immunol.* **20**, 709-760 (2002)
57. T. Tokunaga, T. Yamamoto, S. Yamamoto, How BCG led to the discovery of immunostimulatory DNA. *Jpn. J. Infect. Dis.* **52**(1), 1-11 (1999)
58. C. F. Mojcik, M. F. Gourley, D. M. Klinman, A. M. Krieg, F. Gmelig-Meyling, A. D. Steinberg, Administration of a phosphonothioate oligonucleotide antisense to musine andogenous retroviral MCF *env* causes immune effects *in vivo* in s sequence-specific manner. *Clin, Immunol. and Immunopathol.* **67**(2), 130-136 (1993)
59. H. Liang, Y. Nishioka, C. F. Reisch, D. S. Pisetsky, P. E. Lipsky, Activation of human B Cells by phosphonothioate oligodeoxynucleotide. *J. Clin. Invest.* **98**(5), 1119-1129 (1996)
60. M. Bauer, K. Heeg, H. Wagner, G. B. Lipford, DNA activates human immune cells through a CpG sequence-dependent manner. *Immunology* **97**(4), 699-705 (1999)
61. G. Hartmann, A. M. Krieg, Mechanism and function of a newly identified CpG DNA motif in human primary B cells. *J. Immunol.* **164**(2), 944-953(2000)
62. A. M. Krieg, Therapeutic potential of Toll-like Receptor 9 activation. *Nat. Rev. Drug Discovery* **5**, 471-484 (2006)
63. J. Pohar, D. Lainšček, K. Ivičak-Kocjan, M.-M. Cajnko, R. Jarala, M. Benčina, Short single-stranded DNA degradation products augment the activation of Toll-like receptor 9. *Nature Communications* **8:**15363, 1-13 (2017)
64. M. Peter, K. Bode, G. B. Lipford, F. Eberle, K. Hegg, A. H. Dalpke. Characterization of suppressive oligodeoxynucleotides that inhibit Toll-like receptor-9-mediated activation of innate immunity. *Immunology* **123**(1), 118-128 (2008)
65. I. Gursel, M. Gursel, H. Yamada, K. J. Ishii, F. Takeshita, D. M. Klinman. Repetitive elements in mammalian telomeres suppress bacterial DNA-induced immune activation. *J. Immunol.* **171**, 1393-1400 (2003)
66. J. S. Riley, S. W. G. Tait, Mitochondrial DNA in inflammation and immunity. *EMBO reports* **21**, e49799 (2020)
67. M. W. Gray, G. Burger, B. F. Lang, The origin and early evolution of mitochondria. *Genome Biology* **2**(6), 1018.1-1018.5 (2001)
68. J.-H. Lai, M.-Y. Wang, C.-Y. Huang, C.-H. Wu, L.-F. Hung, C.-Y. Yang, P.-Y. Ke, S.-F. Luo, S.-J. Liu, L.-J. Ho, Infection with the dengue RNA virus activates TLR9 signaling in human dendritic cells. *EMBO Rep.* **19**(8), e46182 (2018)
69. J.H. Lai, D.-W. Wu, C.-H. Wu, L.-F. Hung, C.-Y. Huang, S.-M. Ka, A. Chen, Z.-F. Chang, L.-J. Ho. Mitochondrial CMPK2 mediates immunomodulatory and antiviral activities through IFN-dependent and IFN-independent pathways. *iScience* **24**(6), 102498 (2021)
70. S. Nishimoto, D. Fukada, Y. Higashikuni, K. Tanaka, Y. Hirata, C. Murata, J.-R. Kim-Kaneyama, F. Sato, M. Bando, S. Yagi, T. Soeki, T. Hayashi, I. Imoto, H. Sakaue. M. Shimabukoro, M. Sata, Obesity-induced DNA released from adipocytes stimulates chronic adipose tissue inflammation and insulin resistance. *Sci. Adv.* **2**, e1501332 (2016)

71. A. B. Imaeda, A. Watanabe, M. A. Sohail, S. Mahmood, M. Mohamadnejad, F. S. Sutterwala, R. A. Flavell, W. Z. Mahal, Acetaminophen-induced hepatotoxicity in mice is dependent on TLR9 and the Nalp3 inflammasome. *The Journal of Clinical Investigation* **119(2)**, 305-314 (2009)

72. M. Haghiac, N. L. Vora, S. Basu, K. L. Johnson, L. Presley, D. W. Bianchi, S. Hauguel-De Mouzon, Increased death of adipose cells, a path to release cell free DNA into systemic circulation of obese women, *Obesity (Silver Spring)* **20**(11), 2213-2219 (2012)

73. K.-L. G. Spindler, A. K. Boysen, N. Pallisgård, J. S. Johansen, J. Tabernero, M. M. Sørensen, B. V. Jensen, T. F. Hansen, D. Sefrioui, R. F. Andersen, I. Brandslund, A. Jakobsen, Cell-free DNA in metastatic colorectal cancer: a systematic review and meta-analysis, *Oncologist* **22**(9), 1049-1055 (2017)

74. E. M. Tan H. G. Kunkel, Characteristics of a soluble nuclear antigen precipitating with sera of patients with systemic lupus erythematosus. *J. Immunol.* **96**, 464-471 (1966)

75. J. Tian, A. M. Avalos, S.-Y. Mao, B. Chen, K. Senthil, H. Wu, P. Parroche, S. Drabic, D. Golenbock, C. Sirois, J. Hua, L. L. An, L. Audoly, G. La Rosa, A. Bierhaus, P. Naworth, A. Marshak-Rothstein, M. K. Brow, K. A. Fitzgerald, E. Latz, P. A. Kiener, A. J. Coyle, Toll-like receptor 9-dependent activation of DNA-containing immune complexes is mediated by HMGB and RAGE. *Nat. Immunol.* **8**(5), 487-496 (2007)

76. A. Moreno-Angarita, C. C. Aragón, G. J. Tobón, Cathelicidin LL-37: A new important molecule in the pathophysiology of systemic lupus erythematosus, *J. Transl. Autoimmun*. **3**:100029 (2019)

77. P. Sumby, K. D. Barbian, D. J. Gardner, A. R. Whitney, D. M. Welty, R. D. Long, J. R. Bailey, M. J. Parnell, N. P. Hoe, G. G. Adams, F. R. DeLeo, J. M. Musser, Extracellular deoxyribonuclease made by group A Streptococcus assists pathogenesis by enhancing evasion of the innate immune response. *PNAS* **102**(5), 1679-1684 (2005)

78. B. Dema, N. Charles, Autoantibodies in SLE: specificities, isotypes and receptors. *Antibodies* **5**(1), 2 (2016)

79. T. K. Means, E. Latz, F. Hayashi, M. R. Murali, D. T. Golenbock, A. D. Luster, Human lupus autoantibody-DNA complexes activate DCs through cooperation of CD32 and TLR9. J. Clin. Invest. 115(2): 407-417 (2005)

80. Ann Marshak-Rothstein, Toll-like receptors in systemic autoimmune disease. *Nat. Rev. Immunol.* **6**, 823-835 (2006)

81. G. M. H. Birchenough, M. E. V. Johansson, J. K. Gustafsson, J. H. Bergström, G. C. Hansson, New developments in goblet cell mucus secretion and function. *Mucosol Immunol.* **8**, 712-719 (2015)

82. M. E. V. Johansson, H. Sjövall, G. C. Hansson, The gastrointestinal mucus system in health and disease. *Nat. Rev. Gastroenterol. Hepatol.* **10**(6), 352-361 (2013)

83. A. Macierzanka, A. R. Mackle, B. H. Bajka, N. M. Rigby, F. Nau, D. Dupont. Transport of particles in intestinal mucus under stimulated infant and adult physiological conditions: impact of mucus structure and extracellular DNA. *PLOS One* **9**(4), e92754 (2014)

84. M. E. V. Johansson, H. E. Jakobsson, J. Holemén-Larsson, A. Schütte, A. Ermund, A. M. Rodríguez-Piñeiro, L. Arike, C. Wising, F. Svensson, F. Bäcked, G. C. Hansson, Normalization of host intestinal mucus layer requires long-term microbial colonization. *Cell Host Microbe* **18(5)**, 589-592 (2015)

85. J. M. Williams, C. A. Duckworth, M. D. Burkitt, A. J. M. Watson, B. J. Campbell, D. M. Pritchard, Epithelial Cell Shedding and Barrier Function, *Vet. Pathol.* **52**(3), 445-455 (2015)

86. A. M. Patterson, A. J. M. Watson, Deciphering the complex signaling systems that regulate intestinal epithelial cell death processes and shedding. *Frontiers of Immunology* **8**:841, 1-7 (2017)

87. J. M. Blander, Death in the intestinal epithelium—basic biology and implications for inflammatory bowel disease. *FEBS Journal* **283**, 2720-2730 (2016)

88. P. H. V. Saavedra, L. Huang, F. Ghazavi, S. Kourula, T. V. Berghe, N. Takahashi, P. Vandenabeele, M. Lamkanfi, Apoptosis of intestinal epithelial cells restricts *Clostridium difficile* infection as a model of pseudomembranous colitis. *Nature Comm.* **9**:4846, 1-10 (2018)

89. L. J. Cliffe, N. E. Humphreys, T. E. Lane, C. S. Potten, C. Booth, R. K. Grencis, Accelerated intestinal epithelial cell turnover: a new mechanism of parasite expulsion. *Science* **308**, 1463-1465 (2005)

90. M. Sachet, Y. Y. Liang, R. Oehler, The immune response to secondary necrotic cells, *Apoptosis* **22**, 1189-1204 (2017)

91. C. Günther, H. Neumann, M. F. Neurath, C. Becker, Apoptosis, necrosis and necroptosis: cell death regulation in the intestinal epithelium. *Gut* **62**, 1062-1071 (2013)

92. S. Mehandru, J.-F. Colombel, The intestinal barrier, an arbitrator turned provocateur in IBD. *Nat. Rev. Gastroenterol. Hepatol.* **18**, 83-84 (2020)

93. A. Nishida, R. Inoue, O. Inatomi, S. Bamba, Y. Naito, A. Andoh, Gut microbiota in the pathogenesis of inflammatory bowel disease. *Clin. J. of Gastroenterology* **11**:1-10 (2017)

94. M. Maronek, B. Gromova, R. Liptak, B. Konecna, M. Pastorek, B. Cechova, M. Harsanyova, J. Budis, D. Smolak, J, Radvanszky, T. Szemes, J. Harsanyiova, A. Kralova Trancikova, R. Gardlik, Extracellular DNA correlates with intestinal inflammation in chemically induced colitis in mice. *Cells* **10**: 81: 1-19 (2021)

95. K. N. Couper, D. G. Blount, E. A. Riley, IL-10: the master regulator of immunity to infection. *J. Immunol.* **180**(9), 5771-5777 (2008)

96. S. C. Gribar, C. P. Sodhi, W. M. Richardson, R. J. Anand, G. K. Gittes, M. F. Branca, A. Jakub, X.-H. Shi, S. Shah, J. A. Ozolek, D. J. Hackam, Reciprocal expression and signaling of TLR4 and TLR9 in the pathogenesis and treatment of necorotizing entercolitis. *J. Immunol.* **182**, 636-646 (2009)

97. J. Arciero, G. B. Ermentrout, R. Siggers, A. Afrazi, D. Hackam, Y. Vodovotz, J. Rubin, Modeling the interaction of bacteria and Toll-like receptor-mediated inflammation in necrotizing entercolitis. *Journal of Theoretical Biology* **321**, 83-99 (2013)

98. J. Henao-Mejia, E. Elinav, C. Jin, L. Hao, W. Z. Mehal, T. Strowig, C. A. Thaiss, A. L. Kau, S. C. Eisenbarth, M. J. Jurczak, J.-P. Camporez, G. I. Shulman, J. I. Gordon, H. M. Hoffman, R. A. Flavell, Inflammasome-mediated dysbiosis regulates progression of NAFLD and obesity. *Nature* **482**, 179-186 (2012)

99. M. Good, C. P. Sodhi, J. A. Ozolek, R. H. Buck, K. C. Goehring, D. L. Thomas, A. Vikram, K. Biddy, M. J. Morowitz, B. Firek, P. Lu, D. J. Hackam, *Lactobacillus rhamnosus* HN001 decreases the severity of necrotizing enterocolitis in neonatal mice and preterm piglets: evidence in mice for a role of TLR9. *Am. J. Physiol. Gastrointest. Liver Physiol.* **306**, G1021-G1032 (2013)

100. S. R. Christensen, J. Shupe, K. Nickerson, M. Kashgarian, R. A. Flavell, M. J. Shlomchik, Toll-like receptor 7 and TLR9 dictate autoantibody specificity and have opposing inflammatory and regulatory roles in a murine model of lupus. *Immunity* **25**(3), 417-428 (2006)

101. R. Sender, S. Fuchs, R. Milo, Are we really vastly outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans, *Cell*, **164**(3), 337-340 (2016)

102. C. A. Brennan, W. G. Garrett, *Fusobacterium nucleatum*—symbiont, opportunist and oncobacterium. *Nat. Rev. Microbiol.* **17**, 156-166 (2018)

103. https://www.dsmz.de/microorganisms/medium/pdf/DSMZ_Medium58.pdf [German Collection of Microorganism and Cell Cultures GmBH (DSMZ), Leibniz Institute, Braunschweig, Germany]

104. Graham FL, Smiley J, Russell WC, Nairn R, Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *The Journal of General Virology* **36** (1): 59–74 (1977)

105. https://www.invivogen.com/hek-dual-htlr9 [InvivoGen Inc., Human TLR9 Dual-Reporter HEK293 Cells]

106. https://www.invivogen.com/hek-dual-null [InvivoGen Inc., Hek-Dual Null™ Cells]

107. G. Klein, B. Giovanella, A. Westman, J. S. Stehlin, D. Mumford, An EBV-genome-negative cell line established from an American Burkitt Lymphoma; receptor characteristics. EBV infectibility and permanent conversion into EBV-positive sublines by *in vitro* infection. *Intervirology* **5**, 319-334 (1975)

108. https://www.invivogen.com/ramos-blue-cell [InvivoGen Inc., Ramos-Blue™ Cells]

109. D. Ahmed, P. W. Eide, I. A. Eilertsen, S. A. Danielsen, M. Eknæs, M. Hektoen, G. E. Lind, R. A. Lothe, Epigenetic and genetic features of 24 colon cancer cell lines. *Oncogenesis* **2**: e71, 1-8 (2013)

110. M Rousset, The human colon carcinoma cell lines HT-29 and Caco-2: two in vitro models for the study of intestinal differentiation. *Biochemie* **68**, 1035-1040 (1986)

111. G. C. Hansson, Mucus and mucins in diseases of the intestinal and respiratory tracts. *J. Internal Medicine* **285**(5), 479-490 (2019)

112. W. E. Kloos, M. S. Musselwhite, Distribution and persistence of *Staphylococcus* and *Micrococcus* species and other aerobic bacteria on human skin. *Applied Microbiology* **30**(3), 381-395 (1975)

113. Y. W. Han, W. Shi, G. T.-J. Huang, S. K. Haake, N.-H. Park, H. Kuramitsu, R. Genco, Interactions between periodontal bacteria and human oral epithelial cells: *Fusobacterium nucleatum* adheres to and invades epithelial cells. *Infection and Immunity* **68**(6), 3140-3146 (2000)

114. A. Ikegami, P. Chung, Y. W. Han, Compementation of the *fadA* mutation in *Fusobacterium nucleatum* demonstrates that the surface-exposed adhesin promotes cellular invasion and placental colonization. *Infection and Immunity* **77**(7), 3075-3079 (2009)

115. J.-W. Huh, T.-Y. Roh, Opportunistic detection of *Fusobacterium nucleatum* as a marker for the early gut microbial dysbiosis. *BMC Microbiol.* **20**(1): 208 (2020)

116. O. Gundogdu, B. W. Wren, Microbe Profile: *Campylobacter jejuni*—survival instincts. *Microbiology* **166**(3), 230-232 (2020)

117. A. Y. Guh, P. K.Kutty, *Clostridioides difficile* infection. *Ann. Intern. Med.* **169**(7), ITC-49-ITC-64 (2018)

118. J. E. Grass, L. H. Gould, B. E. Mahon, Epidemiology of foodborne disease outbreaks caused by *Clostridium perfringens*, United States, 1998–2010. *Foodborne Pathog. Dis.* **10**(2), 131-136 (2013)

119. https://international.neb.com/protocols/0001/01/01/digestion-with-nebnext-dsdna-fragmentase-m0348 [New England Biolabs, Inc., Digestion with NEBNext dsDNA Fragmentase]

120. T. Ribarska, P. M. Bjørnstad, A. Y. M. Sundaram, G. D. Gilfillan, Optimization of enzymatic fragmentation is crucial to maximize genome coverage: a comparison of library preparation methods for Illumina sequencing. *BMC Genomics* **23**:92 (2022)

121. https://international.neb.com/tools-and-resources/troubleshooting-guides/troubleshooting-guide-for-cloning [New England Biolabs, Inc., Troubleshooting guide for cloning]

122. M. Rutz, J. Metzger, T. Gellert, P. Luppa, G. B. Lipford, H. Wagner, S. Bauer, Toll-like receptor 9 binds single-stranded CpG-DNA in a sequence- and pH-dependent manner. *Eur. J. Immunol.* **34**, 2541–2550 (2004)

123. J. B. Kaper, J. P. Nataro, H. L. T. Mobley, Pathogenic *Escherichia coli*, *Nat. Rev. Microbiol.* **2**, 123-140 (2004)

124. S. Bauer, C. J. Kirschning, H. Häcker, Human TLR9 confers responsiveness to bacterial DNA via species-specific CpG motif recognition. *PNAS* **98**(16), 9237-9242 (2001)

125. J. F. Sicard, G. Le Bihan, P. Vogeleer, M. Jacques, J. Harel. Interactions of intestinal bacteria with components of the intestinal mucus. *Front. Cell Infect. Mictobiol.* **7**:387 (2017)

126. G. C. Hansson, Mucus and mucins in diseases of the intestinal and respiratory tracts. *J. Internal Medicine* **285**(5), 479-490 (2019)

127. P. J. Sansonetti, A. Phalipon, M cells as ports of entry for enteroinvasive pathogens: mechanisms of interaction, consequences for the disease process. *Seminars in Immunol.* **11**(3), 193-203 (1999)

128. S. C. Corr, C. C. G. M. Gahan, C. Hill, M-cells: origin, morphology and role in mucosal immunity and microbial pathogenesis. *FEMS Immunol. Med. Microbiol.* **52**(1), 2-12 (2008)

129. J. L. Coombes, F. Powrie, Dendritic cells in intestinal immune regulation. *Nat. Rev. Immunol.* **8**(6), 435-446 (2008)

130. C. Palmer, E. M. Bik, D. B. DiGiulio, D. A. Relman, P. O. Brown, Development of the Human Infant Intestinal Microbiota. *PLOS Biology* **5**(7), e177 (2007)

131. M. G. Dominguez-Bello, E. K. Costello, M. Conteras, M. Magris, G. Hidalgo, N. Fierer, R. Knight, Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *PNAS* **107**(26), 11971-11975 (2010)

132. Y. Shao, S. C. Forster, E. Tsaliki, K. Vervier, A. Strang, N. Simpson, N. Kumar, M. D. Stares, A. Rodger, P. Brocklehurst, N. Field, T. D. Lawley, Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth. *Nature* **574**, 117-121 (2019)

133. K. Kordy, T. Gaufin, M. Mwangi, F. Li, C. Cerini, D. J. Lee, H. Adisetiyo, C. Woodward, P. S. Pannaraj, N. H. Tobin, G. M. Aldrovandi, Contributions to human breast milk microbiome and enteromammary transfer of *Bifidobacterium breve*. *PLoS One* **15**(1), e0219633 (2020)

134. D. Mariat, O. Firmesse, F. Levenez, H. Sokol, J. Doré, G. Corthier, J.-P. Furet, The *Firmicutes/Bacteroides* ratio of the human gut changes with age. *BMC Microbiology* **9**, 123 (2009)

135. C. H. King, H. Desai, A. C. Sylvetsky, J. LoTempio, S. Ayanyan, J. Carrie, K. A. Crandall, B. C. Fochtman, L. Gasparyan, N. Gulzar, P. Howell, N. Issa, K. Krampis, L. Mishra, H. Morizono, J. R. Pisegna. S. Rao, Y. Ren, V. Simonyan, K. Smith, S. VadBrat, M. D. Yao, R. Mazumder, Baseline human gut microbiota profile in healthy people and standard reporting template. *PLoS One* **14**(9), e0206484 (2019)

136. O. Ménard, V. Gafa, N. Kapel, B. Rodriguez, M.-J. Butel, A.-J. Waligora-Dupriet, Characterization of immunostimulatory CpG-rich sequences from different *bifidobacterium* species. *Appl. Environ. Microbiol.* **76**(9), 2846-2855 (2010)

137. B. M. Henrick, L. Rodriguez, T. Lakshmikanth, C. Pou, E. Henckel, A. Arzoomand, A. Olin, J. Wang, J. Mikes, Z. Tan, Y. Chen, A. M. Ehrlich, A. K. Bernhardsson, C. H. Mugabo, Y. Ambrosiani, A. Gustafsson, S. Chew, H. K. Brown, J. Prambs, K. Bohlin, R. D. Mitchell, M. A. Underwood, J. T. Smilowitz, J. B. German,S. A. Frese, P. Brodin, *Bifidobacteria*-mediated immune system imprinting early in life. *Cell* **184**, 3884-3898 (2021)

138. H. Zoghouani, C. M. Hoeman, B. Adkins, Neonatal immunity: faulty T-helpers and the shortcomings of dendritic cells. *Trends Immunl.* **30**(12), 585-591 (2009)

139. E. Maggi, The TH1/TH2 paradigm in allergy. *Immunotechnology* **3**(4), 233-244 (1998)

140. M. E. Kirtland, D. C. Tsitoura, S. R. Durham, Mohamed H. Shamji, Toll-like receptor agonists as adjuvants for allergen immunotherapy. *Front. Immunol.* **11**: 599083 (2020)

141. A. K. Kiemer, R. H. Senaratne, J. Hoppstädter, B. Diesel, L. W. Riley, K. Tabeta, S. Bauer, B. Beutler, B. L. Zuraw, Attenuated Activation of Macrophage TLR9 by DNA from Virulent Mycobacteria. *J. Innate. Immun.* **1**(1), 29-45 (2008)

142. S. Nagata, R. Hanayama, K. Kawane, Autoimmunity and the clearance of dead cells. *Cell* **140**, 619-630 (2010)

143. J. Tel, A. J. A. Lambeck, L. J. Cruz, P. J. Tacken, I. J. M. de Vries, C. G. Figdor, Human plasmacytoid dendritic cells phagocytose, process, and present exogenous particulate antigen. *J. Immunol.* **184**(8), 4276-4283 (2010)

144. C. Hagiwara, M. Tanaka, H. Kudo, Increase in colorectal epithelial apoptotic cells in patients with ulcerative colitis ultimately requiring surgery. J. Gastroenterol. Hepatol. **17** (7), 758-764 (2002)

145. A. H. Wyllie, Glucocorticoid-induced thymocyte apoptosis is associated with endogenous endonuclease activation. *Nature* **284**, 555-556 (1980)

146. J. G. Camp, C. L. Frank, C. R. Lickwar, H. Guturu, T. Rube, A. M. Wenger, J. Chen, G. Bejerano, G. E. Crawford, J. F. Rawls, Microbiota modulate transcription in the intestinal epithelium without remodeling the accessible chromatin landscape. *Genome Res.* **24**, 1504-1516 (2014)

147. A. Piovesan, M. C. Pelleri, F. Antonaros, P. Strippoli, M. Caracausi, L. Vitale, On the length, weight and GC content of the human genome, *BMC Res. Notes* **12**: 106 (2019)

148. H. F. Helander, L. Fändriks, Surface area of the digestive tract—revisited. *Scand. J. Gastroenterol.* **49**(6), 681-689 (2014)

149. S. E. Doyle, R. M. O'Connell, G. A. Miranda, S. A: Vaidya, E. K. Chow, P. T. Liu, S. Suzuki, N. Suzuki, R. L. Modlin, W.-C. Yeh, T. F. Lane, G. Cheng, Toll-like receptors induce a phagocytic gene program through p38. *J. Exp. Med.* **199**(1), 81-90 (2004)

150. Y. Mokhtari, A. Pourbagheri-Sagaroodi, P. Zafari, N. Bagheri, S. H. Ghaffari, D. Bashash, Toll-like receptors (TLRs): An old family of immune receptors with a new face in cancer pathogenesis. *J. Cell. Mol. Med.* **25**(2), 639-651 (2021)

151. J. H. Kauppila, T. J. Karttunen, J. Saarnio, P. Nyberg, T. Salo, D. E. Graves, P. P. Lehenkari, K. S. Selander, Short DNA sequences and bacterial DNA induce esophageal, gastric and colorectal cancer cell invasion. *APMIS* **121**(6), 12016 (2012)

152. M. A. Merrell, J. M. Ilvesaro, N. Lehtonen, T. Rorsa, B. Gehrs, E. Rosenthal, D. Chen, B. Shackley, K. W. Harris, K. S. Selander, Toll-like receptor 9 agonists promote cellular invasion by increasing matrix metalloproteinase activity *Mol. Cancer Res.* **4**(7), 437-447 (2006)

153. M. A. Casasanta, C. C. Yoo, B. Udayasuryan, B. E. Sanders, A. Umaña, Y. Zhang, H. Peng, A. J. Duncan, Y. Wang, L. Li, S. S. Verbridge, D. J. Slade, *Fusobacterium nucleatum* host-cell binding and invasion induces IL-8 and CXCL1 secretion that drives colorectal cancer cell migration. *Sci. Signaling* **13**, eaba9157 1-12 (2020)

154. D. J. Dwyer, D. M. Camacho, M. A. Kohanski, J. M. Callura, J. J. Collins, Antibiotic-induced bacterial cell death exhibits physiological and biochemical hallmarks of apoptosis. *Mol. Cell* **46**(5), 561-572 (2012)

155. M. Shitani, Z. K. Sanchez, K. Kimbara, Genomics of microbial plasmids: classification and identification based on replication and transfer systems and host taxonomy. *Front Microbiol.* **6**:242 (2015)

156. Y. Tan, J. C. Kagen, Microbe-inducible trafficking pathways that control Toll-like receptor signaling. *Traffic* **18**(1): 6-17 (2017)

157. J. Pohar, D. Lainšček, Ana Kunšek, M.-M. Cajnko, R. Jerala, C. Benčina, Phosphodiester backbone of the CpG motif within immunostimulatory oligodeoxynucleotides augments activation of Toll-like receptor 9. *Sci. Rep.* **7**:14598 (2017)

158. R. Meddeb, Z. A. A. Dache, S. Thezenas, A. Otandault, R. Tanos, B. Pastor, C. Sanchez, J. Azzi, G. Tousch, S. Azan, C. Mollevi, A. Adenis, S. El Messaoudi, P. Blache, A. R. Thierry, Quantifying circulating cell-free DNA in humans, *Sci. Rep.* **9**: 5220 (2019)

159. W.-H. Shao, P. L. Cohen, Disturbances of apoptotic cell clearance in systemic lupus erythematosus. *Arthritis Res. Ther.* **13**(1): 202 (2017)

160. A. Mahajan, M. Herrmann, L. E. Muñoz, Clearance deficiency and cell death pathways: a model for the pathogenesis of SLE. *Front. Immunol.* **7**: 35 (2016)

161. S. Jahr, H. Hentze, S. Englisch, D. Hardt, F. O. Fackelmayer, R. D. Hesch, R. Knippers, DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res.* **61**(4), 1659-1665 (2001)

162. Y. Sato, Y. Goto, N. Narita, D. S. B. Hoon, Cancer cells expressing Toll-like receptors and the tumor microenvironment. *Cancer Microenvironment* **2**, 205-214 (2009)

163. F. Wong, A. Amir, Mechanics and dynamics of bacterial cell lysis. *Biophys. J.* **116**(12), 2378-2389 (2019)

164. S. Tsuji, M. Ushio, S. Sakurai, T. Minamoto, H. Yamanaka, Water temperature-dependent degradation of environmental DNA and its relation to bacterial abundance. *PLoS One* **12**(4): e0176608 (2017)

165. X. He, M. Parenti, T. Grip, B. Lönnerdal, N. Timby, M. Domellöf, O. Hernell, C. M. Slupsky, Fecal microbiome and metabolome of infants fed bovine MFGM supplemented formula or standard formula with breast-fed infants as reference: a randomized controlled trial. *Sci. Rep.* **9**: 11589 (2019)

166. J. Pohar, C. Yamamoto, R. Fukui, M.-M. Cajnko, K. Miyake, R. Jerala, M. Benčina. Selectivity of human TLR9 for double CpG motifs and implications for the recognition of genomic DNA. *J. Immunol.* **198**, 2093-2104 (2017)

167. J. Pohar, A. K. Krajnik, R. Jerala, M. Benčina. Minimal sequence requirements for oligodeoxyribonucleotides activating human TLR9. *J. Immunol.* **194**, 3901-3908 (2015)

168. M. E. Khan, C. Borde, E. P. C. Rocha, V. Mériaux, V. Maréchal, P. Escoll, S. Goyard, J.-M. Cavaillon, B. Manoury, N. Doyen, TLR9 activation is triggered by the excess of stimulatory versus inhibitory motifs present in *Trypanosomatidae* DNA. *PLOS Neglected Tropical Diseases* **8**(11), e3308 (2014)

169. R. Lande, D. Ganguly, V. Facchinetti, L. Frasca, C. Conrad, J. Gregorio, S. Meller, G. Chamilos, R. Sebasigari, V. Riccieri, R. Bassett, H. Amuro, S. Fukuhara, T. Ito, Y.-J. Liu, M. Gilliet, Neutrophils activate plasmacytoid dendritic cells by releasing self-DNA-peptide complexes in systemic lupus erythematosus. *Sci. Trans. Med.* **3**(73): 73ra19 (2011)

170. M. L. Coleman, E. A. Sahai, M. Yeo, M. Bosch, A. Dewar, M. F. Olson. Membrane blebbing during apoptosis results from caspase-mediated activation of ROCK I. *Nat. Cell Biol.* **3**, 339-345(2001)

171. J. Lee, J.-H. Mo, K. Katakura, I. Alkalay, A. N. Rucker, Y.-T. Liu, H.-K. Lee, C. Shen, G. Cojocaru, S. Shenouda, M. Kagnoff, L. Eckmann, Y. Ben-Neriah, E. Raz, Maintenance of colonic homeostasis by distinctive apical TLR9 signaling in intestinal epithelial cells. *Nat. Cell Biology*, **8**(12), 1327-1336 (2006)

172. M. P. Chan, M. Onji, R. Fukui, K. Kawane, T. Shibata, S.-I. Saitoh, U. Ohto, T. Shimizu, G. N. Barber, K. Miyake, DNase II-dependent DNA digestion is required for DNA sensing by TLR9. *Nature Comm.* **6**:5853, 1-10 (2015)

173. J. Wu, L. Sun, X. Chen, F. DU, H. Shi, C. Chen, Z. J. Chen, Cyclic-GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. *Science*, **339**:6121 (2013)

174. L. Yu, P. Liu, Cytosolic DNA sensing by cGAS: regulation, function, and human diseases. *Signal Transduction and Targeted Therapy* **6**:170 (2021)

175. A. Basit, M.-G. Cho, E.-Y. Kim, D. Kwon, S.-J. Kang, J.-H. Lee, The cGAS/STING/TBK1/IRF3 innate immunity pathway maintains chromosomal stability through regulation of p21 levels. *Exp. Mol. Med.* **52**(4), 643-657 (2020)

176. J. R. Tversky, T. V. Le, A. P. Bieneman, K. L. Chichester, R. G. Hamilton, J. T. Schroeder, Human blood dendritic cells from allergic subject have impaired capacity to produce interferon-α via TLR9. *Clin. Exp. Allergy* **38**(5), 781-788 (2008)

177. J. Visser, J. Rozing, A. Sapine, K. Lammers, A. Fasano, Tight junctions, intestinal permeability, and autoimmunity celiac disease and type 1 diabetes paradigms. *Ann. N. Y. Acad. Sci.* **1165**, 195-205 (2009)

178. D. E. Matei, M. Menon, D. G. Alber, A. M. Smith, B. Nedjat-Shokouhi, A. Fasano, L. Magill, A. Duhlin, S. Bitoun, A. Gleizes, S. Hacein-Bey-Abina, J. J. Manson, E. C. Rosser, the ABIRISK Consortium, N. Klein, P. A. Blair, C. Mauri, Intestinal barrier dysfunction plays an integral role in arthritis pathology and can be targeted to ameliorate disease. *Med* **2**(7), 864-883 (2021)

179. K. E. Pijls, G. H. Koek, E. E. Elamin, H. de Vries, A. A. M. Masclee, D. M. A. E. Jankers, Large intestine permeability is increased in patients with compensated liver cirrhosis

180. P. Halasz, G. Holloway, B. S. Coulson, Death mechanisms in epithelial cells following rotavirus infection, exposure to inactivated rotavirus or genome transfection. *J. General Virology* **91**, 2007-2018 (2010)

181. G. Chakrabarti, X. Zhou, B. A. McClane, Death pathways activated in CaCo-2 cells by *Clostridium perfringens* enterotoxin. *Infection and Immunity* **71**(8), 4260-4270 (2003)

182. M. Hanus, D. Parada-Venegas, G. Landskron, A. M. Wielandt, C. Hurtado, K. Alvariz, M. A. Hermoso, F. López-Köstner, M. De la Fuente, Immune system, microbiota, and microbial metabolites: the unresolved triad in colorectal cancer microenvironment. *Front. Immunol.* **12**:612826 (2021)

183. C. Stockmann, D. Schadendorf, R. Klose, I. Helfrich, The impact of the immune system on tumor: angiogenesis and vascular remodeling. *Front. Oncol.***4**:69, 1-13 (2014)

184. C. Brennan, W. Garrett, Gut microbiota, inflammation, and colorectal cancer. *Annu. Rev. Microbiol.* **70**, 395-411 (2016)

185. E. Ullrich, M. Bonmort, C. Mignot, G. Kroemer, L. Zitvogel, Tumor stress, cell death and the ensuing immune response. *Cell Death & Diffrentiation* **15**, 21-28 (2008)

186. F. Genua, V. Raghunathan, M. Jenab, W. M. Gallagher, D. J. Hughes, The role of gut barrier dysfunction and microbiome dysbiosis in colorectal cancer development. *Front. Oncol.* **11**: 626349 (2021)

187. Q. Luo, L. Zeng, C. Tang, Z. Zhang, X. Chen, C. Zeng, TLR9 induces colitis-associated colorectal carcinogenesis by regulating NF-κB expression levels. *Oncology Letters* **20**(4):110, 1-10 (2020)

188. C. Gao, T. Qiao, B. Zhang, S. Yuan, X. Zhuang, Y. Luo, TLR9 signaling activation at different stages in colorectal cancer and NF-kappaB expression. *OncoTargets and Therapy* **11**, 5963-5971 (2018).

189. E. Furrie, S. Macfarlane, G. Thomson, G. T. Macfarlane, Microbiology & Gut Biology Group, Tayside Tissue & Tumour Bank, Toll-like receptors-2, -3 and -4 expression patterns on human colon and their regulation by mucosal-associated bacteria. *Immunology* **115**, 565-574 (2005)

190. M. Gagnon, A. Z. Berner, N. Chervet, C. Chassard, C. Lacroix, Comparison of the Caco-2, HT-29 and the mucus-secreting HT-29-MTX intestinal cell models to investigate *Salmonella* adhesion and invasion. *J. Microbiological Methods* **94**, 274-279 (2013)

191. C. Kleiveland. Co-culture Caco-2/immune cells. From K. Verhoeckx et al. (eds), *The Impact of Food Bio-Actives on Gut Health*, DOI 10.1007/978-3-319-16104-4_18, 197-205 (2015)

192. A. R. Mridha, F. Haczeyni, M. M. Yeh, W. G. Haigh, G. N. Ioannou, V. Barn, H. Ajamieh, L. Adams, J. M. Hamdon, N. C. Teoh, G. C. Farrell, TLR9 is up-regulated in human and murine NASH: pivotal role in inflammatory recruitment and cell survival. *Clin. Sci. (Lond.)* **131**(16), 2145-2159 (2017)

193. L. A. Minns, L. C. Menard, D. M. Foureau, S. Darche, C. Ronet, D. W. Mielcarz, D. Buzoni-Gatel, L. H. Kasper, TLR9 is required for the gut-associated lymphoid tissue response following oral infection of Toxoplasma gondii. *J. Immunol.* **176**(12), 7589-97 (2006)

194. L. M. Keubler, M. Buettner, C. Häger, A. Bleich, A multihit model: colitis lessens from the interleukin-10-deficient mouse. *Inflammatory Bowel Disease* **21**(8), 1967-1975 (2015)