Interactions between intestinal commensal bacteria and different hosts: two sides of the same story

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Anna Lange

aus Mutlangen

Tübingen

2018

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	10.01.2019
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Julia-Stefanie Frick
2. Berichterstatter:	Prof. Dr. Andreas Peschel

Erklärung:

Ich erkläre hiermit, dass ich die zur Promotion eingereichte Arbeit selbständig verfasst, nur die angegebenen Quellen und Hilfsmittel benutzt und Stellen, die wörtlich oder inhaltlich nach den Werken anderer Autoren entnommen sind, als solche gekennzeichnet habe. Eine detaillierte Abgrenzung meiner eigenen Leistungen habe ich im Abschnitt *„Contributions“* vorgenommen.

……..

Unterschrift

Tübingen den

Table of Contents

Abbreviations

AMP	Antimicrobial peptide
ATP	Adenosine triphosphate
BD-2	β-defensin 2
BMDC	Bone-marrow derived dendritic cell
CCL20	C-C motif chemokine ligand 20
CCR6	C-C chemokine receptor type 6
CD4	Cluster of differentiation 4
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DNA	Deoxyribonucleic acid
DUOX	Dual oxidase
ELISA	Enzyme-linked immunosorbent assay
HEK	Human embryonic kidney
HGT	Horizontal gene transfer
IBD	Inflammatory bowel disease
IFNγ	Interferon γ
IL2$^{-/-}$	Interleukin 2-deficient
IL-22	Interleukin 22
kbp	kilo basepairs
LPS	Lipopolysaccharide
LTA	Lipoteichoic acid
MAMP	Microbe-associated molecular pattern
MD-2	Myeloid differentiation protein 2
mM	Millimolar
mRNA	messenger RNA
NADPH	Nicotinamide adenine dinucleotide phosphate
NOS	Nitric oxidase synthase
NOX2	NADPH oxidase 2
OMV	Outer membrane vesicle
PCR	Polymerase chain reaction
PSA	Polysaccharide A
PRR	Pattern recognition receptor

Rag1$^{-/-}$	Recombination-activating protein 1-deficient
REGIIIβ	Regenerating islet-derived III β
RNS	Reactive oxygen species
ROS	Reactive nitrogen species
TLR	Toll-like receptor
TNF	Tumor necrosis factor
tRNA	transfer ribonucleic acid

Summary

The gut microbiota crucially contributes to a variety of body functions: it supports nutrient digestion and influences the host immune system. It further impacts progression and/or the outcome of a variety of diseases. Inflammatory bowel diseases (IBD) are one of the most prominent microbiota-influenced diseases, which are accompanied by a compositional shift of the microbiota.

A healthy human intestinal microbiota mostly consists of commensal bacteria with *Firmicutes* and *Bacteroidetes* making up the largest proportion. A balanced microbiota composition is important to maintain health benefits for the host. Commensals can be divided into symbionts and pathobionts. Symbionts are generally characterized by their beneficial immunomodulatory influence on the host, whereas pathobionts may negatively affect the host under certain circumstances therefore promoting intestinal inflammation. To characterize two model intestinal commensal bacterial species – symbiotic *Bacteroides vulgatus* mpk and pathobiotic *Escherichia coli* mpk – we sequenced the strains to analyze their respective genomes.

The *B. vulgatus* mpk genome contains many different mobile genetic elements that have caused enormous genome diversification. Additionally we identified gene loci encoding for the formation of different outer membrane vesicles (OMVs) containing capsular polysaccharides that are produced under stress conditions to supporting bacterial survival. Further, we demonstrated that *B. vulgatus* mpk OMVs support immune priming of host cells therefore supporting intestinal immune homeostasis. Unlike *B. vulgatus, E. coli* strains are only poor gut colonizers during intestinal homeostasis but their accumulation is boosted during inflammation. We identified metabolic gene clusters in the *E. coli* mpk genome mediating bacterial survival during intestinal inflammation, most importantly the ethanolamine and 1,2-propanediol utilization clusters.

Since vertebrates represent the most commonly used models to study microbe-host interactions we aimed to provide an alternative invertebrate host model, which is suitable to distinguish symbiotic from pathobiotic intestinal commensals. Therefore, we established an insect replacement model using the *Galleria mellonella* larva, which allows for the prediction of the immunogenic potential by commensal bacteria in mice.

Zusammenfassung

Die Darmmikrobiota nimmt entscheidend Einfluss auf verschiedene Körperfunktionen: sie wirkt unterstützend auf die Verdauung und beeinflusst das Immunsystem des Wirtes. Zudem beeinflusst sie das Fortschreiten und/oder den Ausgang von verschiedenen Erkrankungen. Chronisch-entzündliche Darmerkrankungen (CED) zählen zu den am häufigsten durch die Mikrobiota beeinflussten Erkrankungen, die mit einer veränderten Mikrobiota-Zusammensetzung einhergehen.

Die Darmmikrobiota eines gesunden Erwachsenen setzt sich hauptsächlich aus kommensalen Bakterien zusammen, wobei das Phylum der *Firmicutes* und der *Bacteroidetes* den größten Anteil ausmachen. Eine ausgeglichene Zusammensetzung der Mikrobiota trägt maßgeblich zum Erhalt des gesundheitlichen Wohlbefindens des Wirtes bei. Kommensale können weiterhin in Symbionten und Pathobionten unterteilt werden. Symbionten sind durch vorteilhafte immunmodulatorische Eigenschaften auf den Wirt charakterisiert, während Pathobionten, unter bestimmten Vorrausetzungen, einen negativen Einfluss auf den Wirt nehmen können und dadurch Entzündungen im Darm fördern können. Zur besseren Charakterisierung zweier Modell-Darmkommensale – dem symbiotischen *Bacteroides vulgatus* mpk und dem pathobiontischen *Escherichia coli* mpk – haben wir die Genome beider Bakterienspezies sequenziert und untersucht.

Das Genom des *B. vulgatus* mpk enthält viele verschiedene mobile genetische Elemente, deren Einfluss eine enorme Formbarkeit des Genoms zur Folge hat. Außerdem konnten verschiedene Genloci zur Bildung von *outer membrane vesicles* (OMVs) identifiziert werden, die sogenannte Kapselpolysaccharide enthalten und unter bestimmten Stressbedingungen als Überlebensstrategie von Bakterien gebildet werden. Darüber hinaus konnten wir zeigen, dass OMVs das Immune-Priming der Wirtszellen unterstützen und damit zur Aufrechterhaltung der intestinalen Homöostase beitragen. Im Gegensatz zu *B. vulgatus* ist *E. coli* unter physiologischen Bedingungen kaum in der intestinalen Mikrobiota präsent. Allerdings wird das Wachstum von *E. coli* durch eine Entzündung im Darm verstärkt. Wir konnten verschiedene metabolische Gencluster identifizieren, die das Überleben von *E. coli* während einer Entzündung im Darm unterstützen; am wichtigsten sind hier die Cluster, die das Verstoffwechseln von Ethanolamin und 1,2-Propandiol ermöglichen.

Da Vertebraten die am häufigsten verwendeten Tiermodelle sind, um Bakterien-Wirts-Interaktionen zu untersuchen, wollten wir ein alternatives Invertebraten-Modell zur Verfügung stellen, dass es ermöglicht symbiotische von pathobiotischen Darmkommensalen zu unterscheiden. Deshalb haben wir ein Insektenmodell etabliert, das es erlaubt, eine Vorhersage über das immunogene Potential von kommensalen Bakterien in Mäusen zu treffen.

Publications

a **Lange A.***, Beier S.*, Steimle A., Autenrieth I.B., Huson D.H., Frick J.S.: Extensive Mobilome-Driven Genome Diversification in Mouse Gut-Associated *Bacteroides vulgatus* mpk. Genome Biology and Evolution 8 (4): 1197-207 (2016)

b Maerz J.K., Steimle A., **Lange A.**, Bender A., Fehrenbacher B., Frick J.S.: Outer membrane vesicles blebbing contributes to *B. vulgatus* mpk-mediated immune response silencing. Gut Microbes 9 (1): 1-12 (2018)

c **Lange A.**, Tschörner L., Beck C., Beier S., Schäfer A., Frick J.S.: Survival strategies of *E. coli* mpk in the inflamed intestine. (in preparation)

d **Lange A.**, Schäfer A., Bender A., Steimle A., Parusel R, Frick J.S.: *Galleria mellonella*: a novel invertebrate model to distinguish intestinal symbionts from pathobionts. Frontiers in Immunology 9 (2114) (2018)

Contributions

Publication a

I performed all experiments in the laboratory. Genome analysis was performed together with Sina Beier. Sina Beier and I contributed equally in writing and editing the manuscript together with Daniel Huson, Alex Steimle and Julia-Stefanie Frick.

Publication b

I identified the nine gene loci for OMVs in the *B. vulgatus* genome and had the idea to investigate the anti-inflammatory potential of the OMVs. I took part in designing the project with Alex Steimle and contributed in editing the manuscript.

Publication c

I performed whole genome sequencing of *E. coli* mpk, constructed the *E. coli* mpk Δ*eutR* mutant, and performed all *in vivo* experiments. The manuscript was written me and Julia-Stefanie Frick.

Publication d

The experiments were performed together with Andrea Schäfer. The paper was written by me, Alex Steimle and Julia-Stefanie Frick.

Introduction

1. The gut microbiota

1.1 Gut microbiota composition

Mucosal surfaces of vertebrates are inhabited by a large microbial ecosystem – the microbiota – which is comprised by numbers of bacteria, fungi, viruses, and parasites; especially the gastrointestinal tract is densely populated by trillions of microbes.[1, 2]

The microbiota gets established in newborns right after birth and hence the continuous dynamic process of shaping the gut microbiota composition starts and reaches its highest complexity in human adults.[3] The microbiota of a healthy adult is dominated by two bacterial phyla: *Firmicutes* and *Bacteroidetes*[3, 4]; other less abundant bacterial phyla like *Verrucomicrobia*, *Actinobacteria*, and *Proteobacteria* are also present in the majority of people´s microbiomes.[4] Although the gut microbiota composition is highly variable among distinct individuals, functional gene profiles are similar.[4] For example, all gut ecosystems contain pathways for central amino acid and carbohydrate metabolism but certain specific pathways for nutrient transporters or vitamin synthesis can be restricted to certain microbial species or even a certain strain.[4] Lozupone and colleagues suggest that we all share functional core microbiomes but not a core microbiota. Thus, the overall function of the microbiome is similar but not the bacterial composition.[4] As the composition of the gut microbiota as a dynamic environment is constantly changing within a certain individual, there are various factors that contribute to microbiota composition shifts like age, antibiotic treatments, diet, host genetics, stress or host physical activity.[5]

1.2 Different classes of commensals: symbionts and pathobionts

As previously mentioned, the bacterial gut microbiota composition is diverse. The microbiota members all have in common that they are commensal and nonpathogenic.[2] Commensals can be classified according to different nomenclatures, like phylogenetic groups or according to their metabolic functions. Furthermore, commensals can be classified by their potential effect on the host.[2] Ivanov and Honda provide such a classification into probiotic, autobiotic/symbiotic, and pathobiotic bacteria. Probiotics have a more transient host association, and are

administered to the host in specific amounts. They confer health benefits, and are not necessarily part of the resident microbiota, and can further indirectly affect the beneficial intrinsic microbiota.[2] Symbionts are permanently associated with the host and have important immunomodulatory functions, and directly influence host immune cell homeostasis or function. They are part of the regular microbiota such as *Faecalibacterium prausnitzii*[2], *Bacteroides fragilis*[6] or *Bacteroides vulgatus*[7]. Another important group is represented by the pathobionts, which are also part of the permanent microbiota. In genetically predisposed individuals pathobionts can promote intestinal inflammation.[8] Therefore bacteria of this group can have detrimental effects on the host under certain environmental conditions when microbiota or host immunity is disturbed.[2] Such pathobionts are for example *Escherichia coli*[1], *Helicobacter hepaticus*[9], or *Clostridium difficile*[9].

Our group has extensively worked with *B. vulgatus* strain mpk as a model symbiont and *E. coli* mpk as a model pathobiont. Some years ago, the group identified *E. coli* mpk to induce colitis in gnotobiotic *Il2*$^{-/-}$ mice, whereas *B. vulgatus* prevented from *E. coli*-induced inflammation.[10] Recently, we found that intestinal commensals had an influence on secretion and intracellular activity of host cathepsin S (CTSS). *B. vulgatus* mpk as a symbiotic bacterium, may actively regulate this process, and helps to stabilize physiological CTSS levels to prevent induction of pathological inflammatory responses.[11]

1.3 Microbes in the intestine

The microbiota confers several advantages for the host: digestion of nutrients, immune regulation, and prevention of pathogen colonization or support of growth inhibition of the endogenous potentially harmful pathobionts.[12, 13] The latter is dependent on the concept of colonization resistance.[14] Colonization resistance can be mediated directly by interbacterial competition or indirectly by the host immune system stimulated by commensals (discussed in the next section). Mechanisms of direct competitions will be discussed in this section. Direct colonization resistance includes the competition for nutritional niches.[12] Carbohydrates, which are represented by complex polysaccharides or the more simple monosaccharides, provide an important nutrient and energy source for the microbiota. Each microbiota member has its own preference for certain dietary carbohydrates e.g. the pathogen

Citrobacter rodentium and commensal *E. coli* are restricted to simple sugar consumption, whereas other commensals such as *B. thetaiotaomicron* are capable of consuming complex and simple sugars. Different mouse experiments have shown that colonization of *C. rodentium* can be either reduced through competition with *E. coli* that prefers the same carbohydrates or by commensals that are broad carbohydrate degraders and manipulated by a carbohydrate restrictive diet. Similar competitive advantages are mediated by competition for sialic acid, or succinate. Consumption of both substrates by intrinsic commensals provides colonization resistance against *Clostridium difficile*.[13]

The so called metabolic exclusion of *Proteobacteria* members by fermentation end products can also be observed in the intestine. *Firmicutes* and *Bacteroidetes* strains break down complex polysaccharides by fermentation and produce short-chain fatty acids (SCFA), which stimulate host immune system and reduce the abundance of *Proteobacteria* in the intestine.[14]

Further, direct competition can be mediated by toxins or antimicrobial molecules. Type VI secretion systems (T6SS) were originally suggested to be a bacterial virulence trait which secrets antibacterial toxins. Though, recent evidence shows that the system could also be involved in interbacterial competition. Such T6SSs were identified in *Bacteroides* strains. This suggests a reason why *Bacteroides* spp. are present in a great abundance in the intestine and a possible contribution of this system to ecological homeostasis by growth inhibition.[12, 13] *E. coli* Nissle on the other hand is able to secrete microcins which are small antimicrobial peptides that can kill competitors and are even able to target *Salmonella* Typhimurium and limit their expansion in the intestine.[15]

2. Microbiota-host interactions

The previous section highlighted the gut microbiota, and how microbes contribute to shape the microbiota composition. However, the intestinal microbiota also deeply impacts the host immune system.[12]

In addition to their barrier function intestinal epithelial cells (IECs) are also important components of the immune system. The microbiota is recognized by epithelial receptors such as Toll-like receptors (TLRs) which specifically interact with microbe-associated molecular patterns (MAMPs) such as lipopolysaccharides (LPS),

lipoteichoic acids (LTAs), or peptidoglycans. After recognition, downstream responses are activated depending on the specific structure of the bacterial antigen.[16, 17]

2.1 Microbes and host during intestinal homeostasis

Commensal bacteria are crucial for induction and maintenance of intestinal homeostasis.[18] Upon their recognition they stimulate different innate and adaptive immune responses.[19]

The microbiota stimulates immunoglobulin A (IgA) production and secretion.[1, 16] IgA is an immunoglobulin produced in IgA^+ B cells and plasma cells of mucosal tissues in the intestine. It is transported though epithelial cells and secreted to the intestinal lumen. IgA coats commensals and antigens to prevent their binding to epithelial cells and penetration of the lamina propria.[1]

Innate lymphoid cells (ILCs) are important mediators of mutual interactions between the microbiota and the host immune system and share functional characteristics with T cells.[20] However, ILCs do not express antigen receptors.[21] ILC differentiation and ILC-induced production of IL-22 is reported to depend on the microbiota. IL-22 stimulates expression of regenerating islet-derived (REG) IIIβ and REGIIIγ which have antimicrobial properties.[1, 22] Additional antimicrobial peptides produced by epithelial cells are α- and β-defensins. β-defensin 1 (BD-1) is continuously expressed in the intestine.[19] These mechanisms support the balance of microbiota composition.[1, 19, 22]

T helper 17 (T_H17) cells belong to a certain linage of $CD4^+$ T_H cells. In contrast to T_H1 and T_H2 cells, they are preferentially expressed in the intestine. T_H17 cells are involved in host defense responses as well as development of autoimmune diseases by production of IL-17 and IL-22.[1] Regulatory T (T_{Reg}) cells another $CD4^+$ T_H cell subset can also accumulated in the intestine, and fulfill immunosuppressive functions to maintain homeostasis and prevent autoimmune diseases.[23] T_{Reg} production is stimulated by certain commensal microbiota members. It has been demonstrated that *Bacteroides fragilis* is able to promote differentiation of IL-10-secreting T_{Reg} cells by a distinct component of their outer membrane vesicles (OMVs): polysaccharide A (PSA).[1, 16]

The above described mechanisms help to fortify of the intestinal barrier function and contribute to maintenance of intestinal homeostasis.[1, 16, 22]

2.2 Microbes and host during intestinal diseases

Apart from crucial contributions of the intestinal microbiota to gut homeostasis the microbiota can have an influence on intestinal diseases.[1]

It is well studied that the microbiota represents an important factor which contributes to pathological intestinal inflammation eventually resulting e.g. in inflammatory bowel diseases (IBD). In mice and humans, IBD is characterized by an inadequate activation of immune responses which might be caused by commensals. It was shown that different genetic predispositions of IBD patients are associated with malfunction of bacterial recognition and defective antibacterial mechanisms.[1]

In IBD patients, the epithelial barrier is leaky and innate immune responses are disturbed, and e.g. the expression of pattern recognition receptors (PRR) such as TLR4, which binds LPS, are strongly upregulated.[24, 25] Further, antigen recognition is disturbed, and intestinal cells do not properly recognize commensal bacteria and induce T_H1 and T_H17 pro-inflammatory responses. IECs, which normally balance T cell responses, activate further T cells instead of inducing T cell anergy. IBD patients are also defective of clearance mechanisms towards overreactive and autoreactive T cells, therefore activated T cells persist and do not undergo apoptosis. The balance of effector and regulatory T cells is disrupted. The activated effector T cells lead to production and release of inflammatory cytokines like IL-12, IL-18, TNF-like 1A and IFNγ which cause stimulation and secretion of IL-1, TNFα and IL-6 from macrophages.[25] Previous IL-10 secretion by T_{Reg} cells and the antimicrobial activity of REGIIIγ are therefore inhibited.[1] Such defects can lead to the accumulation of pathobionts like *E. coli* due to lowered colonization resistance.[26, 27] Similar immune responses are induced during pathogen infection.[22] Interestingly, antimicrobial peptide hBD-1 expression is stable during inflammation but a set of β-defensins (BD-2, BD-3, BD-4) is upregulated upon pathogen and pathobiont encounter.[19]

The increased colonization by pathobionts or pathogens leads to a completely altered intestinal environment which influences the composition of the whole microbial community.[28]

Intestinal inflammation involves the accumulation of reactive oxygen and nitrogen species (ROS/RNS) induced by dual oxidase 2 (DUOX2), NADPH oxidase 2 (NOX2) (both ROS production) and inducible nitric oxide synthase (iNOS) (mainly RNS production) at the epithelium.[29] After diffusion into the lumen they are processed to terminal respiratory electron acceptors such as tetrathionate or nitrate. Those electron acceptors become available during inflammation and create a distinct nutritional niche favorable for *Enterobacteriaceae* like *Salmonella* Typhimurium and *E. coli* which can grow by anaerobic respiration.[28] Anaerobic nitrate or tetrathionate respiration couple the reduction of nitrate to nitrite, and tetrathionate to thiosulfate, respectively.[29] The formation of this alternative nutritional niche enables *Salmonella* to consume the non-fermentable substrate ethanolamine which is part of the membrane component phosphatidylethanolamine.[30] *S.* Typhimurium is further able to use 1,2-propanediol, the fermentation end product of fucose and rhamnose degradation, by anaerobic respiration as well as by anaerobic fermentation.[31]

3. Microbiota and other animal disease models

The intestinal microbiota is a decisive factor for health and disease not only for mice and human immune systems like discussed above.

The zebrafish *Danio rerio* harbors a complex microbiota; however, the simplest composition among vertebrate models.[32] It was demonstrated that germ-free zebrafish which were colonized with microbes changed gene expression of more than 200 genes of which 59 genes were conserved between zebrafish and mice. The expressed genes were involved in proliferation of epithelial cells, promotion of nutrient-related metabolism, and innate immune responses. This suggests a high degree of conservation of host responses towards microbes.[33]

Stephens and colleagues observed that the microbiota of zebrafish changes throughout the development from larval, to juvenile and adult stage.[34] The major phylum is comprised by *Proteobacteria* which undergoes compositional changes from the larval to adult stage. Among this phylum the γ-*proteobacteria* were the most abundant class and were especially abundant in larval intestines. Throughout the development, changes within the lesser abundant α-, β- and δ-*proteobacteria* classes could also be observed. Another important phylum is represented by the *Firmicutes* with three major classes: *Fusobacteria* and CK-1C4-19 (group of uncultured

Firmicutes) predominantly detected in adults, and *Bacilli* which were found in stable amounts in all different developmental stages. Further bacteria of the *Actinobacteria* phylum were identified.[34] Although the microbiota composition differs from the mouse or human microbiota the resulting host responses during bacterial colonization are conserved. It was shown that microbiota composition like in mice can mediate disease susceptibility in an intestinal inflammation model pointing out the use of zebrafish as a disease model.[33] Inflammatory bowel disease can be modeled in zebrafish by chemically-induced inflammation using oxazolone. When the animals were treated with the antibiotic colistin sulfate the microbiota was dominated by the γ-*proteobacteria* class and the zebrafish developed inflammation after oxazolone application. On the other hand, when the zebrafish were treated with vancomycin prior to oxazolone treatment the class of *Fusobacteria* was predominant and inflammatory responses were significantly decreased.[32] Like observed in mice and humans the microbiota has a decisive effect on the outcome of a disease.

The fruit fly *Drosophila melanogaster* as an insect model species harbors a less complex microbiota with only 1-30 taxa mostly belonging to the genera *Acetobacter* and *Lactobacillus* in laboratory *Drosophila* stocks. Some species are laboratory-specific such as *Gluconobacter morbifer*, or *Enterococcus faecalis*, whereas *Lactobacillus plantarum* and *Acetobacter pomorum/ pasteurianus* have been found in *Drosophila* of different laboratories.[35] In wild-caught flies a higher variation among *Proteobacteria* was observed. Despite increased species diversity in wild flies *Acetobacter/ Gluconobacter* and *Lactobacillus* can be considered as common microbiota members.[35] There are similarities between the contribution of *Drosophila* and mammalian microbiota to intestinal homeostasis. The resident microbes were found to promote epithelial barrier function and it was suggested that they are important in nutrient allocation and metabolism.[36]

It is reported that *Drosophila* host genes contribute to shape the microbiota composition. Studies have shown that evolutionary conserved pathways such as the immune deficiency pathway responsible for antimicrobial peptide (AMP) production and the DUOX pathway inducing ROS production are triggered in intestinal epithelial cells in response to gut microbes. It is reported that efficient ROS-signaling is indispensible to balance intrinsic microbiota composition. If *Drosophila* is not able to balance the composition of the microbiome the abundance of certain microbiota members can increase and lead to deregulation of intestinal immunity. [37] Recently, a

Drosophila Alzheimer disease model was proposed and the involvement of intestinal dysbiosis was associated with neurodegeneration. Oral application of a non-pathogenic *Erwinia* sp. (Ecc15), which was able to colonize the intestine, lead to enteric dysbiosis and in turn influenced host immune system to recruit immune hemocytes to the brain and cause neuroinflammation.[38]

Apart from well-studied mammalian host models, non-mammalian hosts like zebrafish or *Drosophila* provide alternatives to study complex microbiota-driven diseases.

4. Invertebrate host models

The protection and replacement of animals used for scientific purposes are one of the main goals of the Directive 2010/63/EU as animal experimentation is increasingly associated with ethical problems mainly due to the feeling of pain and suffering of animals during experimentation. These problems increased the need for replacement models with a particular focus on animals with only primitively developed neural systems and almost no feeling of pain.[39] Therefore invertebrates have been increasingly proposed as replace organisms in biomedical research in recent years. Besides ethical aspects, invertebrate models offer further advantages such as the possibility to study larger numbers of individuals, cost-effectiveness and shorter live cycles.[40, 41] Since many models do not tolerate environmental temperatures of 37 °C such as *Caenorhabditis elegans*[39] or *Drosophila melanogaster*[42] the use of most invertebrate models was limited in the research of microbe-host interactions investigating human pathogens or commensal bacteria.

A current frequently used insect model is the larva of the greater wax moth *Galleria mellonella* which is a ubiquitous pest of beehives which causes severe damage due to its destructive way of feeding.[43] It is reported to become a reference model to study bacterial pathogenesis.[39] Although it combines several advantages like cost efficiency, ethics, high throughput analysis and feasibility it lacks a very important feature: a fully annotated genome sequence. Fortunately, different transcriptome sequencing projects are available in the databases.[44]

4.1 Innate immunity parallels

The use of *G. mellonella* as a host model is advantageous as it contains several relevant evolutionary conserved innate immune mechanism as well as insect-specific responses. Insect immunity comprises different components like anatomical and physiological barriers which protect insects from pathogen invasion. The insect body is covered by chitin and serves as protection against infection and mechanical injury. The trachea, the insect lung, is also protected by a chitin layer; further the low humidity inside the trachea creates unfavorable conditions for microbes and prevents colonization. The specific structure of the digestion system contributes further to prevent pathogen colonization: foregut and hindgut are covered with chitin and the biochemical conditions (pH, enzymes) in the digestive system support antimicrobial defense.[45] Insect midguts have a peritrophic membrane similarly like epithelial cells of vertebrates.[45, 46]

One of the most important points when studying host-microbe interactions is the ability to distinguish self from non-self. The ability to do this is evolutionary conserved among vertebrates and invertebrates.[47] Different proteins which serve as pattern recognition receptors were identified in *G. mellonella*: peptidoglycan recognizing proteins, Gram-negative binding proteins, β-1,3-glucan recognition proteins, as well as LPS-binding proteins, apolipophorin and hemolin.[45, 48, 49] Upon microbe recognition insects can induce various cellular and humoral immune responses, which is also homologous in vertebrates.[41, 45, 49] However, and in contrast to vertebrates insects lack any adaptive immune responses and therefore insects require an effective innate immune system.[41]

Cellular immunity involves hemocytes in the hemolymph which is analogous to mammalian blood. Different types of hemocytes were identified in the hemolymph: prohemocytes, plasmatocytes, granular cells, coagulocytes, spherulocytes and oenocytoids. The plasmatocytes and granular cells are involved in mediating phagocytosis, nodulation and encapsulation; phagocytosis is assumed to be similar in insects and mammals.[41] Oenocytoids are involved in carrying prophenoloxidase (PPO) precursors.[45]

Humoral immunity involves constitutive and infection-triggered production of defense molecules such as ROS/RNS and AMPs.[45] In *Drosophila* it was found that DUOX-dependent production of antimicrobial ROS serves as the first line of intestinal

immune defense response. As a second major immune response *Drosophila* induces local production of AMPs to create a complementary response against microbes which are resistant against ROS.[50] *G. mellonella* is able to induce a wide range of AMPs upon microbe challenge such as anti-fungal gallerimycin (defensin-like), cecropin and moricin which have activity against Gram-negative and Gram-positive bacteria and gloverin which interacts with LPS.[41, 45] Secretion of both ROS/RNS and AMPs is homologous in mammals. ROS and RNS also accumulate in response to microbial infection or inflammation.[51] Furthermore, the formation of AMPs such as different β-defensins is triggered upon pathogen infection.[19]

A further humoral response is the PPO pathway which causes melanization of foreign molecules.[45] The latter is described as defense mechanism against pathogens which involves deposition of melanin to encapsulate microbes. This mechanism is analogous to mammalian abscess formation.[41] PPOs belong to the group of type-3 copper proteins, which are distributed among many organisms including vertebrates, plants, and microbes. Interestingly, this group of proteins is named tyrosinase in mammals and microbes, and polyphenol oxidase in plants. Each of these proteins is involved in different physiological processes: in mammals tyrosinase activity is linked with hair and skin color, microbial tyrosinases are related with pathogenicity, plant polyphenol oxidase activity causes browning of e.g. fruits. Only insect PPOs are involved in innate immune responses.[52]

Despite the differences between the insect and mammalian immune system such as the lack of adaptive responses or the different physiological functions of PPO, there are important similarities like microbe recognition or secretion of ROS/RNS and AMPs. These homologies lead to the conclusion that insects are reasonable animal replacement models.

4.2 Infection models

G. mellonella enables the investigation of bacterial virulence factors and the evaluation of antimicrobial compound efficacy. However, even more complex aspects of innate immunity and virulence can be investigated using this animal model.[39]

The most frequently used infection route is *via* hemolymph injection through the skin or in the last proleg. The so-called force-feeding method was suggested to be a accurate way of oral administration of bacteria.[41] In general, the oral administration

better resembles the natural route of infection by intestinal pathogens. However, only limited research was performed using the oral administration model.

G. mellonella is a well-suited *in vivo* model to assess the virulence of *Listeria monocytogenes*. *G. mellonella* survival upon *Listeria* infection was a suitable method to distinguish different strains according to their degree of virulence, which could be correlated with previous infections in the mouse model. Further, it was found that important pathogenesis-related genes of *L. monocytogenes* influence the mortality of *G. mellonella*. In addition, it was shown that *L. monocytogenes* caused the activation of AMP responses in *G. mellonella* and therefore enhanced the defense against infection.[53]

The *G. mellonella* model was recently used to evaluate the efficacy of combined antibiotic treatment to cure infections of multidrug-resistant (MDR) *Pseudomonas aeruginosa*. Prior to *G. mellonella* infection the minimum inhibitory concentrations of different drugs was tested. Dual and triple combinations of antibiotics display an increased effect compared with single component therapies against *in vivo* infection of *G. mellonella* with MDR *P. aeruginosa*.[54]

Aim of this work

In previous studies our group could show that symbiotic or pathobiotic commensals differentially influenced the intestinal immune system of the host and are therefore strongly impacting immune system-related pathologies.[10, 11, 55] We could demonstrate that symbiotic *B. vulgatus* mpk promotes anti-inflammatory host responses and protects from pathological intestinal inflammation in different mouse models of experimental colitis. In contrast, pathobiotic *E. coli* mpk causes a pro-inflammatory host response and exhibits strong colitogenic potential in mouse models of experimental colitis.[10] In addition, we could show that LPS endotoxicity influences the degree of inflammation in *Rag1*$^{-/-}$ mice.[55]

So far the group focused mainly on the characterization of immune responses in the mouse caused by different bacteria. However, the bacterial factors that might induce these responses were not investigated. As a part of this work, light should be shed on the genetic perspectives of gut microbiota symbionts and pathobionts. Therefore the genomes of the two model commensals - *B. vulgatus* mpk and *E. coli* mpk – were sequenced and analyzed. And as there are two sides of gut-microbe interactions, we looked for hints to explain the different immunomodulatory potential of the two commensal strains according to their genome information (i). Furthermore, we were interested in microbial metabolism of the pathobiotic *E. coli* strain which should be characterized in the mouse intestine during health and disease. We identified and investigated survival strategies of *E. coli* mpk in the inflamed intestine (ii).

In the last part of this work the potential of an alternative host model, the larvae of the greater wax moth *Galleria mellonella,* is assessed. The question if different mammalian gut commensals influence both vertebrate and invertebrate hosts in similar ways should be answered (iii). We established a larvae force-feeding model to resemble the natural route of bacterial intake and assessed similarities with our applied mouse model, and our already available results.

Results of the following accepted publications and manuscripts are summarized and discussed in the individual "Results and discussion" sections:

a **Lange A.***, Beier S.*, Steimle A., Autenrieth I.B., Huson D.H., Frick J.S.: Extensive Mobilome-Driven Genome Diversification in Mouse Gut-Associated *Bacteroides vulgatus* mpk. Genome Biology and Evolution 8 (4): 1197-207 (2016)

b Maerz J.K., Steimle A., **Lange A.**, Bender A., Fehrenbacher B., Frick J.S.: Outer membrane vesicles blebbing contributes to *B. vulgatus* mpk-mediated immune response silencing. Gut Microbes 9 (1): 1-12 (2018)

c **Lange A.**, Tschörner L., Beck C., Beier S., Schäfer A., Frick J.S.: Survival strategies of *E. coli* mpk in the inflamed intestine. (Manuscript in preparation)

d **Lange A.**, Schäfer A., Bender A., Steimle A., Parusel R, Frick J.S.: *Galleria mellonella*: a novel invertebrate model to distinguish intestinal symbionts from pathobionts. Frontiers in Immunology 9 (2114) (2018)

Results and discussion

1. Characteristics of *Bacteroides vulgatus* mpk - a model gut symbiont: mobile genetic elements and anti-inflammatory potential of outer membrane vesicles

Bacteroides strains are common members of one of the two predominant gut bacterial phyla: the *Bacteroidetes*.[56, 57] At least 25 different *Bacteroides* species can be found in the intestinal microbiota with *Bacteroides vulgatus* among them.[58] The population of *Bacteroides* is versatile because the species exhibit phenotypic variations. One group of *B. fragilis* strains carries an enterotoxin[59]; according to their genome information many strains are well equipped to degrade different polysaccharides[58] and harbor many mobile genetic elements[60], and some of the *Bacteroides* species are reported to have immunomodulatory function and influence host immunity.[2, 10]

The genome of *B. vulgatus* mpk contains 5.17 Mbp and lacks any extrachromosomal structures. Many gene clusters for carbohydrate utilization – so-called starch utilization systems or polysaccharide utilization loci – could be identified (publication a, supplementary table S3), which are typical for *Bacteroides* strains.[61] Furthermore, many different kinds of mobile genetic elements were found along the genome. Up to 8 % of all genes could be assigned to mobile genetic elements: we could identify 16 insertion sequence (IS) 21-like and 14 IS4-like elements, a complete chromosomally-encoded conjugative transposon (CTn) and two truncated CTns. A complete Mu phage was found, and many genes coding for transposases, integrases, and mobile genetic element proteins (publication a, table 1). Even a CRISPR/Cas system could be identified. Those mobile DNA structures seem to be characteristic for *B. vulgatus* strains because similar structures are also spread over the *B. vulgatus* ATCC8482 type strain, but with little less abundance. The only exception was the CRISPR/Cas system which was absent in the type strain.

External-driven genome rearrangements

Comparative homology analysis of *B. vulgatus* mpk and ATCC8482 revealed that most of the shared protein-coding genes are mobile genes. This observation was also true for comparisons of *B. vulgatus* and other *Bacteroides* spp. It was reported that especially CTns have a major impact on genome plasticity and horizontal gene transfer (HGT) events.[58, 62] The chromosomally encoded CTns belong to the group of integrated conjugative elements (ICEs) and encode all the required genes to transfer

themselves to another cell or sometimes to facilitate the transfer DNA from a bacterial 'donor' species carrying a plasmid to a recipient without a plasmid or ICE.[63] This was prompting to analyze the potential impact that the encoded CTn might have caused in the *B. vulgatus* mpk genome. Protein homology analysis revealed that the complete CTn found in the *B. vulgatus* mpk genome was similar to a CTn present in *B. fragilis* YCH46 and highly similar to the CTn341 of a human *Bacteroides* isolate.[64] This CTn341-like structure was also found in *B. vulgatus* ATCC8482 but with less similarities (Figure 1 and publication a, supplementary table S5). When compared to the CTn of *B. fragilis* the CTnMPK lacks a *tetQ* gene between *rteA* and *bmhA* but contains an insertion of 11 genes (BvMPK_0101-0117) at this position (Figure 1, publication a, supplementary table S6). These 11 genes were absent in all other compared *Bacteroides* genomes. Among those 11 genes were transcriptional regulators, one glyxoxalase protein and a β-lactamase gene. The probably inserted β-lactamase gene prompted us to analyze the antibiotic resistance pattern because this gene was absent in *B. vulgatus* ATCC8482 (publication a, table 2). Resistance to different antibiotics was tested in *B. vulgatus* mpk and ATCC8482. Indeed, a different antibiotic resistance pattern could be determined: *B. vulgatus* mpk had an additional resistance against second, third and fourth generation cephalosporins, which was not true for the ATCC8482 strain. Therefore we propose that the gene BvMPK_0116 coding for the β-lactamase was acquired by HGT.

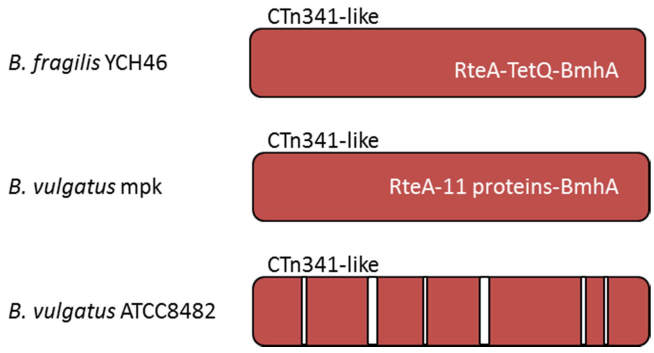


Figure 1: Comparison of different *Bacteroides* Conjugative Transposons (CTn). The genome of *B. fragilis* YCH46 harbors a CTN341-like structure containing a succession of RteA-TetQ-BmhA. *B. vulgatus* mpk genome contains a similar CTn341-like gene sequences, which comprises an 11 gene insertion between *rteA* and *bmhA*. *B. vulgatus* ATCC8482 contains CTn341 as well but lack certain genes inside the CTn341-like cluster.

A 40 kbp stretch could be identified that represented characteristics of a genomic island typical mosaic-like gene distribution of the genes and the random gene arrangement.[65] The stretch is similar to a region found in *B. thetaiotaomicron* and one *B. dorei* isolate. The similarity of the region, even on nucleotide level, between *B. vulgatus* mpk and *B. thetaiotaomicron* was surprising. After a database comparison no other significantly similar hits among other *Bacteroides* sp. could be identified. Therefore, we suggest that the transfer of the region might have occurred in a common ancestor of *B. thetaiotaomicron* and *B. dorei* and got partially lost in the following generation of strains.

Internal genome rearrangements

Besides potential HGT events that indicate an external influence on the genome, there were also events identified that suggest internal genome rearrangements. Two different types of IS elements were identified: IS4-like and IS21-like elements (publication a, supplementary table S8). IS4-type elements are composed of only one gene; IS21-type of two genes encoding a transposase and an ATP-binding protein. IS4-like elements occur in both *B. vulgatus* mpk and ATCC8482 in about the same frequency. IS21-like elements on the other hand could be found more often in the mpk strain, which encodes 16 of this type. Three of these transposase-ATP-binding protein pairs (BvMPK_1852/53, BvMPK_1710/11, BvMPK_1044/45) are paralogs with high similarity. BvMPK_1044/45 could be the result of a gene duplication and transition event because *traG* and *traH* genes were found upstream of the ATP-binding protein, which are also part of the CTn upstream of BvMPK_1852/53. *TraG* and *traH* might have been copied and transited along with BvMPK_1044/45. The sequence of these three pairs is similar to the two pairs found in the ATCC8482 strain. The other 13 remaining IS21-like elements are paralogs with transposase-ATP-binding protein pairs different from the three previously mentioned pairs. IS elements are reported to copy themselves and to be randomly incorporated at another position in the genome.[66] The result of such copy-paste events might be alternated gene expression or DNA inversions events.[65] Furthermore, it is reported that IS elements occur more often in genomes that are influenced by lower evolutionary pressure and drive reductive evolution.[67] In *Shigella flexneri* it was found that IS elements could be correlated with gene loss and areas of recombination.[67]

Further hints for internal genome diversification were identified with groups of hypothetical proteins and directly neighboring transposases that share different protein homology and are spread over the genome. Incidences of those protein pairs were compared in both *B. vulgatus* mpk and ATCC8482. In *B. vulgatus* mpk we could identify 18 and in ATCC8482 six of such pairs. A ClustalΩ alignment of all 24 hypothetical proteins adjacent to the transposases showed that the proteins can be divided into four different groups. A multiple sequence alignment demonstrates the homology among the four different groups (publication a, figure 4A). The alignment shows that group 1 includes seven proteins from *B. vulgatus* mpk, group 2 contains other seven *B. vulgatus* mpk proteins, group 3 consists of four *B. vulgatus* mpk proteins and two ATCC8482 proteins; group 4 only represents proteins found in *B. vulgatus* ATCC8482. One may argue why IS elements should be copied and spread over genomes since the insertions in the genome would create the risk of changed gene expression. Further, the question is: what could be the driving force for genome rearrangements and resulting fundamental changes in a genome? Pál and Papp report that the action of IS elements and other mutator structures was increased during certain stress conditions.[68] This suggests that these structures might indirectly contribute to bacterial survival under stress conditions.

Integration hotspots along the genome

In a next step the movement of the transposases was visualized to observe any pattern for how or where the transposases might be introduced in the genome (publication a, figure 4B). According to figure 4B the transposases seem to avoid integration into certain parts of the genome. Therefore, we investigated if there are certain rearrangement hotspots in the genome as it was recently proposed for *Staphylococcus aureus*.[69] After correlating protein paralogy and the occurrence and location of the mobile genetic elements, we identified a region with a higher frequency of gene transition sites and density of mobile genes. This region is located between 1,125 and 3,300 kbp on the chromosome (publication a, figure 5). Based on the information that *Bacteroides* incorporated transposases at the 3´ end of leucine or serine tRNA[70, 71], we screened the genome to identify tRNAs adjacent to mobile elements. 17 (out of 79 tRNAs) were associated with an integrase or transposase and eleven of those were located in the 1,125 and 3,300 kbp (publication a, supplementary table S9). Therefore, we suggest a connection of tRNA-mobile

element pair and transition sites; thus we provide evidence for a hotspot region for genome rearrangements.

Outer membrane vesicle (OMV) gene clusters are spread across the B. vulgatus mpk genome

Apart from being genetically versatile, *B. vulgatus* mpk showed another interesting feature: Nine distinct loci for capsular polysaccharide synthesis (CPS) were identified in *B. vulgatus* mpk (publication a, supplementary table S4). These loci are among the most polymorphic regions among *Bacteroides* strains and reported to be quite common *Bacteroides* traits.[72] Polysaccharide A (PSA) of *B. fragilis* NCTC9343 was reported to be packed in OMVs and to exhibit immunomodulatory effects and prevent mice from experimental colitis, encephalomyelitis, an experimental model for human multiple sclerosis and central nervous system demyelinating disease.[73, 74] The *B. fragilis* genome contains eight different gene clusters for CPS (PSA-PSH).[75] Generally, CPS gene loci are comprised of *upcY* and *upcZ* homologs, which are the regulatory components of the locus. Downstream of the regulators the loci encode different glycosyltransferases and various carbohydrate transporters.[72]

Due to the genome data, we propose that *B. vulgatus* mpk might produce also capsular polysaccharides, which are packed into OMVs. Although we conclude from the available genome information from *B. vulgatus* mpk and *B. fragilis* that the composition of the produced polysaccharides must be different because the genes and the order of the gene located in the respective capsular polysaccharide gene clusters in both strains are different. These differences were prompting to identify and characterize the immunomodulatory potential of *B. vulgatus* OMV in more detail.

OMVs are able to induce tolerant CD11c$^+$ cells

Previous work showed that live *B. vulgatus* mpk altered immune responses in an anti-inflammatory manner mainly mediated by tolerant CD11c$^+$ colonic lamina propria cells.[10, 76] These findings match with the typical definition of a semi-mature or tolerant phenotype of CD11c$^+$ cells such as bone marrow-derived dendritic cells (BMDCs): inability to promote pro-inflammatory responses, non-responsiveness towards maturation stimuli or active induction of regulatory T cells.[77] We could demonstrate that even only *B. vulgatus* OMVs and no living bacteria were needed to induce tolerance in CD11c$^+$ cells. Incubation of BMDCs with either live *B. vulgatus* or the

isolated OMVs resulted in similar surface expression of MHCII, CD40, CD80, and CD86. The expression of those surface markers were characteristically lower in the tolerant $CD11c^+$ BMDC phenotype, compared to the mature $CD11c^+$ cells induced after *E. coli*-stimulation (publication b, figure 1D). Further experiments could show that even after a first immune priming event with OMVs or the live bacteria, and a second immune challenge with *E. coli*, the tolerant phenotype of the cells was maintained and MHCII surface expression was lower compared to the *E. coli*-primed and *E. coli*-challenged cells. In addition, OMV or *B. vulgatus* priming prevented from excess TNFα secretion (publication b, figure 1F). Thus, the immunomodulatory effect mediated by *B. vulgatus* OMVs is enough to induce semi-maturation, which prevents from further maturation and pro-inflammatory responses.

OMV-induced immunomodulatory effects are mediated by host TLR2 and TLR4

B. vulgatus OMVs were able to activate the TLR2 receptor in mTLR2 overexpressing human embryonic kidney (HEK) cells and the mTLR4 receptor complex in CD14/mTLR4/MD-2 overexpressing HEK cells, respectively (publication b, figure 2A). *Tlr2*$^{-/-}$, *Tlr4*$^{-/-}$, and *Tlr2*$^{-/-}$x*Tlr4*$^{-/-}$ BMDCs were stimulated with live *B. vulgatus*, *E. coli* or *B. vulgatus* OMVs. MHCII and CD40 surface expression was analyzed, as well as secretion of TNFα and IL-6. The loss of both TLR2 and TLR4 lead to complete impairment of antigen recognition, and low levels of MHCII and CD40 surface expression, and to low levels of secreted TNFα and IL-6 (publication b, figure 2C). *Tlr2*$^{-/-}$ BMDCs still express these maturation-associated surface markers, and pro-inflammatory cytokines initially suggested after HEK stimulation. This suggests that both TLR4 and TLR2 signaling might be required for BMDC semi-maturation. To verify the actual involvement of TLR2 we investigated the role of TLR2 recognition for induction of BMDC semi-maturation. $TLR2^{-/-}$ BMDCs were primed with live *B. vulgatus*, *E. coli* or OMVs for 24 h and challenged for additional 16 h with *E. coli* (publication b, figure 2D). MHCII expression, TNFα and IL-6 secretion was significantly higher in the *E. coli* challenged samples compared to the mock control (publication b, figure 2E). Similar observations could be made in *B. vulgatus*- and OMV-primed and *E. coli*-challenged cells. Thus, TLR2 signaling is required to induce semi-mature BMDCs to prevent pro-inflammatory responses. These results correspond with previously published results of Kasper and colleagues, who showed

that the anti-inflammatory action of PSA-OMV is mediated by host TLR2 and TLR4 receptors.[73, 78]

OMVs are internalized in CD11c$^+$ cells

Intestinal microbes are separated from the gut epithelium by a thick mucus layer.[79] But how can the bacteria influence the host immune system? It would require a sort of intimate contact of at least compounds that pass the mucus layer and cross the epithelia barrier. OMVs might be such “compounds” that are able to cross the epithelial barrier. Therefore the internalization of OMVs into BMDCs was assessed. The uptake of fluorescein isothiocyanate (FITC)-labeled OMVs of cultured BMDCs was monitored at different time points and with two different OMV concentrations (publication b, figure 1G). After an incubation of 30 min the population of CD11c$^+$ and FITC-OMV$^+$ cells increased when incubated with 5 µg/µl OMVs. The internalization of *B. vulgatus* OMVs corresponds with a study, which showed that OMVs of probiotic *E. coli* Nissle 1917 entered a number of epithelial cell lines and exhibited anti-inflammatory effects.[80]

Taken together the general characteristics of the intestinal symbiont *B. vulgatus* mpk are: a genome containing many mobile genetic elements, which lead to genome diversification potentially induced under stressful conditions, and OMVs containing capsular polysaccharides, that are also reported to be supportive to bacterial survival under stress conditions.[81] Further, we could show that *B. vulgatus* mpk OMVs support immune priming in host cells, which lead to cellular homeostasis, and most probably also influence intestinal homeostasis.

2. *Escherichia coli* mpk – model microbe acts as intestinal pathobiont: Survival strategies in the inflamed intestine.

Besides the common and abundant bacterial phylum *Bacteroidetes*, the gut microbiome harbors a very limited amount of bacteria of the phylum *Proteobacteria*. The *Proteobacteria* include species such as *Salmonella* Typhimurium or commensal and pathogenic *E. coli*. Despite its relative low abundance, *Escherichia coli* plays a pivotal role in the manifestation of intestinal disorders such as IBD. In IBD patients, the microbiome undergoes a compositional shift characterized by a decrease in *Bacteroidetes* and *Firmicutes* and a higher abundance of *Proteobacteria* such as *E. coli*.[82, 83] During intestinal homeostasis, *E. coli* is generally a poor gut colonizer but its growth is boosted under inflammatory conditions. Intestinal inflammation decreases the colonization resistance of *E. coli* and other resident microbes leading to *E. coli* expansion.[26] This is due to critical environmental changes during inflammation. During inflammation the host´s immune response produces ROS and RNS and nitric oxide. Nitric oxide and ROS react with each other to form nitrate. Such nitrate oxides provide electron acceptors for anaerobic respiration.[84, 85] Anaerobic respiration provides a competitive advantage for bacteria to consume fermentation end-products or non-fermentable substrates as carbon sources, in particular for facultative anaerobic bacteria. This enables these bacteria to avoid competing for fermentable substrates with obligate anaerobic bacterial members.[28] Similar observations have been made during *Salmonella* infection. It was reported that *Salmonella* gains a growth advantage in the inflamed intestine by the consumption of non-fermentable ethanolamine or 1,2-propanediol, an end-product of fucose and rhamnose fermentation.[30, 31] *Salmonella* is able to grow during anaerobic respiration using tetrathionate as electron acceptor, which finally enables ethanolamine utilization.[30] Interestingly, for 1,2-propanediol utilization anaerobic respiration is not required.[31]

In our study we investigated if a mouse commensal *E. coli* isolate – *E. coli* mpk – is also able to gain a fitness advantage during inflammation by consumption of ethanolamine and 1,2-propanediol. This would explain the increased abundance of *E. coli* during IBD in the intestine and help to understand etiology of IBD, which is still not fully unraveled.

Ethanolamine but not 1,2-propanediol utilization is ubiquitous in different E. coli strains

E. coli mpk, a mouse commensal fecal isolate, was evaluated to be nonpathogenic; however, it was demonstrated to cause intestinal inflammation in *Il2*$^{-/-}$ mono-associated gnotobiotic mice.[10] Furthermore, commensal *E. coli* was found to accumulate during intestinal inflammation.[26 86]

To better characterize *E. coli* mpk we performed whole genome sequencing. Sequencing revealed that the *E. coli* mpk genome harbors an ethanolamine utilization gene cluster (eutRKLCBAHGJENMDTQPS), a 1,2-propanediol degradation cluster (pocRpduFABCDEGHJKLMNOPQSTUV) and contains a fucose and rhamnose degradation pathway, which enables the production of the fermentation end-product 1,2-propanediol. It further contains gene clusters for nitrate respiration (*narGHJI, narZYWV, napFDAGHBC*), which might enable the use of RNS as terminal electron acceptors during anaerobic respiration. Additionally, *E. coli* contains a fumarate reductase as an alternative electron acceptor.[87] However, it does not contain a tetrathionate respiration pathway like *Salmonella* Typhimurium. Further *E. coli* mpk encodes a cobalamin synthesis operon which produces the necessary cofactor for propanediol dehydratase, the initial enzyme of 1,2-propanediol degradation pathway.[88]

We could identify many metabolic pathways in nonpathogenic *E. coli* mpk that are also found in the enteropathogen *Salmonella* Typhimurium. In order to prove that these pathways not only exist in *E. coli* mpk, several other *E. coli* strains were screened for the presence of ethanolamine and 1,2-propanediol utilization pathways. We analyzed *E. coli* isolates from IBD patients (LF82, UM146, NRG857c), commensal *E. coli* strains (K12 MG1655, Nissle 1917, HS, ED1a, NCTC86, SE11, SE15), pathogenic *E. coli* strains (0127:H6 E2348/69, O104:H4 str. C227-11, O104:H4 str. 2011C-3493) and *Salmonella* strains (*Salmonella enterica* serovar Typhimurium LT2, *Salmonella enterica* serovar Typhi, *Salmonella bongori*). The nucleotide sequences of both pathways were extracted if they were present in the genome and a multiple sequence alignment was performed.

A cluster for ethanolamine utilization was identified in all analyzed *E. coli* strains; however the commensal *E. coli* MG1655 strain encoded an ethanolamine degradation pathway interrupted by a prophage. This insertion may have resulted in

a dysfunctional pathway. *S.* Typhimurium and *S.* Typhi also harbor this ethanolamine cluster, whereas no such cluster was identified in *S. bongori*.

In order to analyze how closely related the ethanolamine utilization clusters are among different nonpathogenic and pathogenic strains of the *Enterobacteriaceae* family, a multiple sequence alignment was performed, including the commensal *E. coli* strains, *E. coli* isolates from IBD patients, pathogenic *E. coli* strains, *S.* Typhimurium and *S.* Typhi. The comparison revealed a clustering of the tested strains into three groups (publication c, figure 1A). The first group consists of the *Salmonella* and enterohemorrhagic *E. coli* (EHEC) strains, and *E. coli* UM146 isolated from an IBD patient. *E. coli* mpk and other nonpathogenic *E. coli* strains belong into the second group. A third group consists of additional nonpathogenic *E. coli* strains. It includes *E. coli* NRG857c and LF82, which have been isolated from IBD patients, the probiotic strain *E. coli* Nissle 1917 and *E. coli* ED1a. *E. coli* SE15 cannot be assigned to one of the groups. It is rather located between the two groups consisting of nonpathogenic *E. coli* strains. Interestingly, the two IBD patient isolates – *E. coli* LF82 and NRG857c – are very closely related within a commensal *E. coli* group and *E. coli* mpk is related to nonpathogenic strains. These findings are in accordance with reports that identified commensal *E. coli* to accumulate during intestinal inflammation. The ethanolamine utilization cluster of *E. coli* UM146 another isolate from an IBD patient was identified to be closer related to pathogenic strains although being a commensal. This supports the pathobiotic character of commensals that act “pathogenic” during colitis by similar metabolic properties.

Many different bacterial genomes encode ethanolamine utilization gene clusters especially intestine-associated pathogens such as *Salmonella*, but also commensal *Escherichia* and *Pseudomonas* spp encode this gene cluster. In addition, this gene cluster is encoded in genomes of Gram-positive commensal *Enterococcus* and *Clostridium* spp.[89] The intestine constantly provides an ethanolamine source as a compound of phosphatidylethanolamine, which is part of both mammalian and bacterial cell membranes.[90] Physiological concentrations of ethanolamine in the intestine are between 1 and 2 mM.[91]

We could identify an 1,2-propanediol utilization cluster in the *S. enterica* serovars Typhimurium and Typhi; however an 1,2-propanediol utilization cluster was absent from the *S. bongori* genome. The 1,2-propanediol utilization gene cluster was identified in pathogenic *E. coli* O104:H4 str. 2011C-3493, and *E. coli* NRG857c and

LF82, which have been isolated from IBD patients. This leads to the assumption that the presence of this cluster is not a ubiquitous feature such as the presence of an ethanolamine utilization cluster.

The 1,2-propanediol utilization clusters of the *Salmonella* serovars are closely related (publication c, figure 1B), and the 1,2-propanediol gene cluster in *E. coli* NRG857c and LF82 are also closely related. The cluster of *E. coli* mpk and EHEC strain 2011C-3493 cannot be grouped with IBD patient-related strains or the *Salmonella* 1,2-propandiol utilization cluster.

1,2-propanediol utilization is a common degradation pathway in *S. enterica* serovars.[31] This gene cluster is also encoded in some pathogenic and IBD-related *E. coli* strains. Both ethanolamine and 1,2-propanediol utilization are absent from *S. bongori* the second occurring species of the *Salmonella* genus.[31, 92] Through evolution, the *Salmonella* lineage was reported to diverge from *E. coli* and *S. enterica* was separated from *S. bongori* later on.[92] Interestingly, intensive horizontal gene transfer (HGT) between members of the *Enterobacteriaceae* family is quite common, especially between *S. enterica* and *E. coli.*[93] Thus, we propose that the pathways became part of the *S. enterica* genome during its evolution: *S. bongori* might have lost ethanolamine and 1,2-propanediol utilization pathways after splitting from *E. coli* and *S. enterica* acquired both gene clusters by HGT. Some *E. coli* strains might have lost the 1,2-propanediol degradation cluster since it is not frequently present in *E. coli* strains. Furthermore, this pathway does not represent an advantageous fitness factor when colonizing a homeostatic intestine. We and others suggest that 1,2-propanediol degradation is more prevalent in strains isolated from IBD patients.[31] In order to confirm this observation, more *E. coli* strains from IBD patients, other commensal and pathogenic strains need to be analyzed. The fact that 1,2-propanediol utilization clusters are more common in IBD-associated strains might be due to the characteristic outgrowth of *Enterobacteriaceae,* especially *E. coli* strains during inflammation.[28] Therefore, we assume that only the best adapted *E. coli* strains have a survival and fitness advantage under inflammatory or non-homeostatic conditions in the intestine. It might be possible that such strains are not able to be isolated and identified from faeces of healthy donors as *E. coli* strains are not abundant in human feces.

Ethanolamine and 1,2-propanediol utilization contribute to E. coli in vitro competitiveness

In order to determine if either ethanolamine, 1,2-propanediol utilization or both pathways represent a growth advantage for *E. coli* mpk, we constructed different *E. coli* mpk knockout strains. We constructed (1) a strain lacking the *eutR* gene (*E. coli* Δ*eutR*), which regulates expression of the ethanolamine utilization operon, (2) a knockout strain lacking *pduD* (*E. coli* Δ*pduD*), which is part of the initial degradation enzyme *pduCDE*-encoded 1,2-propanediol dehydratase and (3) a double knockout strain (*E. coli* Δ*eutR* Δ*pduD*) with depleted ethanolamine and 1,2-propanediol degradation pathways.

In order to test whether the growth of the knockout strains was affected under aerobic conditions, bacterial growth was assessed by performing aerobic growth assays with different media. None of the *E. coli* strains provided any growth defects in rich medium (data not shown).

However, all strains including the wild type strain poorly grew in aerobic growth assays using minimal medium supplemented with 5mM ethanolamine or 1,2-propanediol (publication c, supplementary figure S1A and B). Furthermore, we used a minimal medium containing scraped mucus from either the ileum or from caecum and colon to better mimic the available nutrients in the intestine. Growth of all *E. coli* stains was assayed aerobically in a plate reader. Both minimal mucus media provided the strains with enough substrates to grow. The ileum medium allowed for better growth compared to the caecum/colon medium (publication c, supplementary figure S1C and D). Interestingly, both mucus media showed similar growth characteristics of the wild type and knockout strains. The wild type strain started to grow earlier than the knockout strains. *E. coli* Δ*eutR* Δ*pduD* started to grow later compared to the wild type and the single knockout strains; both single knockout strains showed less effective growth than the wild type. In the ileum medium an intermediate growth phenotype of the single knockout strains can be observed, which started to grow later than the wild type but earlier than the double knockout strain.

As we observed different growth characteristics in the mucus media, we decided to assess whether the mucus media allows anaerobic competitive growth analysis *in vitro* to evaluate fitness advantage of all knockout strains compared to the wild type strain. A 1:1 mixture of *E. coli* wild type and a respective mutant was used to inoculate the mucus media. The resulting competitive cultures were incubated

anaerobically and 40 mM potassium nitrate was added as electron acceptor. Bacterial growth was observed after 24 h by determination of the colony-forming units (CFUs) of the wild type and the mutant strain. The strains were discriminated according to their antibiotic resistance. The wild type strain contained a kanamycin resistance and the mutant strains harbored a chloramphenicol resistance. Therefore, the CFUs of the wild type strain were recovered from agar plates containing kanamycin and the mutant strains from agar plates containing chloramphenicol, respectively. The resulting competitive indices (CIs) were determined by the CFUs of the wild type strain divided by the CFUs of the mutant. The competition assays using ileum and caecum/colon medium suggest that aerobic conditions do not resemble conditions in a physiological environment (publication c, figure 2A and B, left panels), because all determined CIs were approximately 1, meaning similar CFUs of wild type and competitor mutant strains have been recovered from the culture. Therefore, we conclude that the strains did not compete with each other.

However, anaerobic growth conditions in the ileum mucus medium seem to provide a more competitive environment, since the CI of wt/ Δ*eutR* Δ*pduD* was significantly higher compared to CI of aerobic wt/ Δ*eutR* Δ*pduD* culture (publication c, figure 2A and B, right panels). The *E. coli* Δ*eutR* Δ*pduD* strain showed a significantly enhanced CI, when the culture was incubated anaerobically without electron acceptor. The single knockouts exhibit no growth advantage whether the electron acceptor was added or not. Similar results were observed in the caecum/colon medium. The assessment of the *in vitro* competition assay allows for a first estimation of the competitiveness of the *E. coli* Δ*eutR* Δ*pduD* strain; however the competitiveness of the single knockout strains cannot be determined.

Ethanolamine and 1,2-propanediol utilization supports E. coli survival during acute intestinal inflammation

To monitor the fitness of *E. coli* wild type and mutant strains under inflammatory conditions *in vivo*, we orally administered C57BL/6 specific-pathogen-free (SPF) mice with a 1:1 mixture of kanamycin-resistant *E. coli* wild type and chloramphenicol-resistant *E. coli* wild type (reference group) or kanamycin-resistant *E. coli* wild type and chloramphenicol-resistant *E. coli* Δ*eutR* Δ*pduD* (experimental group) in the drinking water. As described earlier, the strains were discriminated according to their antibiotic resistance. For evaluation of the reference group and experimental group,

we assessed the CFUs of two differentially labeled wild type strains. Throughout the experiments we compared the recovered number of CFUs of the strains in the intestine i.e. the faeces. Furthermore, we evaluated the fitness of the strains for 7 days during acute dextran sodium sulfate (DSS)-induced inflammation. Stool was collected daily and CFUs of the bacterial strains were determined. Additionally, the progress and degree of intestinal inflammation was assessed by monitoring the bodyweight and the disease activity index (DAI). Three days after start of DSS administration an increase of the total bacterial CFUs was observed (publication c, figure 3A and B), which correlates with decreased bodyweight and increased DAI, therefore indicating the presence of intestinal inflammatory processes (publication c, figure 3C). 5 days after the start of DSS administration total bacterial load decreased in all groups. However, the wt/ Δ*eutR* Δ*pduD*-administered group provided slightly lower CFUs of the knockout strain compared to the wild type strain. This observation is in line with a further increased DAI and the increased incidences of soft stool or diarrhea since diarrhea leads to a flushing out of bacteria, which might explain the detected decreased total bacterial loads. *E. coli* Δ*eutR* Δ*pduD* CFUs continued to decrease until day 6 after start of DSS administration. At the end of the experiment, we observed distinctly higher CFUs of the wild type strain compared to the mutant strain in the experimental group. On the other hand, the wt/ wt-administered group maintained similar CFU levels of both different labeled wild type strains throughout the experiment.

At the end of the experiment, we determined CFUs not only from colonic faeces, but also from ileal and caecal content. Interestingly, we identified differences in the recovered CFUs from each strain, assuming varying competitive conditions in different intestinal sections. In the ileal contents the CI of wt/ wt was approximately 1, and CI of the wt/ Δ*eutR* Δ*pduD* was approximately 60, but the differences of the means were not significant (publication c, figure 3D). Whereas, in the caecal and colonic content we could observe significant differences of CI means from the wt/ wt reference group and the wt/ Δ*eutR* Δ*pduD* group (publication c, figure 3E and F). Therefore, we propose increasing competitive conditions from the ileum to caecum and colon among *E. coli* strains. *E. coli* wild type benefits from ethanolamine and 1,2-propanediol utilization in the inflamed intestine and *E. coli* Δ*eutR* Δ*pduD* was significantly less competitive.

Results obtained from *in vivo* observations of wt/ Δ*eutR* Δ*pduD* competition during acute intestinal inflammation in mice verified a significant growth advantage of the wild type strain. However, the competitive advantage of the remaining single knockout strains remains to be analyzed. For the first time we could show that onset of bacterial competition can be observed initially in the terminal ileum, and is further increased in the caecum. It has been reported that also in Crohn´s disease patients with ileal inflammation the abundance of *Enterobacteriaceae* was observed to be elevated.[94] The CIs observed in the colonic faeces might be the result of competition occurred in previous intestinal segments.

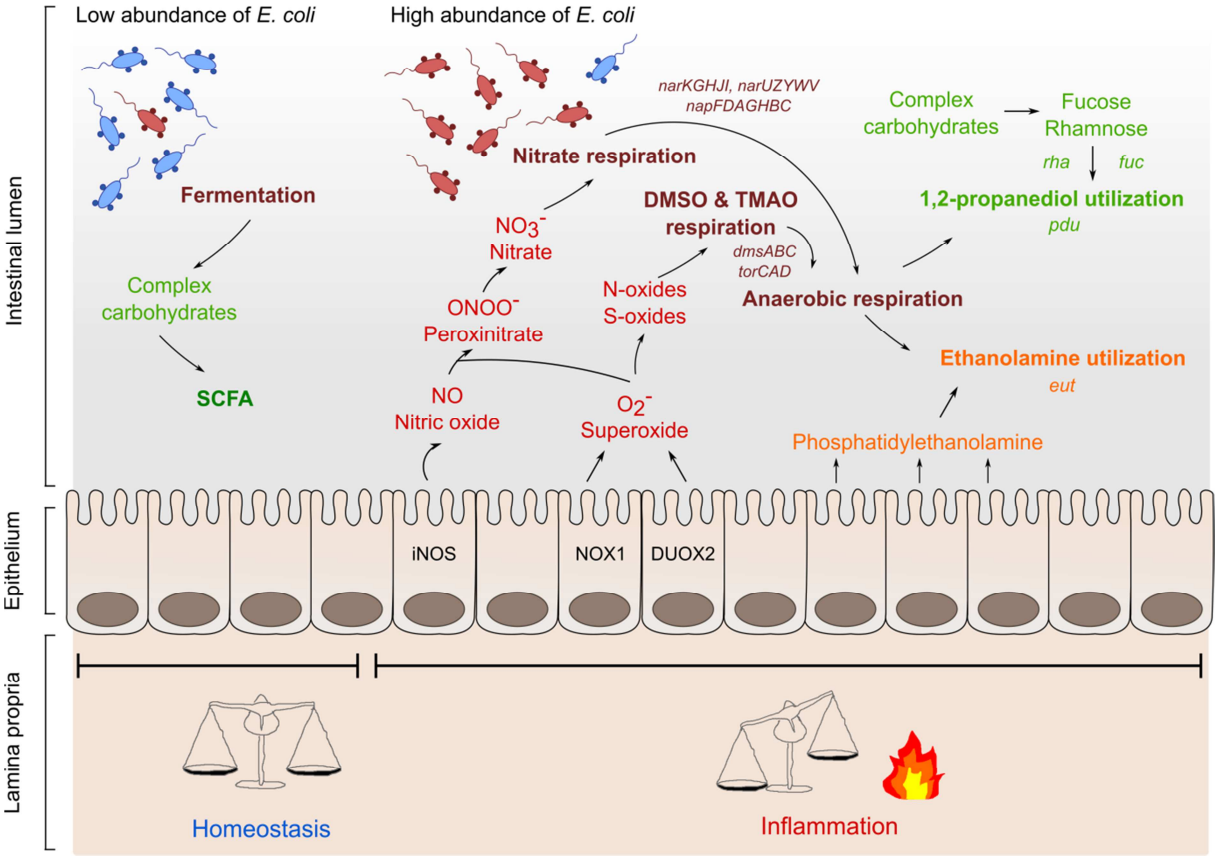


Figure 2: *E. coli* mpk metabolism during intestinal homeostasis and inflammation: In the healthy physiological intestinal environment the abundant obligate anaerobic *Bacteroidetes* and *Firmicutes* produce short-chain fatty acids (SCFAs), which contribute to metabolic exclusion of *Proteobacteria* including *E. coli* (left). Inflammation leads to induced expression of iNOS, NOX1 and DUOX2 enzyme that promote formation of superoxide, nitric oxide and peroxynitrite. These molecules can react and are processed to nitrate, N-oxides or S-oxides, which in turn support the growth and accumulation of *Proteobacteria* such as the facultative anaerobic *E. coli* (right). The changed environmental condition facilitates ethanolamine utilization and enable the degradation of the fermentation end-product 1,2-propanediol.

So far we could find proof for ethanolamine and 1,2-propanediol utilization to provide survival advantage in the inflamed mouse intestine (Figure 2). It remains to be studied whether only one of those pathways confers similar growth benefits compared to *E. coli* Δ*eutR* Δ*pduD*, or if both pathways and a possible switch from one to the other pathways might be beneficial. These observations will help to gain further insights into the involvement of *Enterobacteriaceae* in IBD development.

3. Commensals influence both vertebrate and invertebrate innate immune system similarly

As already mentioned in the two previous chapters, the intestinal microbiota, or rather certain members, play an important role in both maintenance of homeostasis and development of intestinal inflammation in mammals. In this context, gut commensal bacteria can be classified as either symbionts which are beneficial for the host or pathobionts which promote intestinal inflammation in a genetically predisposed organism.[8] Working on this topic usually requires performing experiments using animal models, with rodents being of particular interest. The use of invertebrates however, gives less ethical concerns. In fact, certain invertebrate models such as insect models are currently evaluated for their suitability to replace rodent models, at least for prior screening purposes.

Interestingly, it has been demonstrated that insect intestinal commensal bacteria are also not always symbiotic and beneficial. Recently, it was shown that commensal *Gluconobacter morbifer* provides pathobiotic properties in *Drosophila. G. morbifer* turns “pathogenic” when the commensal microbiota composition is disturbed leading to an increased abundance of this strain. This leads to deregulated intestinal immunity.[37] Hence, insects might be suitable replacement organisms for discriminating intestinal mammalian symbionts from pathobionts.

In order to provide an alternative insect host model, we started to work on a replacement model using the larvae of the greater wax moth *Galleria mellonella* which belongs systematically to the order *Lepidoptera*. In our work we present a comprehensive characterization of different innate immune mechanisms towards symbiotic and pathobiotic intestinal commensals in both *G. mellonella* and the vertebrate mouse model. The *G. mellonella* larvae were orally administered with symbiotic *Bacteroides vulgatus* mpk, mediating a beneficial immunomodulatory effect in mouse models of IBD, hence preventing from intestinal inflammation or the pathobiotic *Escherichia coli* mpk, which ignites inflammation.[10, 11]

Invertebrates and vertebrates share evolutionary conserved components of innate immunity

Previous *in silico* analysis of different components of innate immune responses, such as enzymes being involved in ROS signaling and AMP secretion, verified the close relation between innate mechanisms of vertebrates and invertebrates. We analyzed

sequence homologies between *G. mellonella* markers involved in production of oxidative stress responses (NOS, NOX, and glutathione S-transferases (GST)), LPS recognition (apolipophorin, hemolin) and a defensin-like antimicrobial peptide (gallerimycin) with respective mouse and human proteins. *G. mellonella* sequences were obtained from a published transcriptome dataset and compared to *Mus musculus* and *Homo sapiens* genomes using BLASTX algorithm.[44] The decision whether proteins were potentially homologous was based on published criteria: E-values below 10^{-6} and identity of fragments higher than 30%.[95, 96] One transcript could be identified encoding for NOS in *G. mellonella*. Unfortunately, the full length sequence of *G. mellonella* NOS protein was not available, the query coverage was approximately 90 % and identity greater than 40 %. The transcript reaches E-values of $1x10^{-6}$ when aligned with BLASTX to the mouse genome, and E-values between $1x10^{-7}$ when aligned to the human genome (publication d, table 1). We conclude that the sequences are homologous. Further, we detected one transcript encoding for *G. mellonella* NOX. The BLASTX alignment to the mouse and human genome revealed similarities and resulted in E-values below 10^{-6} (publication d, table 1). However, homology cannot be confirmed resolutely as the query coverage was only 64 % with 34 % identity for mammalian NOX proteins and 62 % query coverage with 25 % identity for mouse and human DUOX proteins.

The transcriptome analysis revealed 19 different available GST variants.[44] We found matches with mammalian GST proteins with E-values below 10^{-6} (publication d, supplementary table S3) after BLASTX alignments, but query coverage was sometimes even lower than 30 %. Only five GST transcripts had query coverage of more than 70 %. Those proteins were found to have reasonable E-values above 10^{-10} and identity of more than 25 %.

In the available transcriptome dataset we could further identify relevant LPS-binding molecules. Specifically, five different transcripts encoding for the LPS recognition molecule apolipophorin were detected as well as two transcripts encoding for hemolin, another LPS-binding opsonin. However, no mammalian homologue could be identified.

The *G. mellonella* defensin-like AMP gallerimycin protein sequence was available from the NCBI database and we observed similarities with mouse and human β-defensin-2 (BD-2) in more detail. We also integrated other insect gallerimycins available from NCBI into our analysis. Different sequences were chosen from other

lepidopteran species. A multiple sequence alignment revealed similar regions of hydrophobic (red) and hydrophilic (blue) areas in all gallerimycin sequences independent from the organism (publication d, supplementary figure S1). This observation suggests that the similar structure accounts for similar functions such as the capability to bind certain interacting molecules. Further, we generated a phylogenetic tree to evaluate relations between the insect gallerimycin protein sequences and the mouse and human BD-2 sequences (publication d, figure 1). Evidently, mouse and human BD-2 peptides showed the closest relation to each other and are also related to the lepidopteran gallerimycin sequences including *G. mellonella* gallerimycin. Additionally, we analyzed sequence similarities of gallerimycin with the antimicrobial cytokine CCL20, because mouse CCL20 shares structural and functional homology with hBD-2.[19] We aligned *G. mellonella* gallerimycin, mBD-2, hBD-2, mouse and human CCL20. Interestingly, as shown in the previous alignment, distribution of similar hydrophobic and hydrophilic amino acids were observed (publication d, supplementary figure S2).

The *in silico* analysis of different components of innate immune responses, such as enzymes relevant in ROS signaling and AMP secretion, verified how closely related innate mechanisms among the vertebrate and insect models are and revealed evolutionary conservation of those genes involved in creating the responses. Therefore, we hypothesize that both vertebrates and insects activate similar components of the innate immune system in response to bacterial exposition.

G. mellonella and mice show similarly increased production of LPS-binding recognition molecules in response to pathobionts

In previous work we showed that the pathobiont *E. coli* mpk induces colitis in *Il2*$^{-/-}$ mice, whereas the symbiont *B. vulgatus* mpk maintains homeostasis.[10] Intestinal commensals can either be symbionts or pathobionts. This distinction becomes crucial in IBD research as IBD comes along with increased abundance of pathobionts.[1] *Bacteroides* species on the other hand are reported to exhibit different immunomodulatory effects.[11, 97] We monitored if the symbiont *B. vulgatus* mpk or the pathobiont *E. coli* mpk were differentially recognized by *G. mellonella* or mouse PRRs. Both bacteria are Gram-negative and contain LPS structures anchored in their outer membrane. Thus, we monitored mRNA expression and/or protein secretion of the LPS-binding molecules apolipophorin (ApoIII) and hemolin in *G. mellonella,* and

CD14 in the mouse model. CD14 is a functionally analogous component of mammalian cells involved in microbe recognition *via* TLR4.[98] In previous work, we already conveyed that *E. coli* and *B. vulgatus* are differentially recognized by TLR4-overexpressing HEK cells. *E. coli* leads to much stronger stimulation of TLR4 than *B. vulgatus* (publication b).

G. mellonella larvae were orally administered with *E. coli* or *B. vulgatus* and RNA was isolated from individual larvae that were incubated for 1 h, 2 h, 3 h, 4 h, 5 h or 6 h. Larvae, which were administered with *E. coli,* provided increased mRNA expression of both *ApoIII* and hemolin genes after 2 h and mRNA expression remained elevated over time, compared to administration with the symbiont *B. vulgatus* (publication d, figure 2A, B). It is reported that *G. mellonella* manages to induce immune responses with distinct intensities in response to different microbes indicating differentially regulated MAMP recognition.[99] However, all available information on microbe recognition in *G. mellonella* was obtained from studies involving systemic hemolymph infection. To date, there is no evidence for differential recognition of Gram-negative microbes *via* the intestine in *G. mellonella*. Interestingly, we also found upregulated hemolin gene expression in response to both *E. coli* and *B. vulgatus* oral administration. Hemolin is abundant in insect hemolymph and interacts with LPS and LTA.[48] Furthermore, hemolin is a member of the immunoglobulin superfamily and is also proposed to function as a recognition molecule.[48] Both apolipophorin and hemolin act as opsonins.[41] The observation that both *E. coli* and *B. vulgatus* were able to induce hemolin mRNA expression suggests that certain bacterial compounds might have crossed the intestinal barrier leading to the involvement of immune responses in the hemolymph.

In order to compare the intestinally-driven bacterial recognition of *E. coli* and *B. vulgatus* in the *G. mellonella* model we decided to analyze bacterial recognition also in mouse intestinal epithelial cells *in vitro*. We stimulated small intestinal mICcl2 cells for 0 h, 4 h, and 24 h with either *E. coli* or *B. vulgatus*. Likewise small intestinal mouse mICcl2 cells stimulated with *E. coli* provided higher *Cd14* gene expression after 4 h (publication d, figure 2C). The corresponding cell culture supernatants of the mICcl2 cells were assayed for secretion of CD14 protein by ELISA. CD14 levels were significantly enhanced in supernatants of *E. coli*-stimulated cells after 24 h compared to *B. vulgatus* stimulation (publication d, figure 2D). Further, we stimulated BMDCs generated from C57BL/6 mice with either *E. coli* or *B. vulgatus* for 0.5 h, 2 h and 4 h

and observed significantly higher gene expression levels of *Cd14* in *E. coli*-stimulated cells compared to *B. vulgatus* stimulation (publication d, figure 2E). Dendritic cells (DCs) have a key role in innate microbe recognition and influence adaptive immune responses.[100] Intestinal DCs are decisive to differentiate between commensals and potential pathogens, and decide on maintenance of homeostasis or active immune responses.[101]

In line with these *in vitro* generated data, we already reported about increased intestinal *Cd14* expression levels in gnotobiotic *Il2*$^{-/-}$ mice mono-associated with *E. coli* compared to gnotobiotic *Il2*$^{-/-}$ mice mono-associated with *B. vulgatus*.[10] In mammals LPS released from Gram-negative bacteria is bound by the extracellular LPS-binding protein and opsonized by CD14 molecule, a surface-expressed opsonic TLR4 co-receptor. The formation of this complex leads to the transfer of LPS to the MD-2 accessory molecule, which is associated with the TLR4 receptor, and finally to oligomerization of the TLR4 receptor, resulting in further immune signaling.[102, 103]

Apart from the recognition of conserved microbial surface structures, insects and mammals share similar effector components that are secreted upon microbial challenge. These components include e.g. ROS/RNS and different AMPs.[37, 47, 104] However, there is only limited research performed in insects focusing on their intestinal immune responses upon microbial challenge. Additionally the majority of information available on insect intestinal immune responses was obtained from *Drosophila*-related research. This is a limitation since *Drosophila*, unlike *G. mellonella*, does not belong to the *Lepidoptera* family. Moreover, *Drosophila* evidently does not induce an antimicrobial response upon a LPS signal[105], which is a very important bacterial factor to be perceived e.g. in the mammalian intestinal innate immune system.

Increased ROS-related gene expression by intestinal commensals in G. mellonella and mice

Rapid environmental changes may force an organism to quickly adapt by triggering a cytoprotective survival program. Such a program may lead to increased production of ROS/RNS causing oxidative stress.[106] In previous work, we demonstrated that stimulation of mouse BMDCs with *E. coli* leads to higher accumulation of intracellular ROS compared to BMDCs stimulated with *B. vulgatus*.[11] In order to test whether different commensals cause different ROS-related gene expression in *G. mellonella*

and gnotobiotic mice, *E. coli* or *B. vulgatus* were orally-administered to both organisms. In addition, small intestinal mICcl2 cells were stimulated with either *E. coli* or *B. vulgatus*. Gene expression of *Nos* and *Nox* was determined from RNA isolated from either whole larval individuals or from ileal and colonic sections of bacteria-administered mono-associated T cell-transplanted *Rag1*$^{-/-}$ mice. The *Rag1*$^{-/-}$ model allows us to investigate innate immune responses. The T cell transfer helps to assess the microbe-driven host responses that depend on both innate and adaptive immune responses to maintain gut homeostasis[107] and allows the differentiation of symbiotic or pathobiotic bacteria.

Oral administration of *E. coli* to *G. mellonella* induced significantly higher *Nos* gene expression compared to *B. vulgatus* administration after 2 h and significant higher *Nox-4* expression after 3 h (publication d, figure 3A, B). After 5 h of bacterial administration, when initial *Nos* and *Nox-4* gene expression levels had decreased, the larvae induced *Gst* gene expression (publication d, figure 3C). It was recently reported that activation of ROS-dependent signaling is involved in the maintenance of intestinal immune homeostasis in *Drosophila*. Xiao and colleagues showed that DUOX-related signaling was able to balance the load of the intestinal microbiota.[104] The DUOX enzyme belongs to the family of NADPH oxidases.[108] Oxidative stress responses were antagonized by induced *Gst* gene expression, which acts as antioxidative molecule.[109, 110] We assumed that the produced ROS and RNS were not intended to substantially harm the host, as the stress response was antagonized to keep cellular homeostasis.

Intestinal infections and inflammation in mammals lead to production of reactive oxygen and nitrogen species.[29] Both ROS and RNS contribute to suppression of microbial growth and intestinal homeostasis in mammals.[109] We were also able to detect elevated *Nos2* gene expression upon microbial challenge in mouse small intestinal mICcl2 cells *in vitro*. The cells were stimulated with *E. coli* and *B. vulgatus* for 0 h, 4 h, and 24 h and *Nos2* expression was determined by quantitative PCR. *Nos2* expression of *E. coli*-stimulated cells was significantly increased after 4 h and 24 h (publication d, figure 3D). *Gstp1* gene expression on the other hand was not regulated at any time point (publication d, figure 3E).

Assessment of *in vivo* relevance of oxidative stress in response to symbiont or pathobiont association was performed in gnotobiotic *Rag1*$^{-/-}$ mice. *Rag1*$^{-/-}$ mice were stably mono-associated with *E. coli* or *B. vulgatus* and subsequently transplanted

with naïve T cells from wild type donor mice. After 4 weeks of T cell transfer, RNA was isolated from ileal and colonic sections and *Nos2* gene expression was determined. *Nos2* gene expression in response to *E. coli*-association was more pronounced compared to *B. vulgatus* association (publication d, figure 3F, G), whereas *Gstp1* expression was not induced (publication d, figure 3H, I). *Nos2* is mainly responsible for nitric oxide production, but it is further able to produce ROS, when concentrations of tetrahydrobiopterin or the co-substrate L-arginine is available in low quantities.[109] *Rag1*$^{-/-}$ mice are a well-established model for the influence of commensal bacteria on intestinal immune balance. The commensal-induced innate immune response differs depending on the activation potential of the commensal and results in proliferation and polarization of T cells, which in turn might limits the commensal induced innate activation.[55] Although the *Rag1*$^{-/-}$ mice were T cell transplanted we observed innate immune responses comparable to that of *G. mellonella*, suggesting that there was no downregulation of the innate response by the transplanted T cells.

Pathobionts induce a stronger antimicrobial response compared to symbionts in both invertebrate and vertebrate hosts

Some bacteria can overcome the oxidative stress responses. However, these bacteria can be eliminated by AMPs, which serve as complementary second line antibacterial defense response.[37, 108] Thus, epithelial cells express AMPs as innate effector molecules to fortify the intestinal barrier and to maintain intestinal homeostasis.[19] Comparable to vertebrates, insects are able to secrete AMPs into the gut lumen. However, and in contrast to vertebrates, insects lack adaptive responses, therefore innate immune mechanisms are the major component of host defense mechanism.[111] Thus, we investigated AMP production in *G. mellonella* in response to oral administration of bacteria. In *G. mellonella,* gene expression of the defensin-like AMP gallerimycin was increased after 4 h and 5 h in response to *E. coli* administration, whereas gene expression in *B. vulgatus*-administered larvae was significantly lower (publication d, figure 4A). Interestingly, we observed a substantially decreasing bacterial load in *G. mellonella* larvae within 6 h after administration (publication d, supplementary figure S3). Additionally, we could also detect increased gene expression of other AMPs in *G. mellonella*, such as gloverin, cecropin, moricin and lysozyme in response to *E. coli* administration compared to *B. vulgatus*

administration (publication d, supplementary figure S4A-D). In summary, commensal-administered *G. mellonella* larvae produced a wide range of different antimicrobial peptides, whereas *E. coli* administration leads to significanty higher AMP production. We analyzed five out of 18 known AMPs from different AMP classes such as defensin-like peptides, cecropins, or moricins which are generally frequently investigated. When the *G. mellonella* immune system is challenged by a certain stimulus, there is always a set of structurally and functionally distinct AMPs which is produced to induce a strong and effective antimicrobial immune response. It was shown that such AMP co-production along with the synergy of AMPs was an effective defense response due to potentiated antimicrobial effects.[112]

Mice and humans also secrete AMPs to maintain intestinal homeostasis. Different classes of antimicrobial peptides such as C-type lectins, cathelicidins or defensins are secreted by mammalian intestinal cells. Cathelicidins and defensins represent the most abundant types of AMPs.[113] We were interested if mammals secrete a comparable pattern of AMPs in response to *B. vulgatus* and *E. coli* challenge as we observed in *G. mellonella*. Intestinal defensins are subdivided into α- and β-defensins. Their expression is, at least in part, dependent on activation of innate PRRs. α-defensins (HD5 and HD6) are produced by specialized Paneth cells. However, β-defensins are produced by a variety of different epithelial cells including enterocytes. Therefore, these β-defensins can be considered one of the most important AMP groups.[114] In humans and mice, BD-1 is constitutively expressed whereas BD-2, BD-3 and BD-4 show increased expression in response to pathogen attack.[19] BD-2 and the antimicrobial chemokine CCL20 share structural homology and are able to bind to the CCR6 receptor.[19, 115] CCR6-CCL20 and CCR6-BD-2 respectively, play an important role in maintenance of intestinal homeostasis and restrict overgrowth of commensal bacteria.[19, 116]

We could observe similar AMP responses in mouse and human cells: we detected higher secretion of CCL20 protein in supernatants of *E. coli*-stimulated mouse small intestinal mICcl2 cells compared to *B. vulgatus*-stimulated cells (publication d, figure 4B). Further we found human colonic epithelial SW480 cells to produce AMPs in response to symbiont or pathobiont stimulation after 0 h, 4 h and 24 h. We could demonstrate that in contrast to *B. vulgatus*, stimulation with *E. coli* leads to secretion of human BD-2 after 24 h in SW480 cell culture supernatants (publication d, figure 4C). Secretion of human CCL20 protein was measured by ELISA. Significantly more

CCL20 was produced after 24 h by *E. coli*-treated cells as compared to *B. vulgatus*-stimulated cells (publication d, figure 4D).

In both animal models, *G. mellonella* and mouse, we showed that *E. coli* induced a higher production of AMPs compared to *B. vulgatus*. In *G. mellonella,* all detected AMPs provided increased gene expression in *E. coli*-administered larvae compared to *B. vulgatus* administration, which indeed points towards a distinct recognition and defined innate immune response in response to distinct Gram-negative commensals. This is in line with previous results obtained with TLR4-overexpressing HEK cells where those bacteria lead to differential stimulation of PRRs and IL-8 secretion (publication b). *In vitro* using mouse small intestinal cells we could determine increased amounts of the antimicrobial cytokine CCL20 after *E. coli* stimulation. We could confirm these results in human epithelial cells and even detect low amounts of the structurally similar AMP hBD-2 after stimulation with *E. coli*. AMP production is TLR-dependent and hBD-2 accumulates after TLR2, TLR3, and TLR4 recognition.[19] Both BD-2 and CCL20 are reported to be important for maintenance of intestinal homeostasis to limit the overgrowth of commensal bacteria.[115, 116] It was shown that both hBD-2 and CCL20 were induced upon Caco-2 cell infection with enteropathogenic *E. coli* and are highly abundant in IBD patients.[19] We propose that the high accumulation of these two types of antimicrobial molecules might also be the consequence of the increased abundance of pathobiotic *E. coli* and its potential to stimulate TLR4. Interpretation of our data generated using the *G. mellonella* model suggests that the monitored immune reactions occur in different sequential timeframes and can be divided into faster and later responses: ROS- and RNS-related gene expression was upregulated quite early between 2 h and 3 h, and antagonized by GST after 5 h, whereas the secretion of AMPs was induced only after oxidative stress responses have been completed. Thus, recognition of the associated bacteria must have occurred previously, since expression of apolipophorin, which recognizes and interacts with LPS structures, was induced already after 2 h.

In summary, we provide evidence for the first time that *G. mellonella* distinctly recognizes different commensals, or their respective MAMPs, by intestinal PRRs finally leading to production of innate effector molecules. This leads in turn to oxidative stress responses and production of AMPs. Further, we could uncover hints for differential intestinal commensal recognition that allow *G. mellonella* for

discrimination of Gram-negative symbionts from pathobionts. Administration of the symbiont *B. vulgatus* leads to less pronounced immune responses (Figure 5, right), whereas administration of *E. coli* initiates stronger immune responses (Figure 5, left). Nevertheless, *E. coli* does not necessarily exhibit pathobiotic features in *G. mellonella*. However, *E. coli* and *B. vulgatus* administration resulted in different immune responses with different intensities therefore confirming different immunogenic potential of the bacteria that was also observed in mice. The findings obtained from insects were therefore comparable to data obtained from vertebrates. We conclude that intestinal innate immune responses among *G. mellonella* and mammals are evolutionary conserved in response to challenge with intestinal commensals. Our results indicate that invertebrates can initiate proper immune responses to excessive loads of microbes and clear bacterial dysbiosis.

Symbionts and pathobionts and their recognition play crucial roles in intestinal homeostasis of IBD[1] and with the invertebrate *G. mellonella* model we provide an alternative model to screen commensals for their immunogenic potential. Hence, we suggest *G. mellonella* as an invertebrate replacement model to discriminate intestinal commensals according to their immunogenic potential. Our future work will aim on the investigation of further microbes and their potential to activate innate immune responses *via* the intestinal tract in the insect model *G. mellonella*.

References

1. Kamada N, Seo SU, Chen GY, Nunez G. Role of the gut microbiota in immunity and inflammatory disease. Nat Rev Immunol 2013; 13:321-35.

2. Ivanov, II, Honda K. Intestinal commensal microbes as immune modulators. Cell host & microbe 2012; 12:496-508.

3. Ottman N, Smidt H, de Vos WM, Belzer C. The function of our microbiota: who is out there and what do they do? Frontiers in cellular and infection microbiology 2012; 2:104.

4. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. Diversity, stability and resilience of the human gut microbiota. Nature 2012; 489:220-30.

5. Cerda B, Perez M, Perez-Santiago JD, Tornero-Aguilera JF, Gonzalez-Soltero R, Larrosa M. Gut Microbiota Modification: Another Piece in the Puzzle of the Benefits of Physical Exercise in Health? Frontiers in physiology 2016; 7:51.

6. Round JL, Mazmanian SK. Inducible Foxp3+ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. Proc Natl Acad Sci U S A 2010; 107:12204-9.

7. Kraal L, Abubucker S, Kota K, Fischbach MA, Mitreva M. The prevalence of species and strains in the human microbiome: a resource for experimental efforts. PLoS One 2014; 9:e97279.

8. Ayres JS. Inflammasome-microbiota interplay in host physiologies. Cell host & microbe 2013; 14:491-7.

9. Chow J, Tang H, Mazmanian SK. Pathobionts of the gastrointestinal microbiota and inflammatory disease. Curr Opin Immunol 2011; 23:473-80.

10. Waidmann M, Bechtold O, Frick JS, Lehr HA, Schubert S, Dobrindt U, Loeffler J, Bohn E, Autenrieth IB. *Bacteroides vulgatus* protects against *Escherichia coli*-induced colitis in gnotobiotic interleukin-2-deficient mice. Gastroenterology 2003; 125:162-77.

11. Steimle A, Gronbach K, Beifuss B, Schafer A, Harmening R, Bender A, Maerz JK, Lange A, Michaelis L, Maurer A, et al. Symbiotic gut commensal bacteria act as host cathepsin S activity regulators. Journal of autoimmunity 2016; 75:82-95.

12. Sassone-Corsi M, Raffatellu M. No vacancy: how beneficial microbes cooperate with immunity to provide colonization resistance to pathogens. J Immunol 2015; 194:4081-7.

13. Kim S, Covington A, Pamer EG. The intestinal microbiota: Antibiotics, colonization resistance, and enteric pathogens. Immunol Rev 2017; 279:90-105.

14. Spees AM, Lopez CA, Kingsbury DD, Winter SE, Baumler AJ. Colonization resistance: battle of the bugs or Menage a Trois with the host? PLoS pathogens 2013; 9:e1003730.

15. Sassone-Corsi M, Nuccio SP, Liu H, Hernandez D, Vu CT, Takahashi AA, Edwards RA, Raffatellu M. Microcins mediate competition among Enterobacteriaceae in the inflamed gut. Nature 2016; 540:280-3.

16. Cerf-Bensussan N, Gaboriau-Routhiau V. The immune system and the gut microbiota: friends or foes? NatRevImmunol 2010; 10:735-44.

17. Sukhithasri V, Nisha N, Biswas L, Anil Kumar V, Biswas R. Innate immune recognition of microbial cell wall components and microbial strategies to evade such recognitions. Microbiological research 2013; 168:396-406.

18. Nell S, Suerbaum S, Josenhans C. The impact of the microbiota on the pathogenesis of IBD: lessons from mouse infection models. NatRevMicrobiol 2010; 8:564-77.

19. Muniz LR, Knosp C, Yeretssian G. Intestinal antimicrobial peptides during homeostasis, infection, and disease. Frontiers in immunology 2012; 3:310.

20. Britanova L, Diefenbach A. Interplay of innate lymphoid cells and the microbiota. Immunol Rev 2017; 279:36-51.

21. Eberl G, Colonna M, Di Santo JP, McKenzie AN. Innate lymphoid cells. Innate lymphoid cells: a new paradigm in immunology. Science 2015; 348:aaa6566.

22. Kamada N, Chen GY, Inohara N, Nunez G. Control of pathogens and pathobionts by the gut microbiota. Nat Immunol 2013; 14:685-90.

23. Vignali DA, Collison LW, Workman CJ. How regulatory T cells work. Nat Rev Immunol 2008; 8:523-32.

24. Cario E. Toll-like receptors in inflammatory bowel diseases: a decade later. Inflamm Bowel Dis 2010; 16:1583-97.

25. Baumgart DC, Carding SR. Inflammatory bowel disease: cause and immunobiology. Lancet 2007; 369:1627-40.

26. Spees AM, Wangdi T, Lopez CA, Kingsbury DD, Xavier MN, Winter SE, Tsolis RM, Baumler AJ. Streptomycin-induced inflammation enhances Escherichia coli gut colonization through nitrate respiration. mBio 2013; 4.

27. Loh G, Blaut M. Role of commensal gut bacteria in inflammatory bowel diseases. Gut microbes 2012; 3:544-55.

28. Winter SE, Baumler AJ. Why related bacterial species bloom simultaneously in the gut: principles underlying the 'Like will to like' concept. Cellular microbiology 2014; 16:179-84.

29. Baumler AJ, Sperandio V. Interactions between the microbiota and pathogenic bacteria in the gut. Nature 2016; 535:85-93.

30. Thiennimitr P, Winter SE, Winter MG, Xavier MN, Tolstikov V, Huseby DL, Sterzenbach T, Tsolis RM, Roth JR, Baumler AJ. Intestinal inflammation allows Salmonella to use ethanolamine to compete with the microbiota. Proc Natl Acad Sci U S A 2011; 108:17480-5.

31. Faber F, Thiennimitr P, Spiga L, Byndloss MX, Litvak Y, Lawhon S, Andrews-Polymenis HL, Winter SE, Baumler AJ. Respiration of Microbiota-Derived 1,2-propanediol Drives Salmonella Expansion during Colitis. PLoS pathogens 2017; 13:e1006129.

32. Kostic AD, Howitt MR, Garrett WS. Exploring host-microbiota interactions in animal models and humans. Genes & development 2013; 27:701-18.

33. Brugman S. The zebrafish as a model to study intestinal inflammation. Developmental and comparative immunology 2016; 64:82-92.

34. Stephens WZ, Burns AR, Stagaman K, Wong S, Rawls JF, Guillemin K, Bohannan BJ. The composition of the zebrafish intestinal microbial community varies across development. The ISME journal 2016; 10:644-54.

35. Broderick NA, Lemaitre B. Gut-associated microbes of Drosophila melanogaster. Gut microbes 2012; 3:307-21.

36. Wong AC, Vanhove AS, Watnick PI. The interplay between intestinal bacteria and host metabolism in health and disease: lessons from Drosophila melanogaster. Disease models & mechanisms 2016; 9:271-81.

37. Kim SH, Lee WJ. Role of DUOX in gut inflammation: lessons from Drosophila model of gut-microbiota interactions. Frontiers in cellular and infection microbiology 2014; 3:116.

38. Wu SC, Cao ZS, Chang KM, Juang JL. Intestinal microbial dysbiosis aggravates the progression of Alzheimer's disease in Drosophila. Nature communications 2017; 8:24.

39. Trevijano-Contador N, Zaragoza O. Expanding the use of alternative models to investigate novel aspects of immunity to microbial pathogens. Virulence 2014; 5:454-6.

40. Bender JK, Wille T, Blank K, Lange A, Gerlach RG. LPS structure and PhoQ activity are important for Salmonella Typhimurium virulence in the Galleria mellonella infection model [corrected]. PLoS One 2013; 8:e73287.

41. Tsai CJ, Loh JM, Proft T. Galleria mellonella infection models for the study of bacterial diseases and for antimicrobial drug testing. Virulence 2016:1-16.

42. Nathan S. New to Galleria mellonella: modeling an ExPEC infection. Virulence 2014; 5:371-4.

43. Kwadha CA, Ong'amo GO, Ndegwa PN, Raina SK, Fombong AT. The Biology and Control of the Greater Wax Moth, Galleria mellonella. Insects 2017; 8.

44. Vogel H, Altincicek B, Glockner G, Vilcinskas A. A comprehensive transcriptome and immune-gene repertoire of the lepidopteran model host Galleria mellonella. BMC genomics 2011; 12:308.

45. Wojda I. Immunity of the greater wax moth Galleria mellonella. Insect science 2016.

46. Huang JH, Jing X, Douglas AE. The multi-tasking gut epithelium of insects. Insect biochemistry and molecular biology 2015; 67:15-20.

47. Buchmann K. Evolution of Innate Immunity: Clues from Invertebrates via Fish to Mammals. Frontiers in immunology 2014; 5:459.

48. Yu XQ, Kanost MR. Binding of hemolin to bacterial lipopolysaccharide and lipoteichoic acid. An immunoglobulin superfamily member from insects as a pattern-recognition receptor. Eur J Biochem 2002; 269:1827-34.

49. Brivio M, Mastore M, Pagani M. Parasite-host relationship a lesson from a professional killer. Invertebrate Survival Journal 2005; 2:41-53.

50. Zug R, Hammerstein P. Wolbachia and the insect immune system: what reactive oxygen species can tell us about the mechanisms of Wolbachia-host interactions. Frontiers in microbiology 2015; 6:1201.

51. Fang FC. Antimicrobial reactive oxygen and nitrogen species: concepts and controversies. Nat Rev Microbiol 2004; 2:820-32.

52. Lu A, Zhang Q, Zhang J, Yang B, Wu K, Xie W, Luan YX, Ling E. Insect prophenoloxidase: the view beyond immunity. Frontiers in physiology 2014; 5:252.

53. Mukherjee K, Altincicek B, Hain T, Domann E, Vilcinskas A, Chakraborty T. Galleria mellonella as a model system for studying Listeria pathogenesis. Appl Environ Microbiol 2010; 76:310-7.

54. Krezdorn J, Adams S, Coote PJ. A Galleria mellonella infection model reveals double and triple antibiotic combination therapies with enhanced efficacy versus a multidrug-resistant strain of Pseudomonas aeruginosa. Journal of medical microbiology 2014; 63:945-55.

55. Gronbach K, Flade I, Holst O, Lindner B, Ruscheweyh HJ, Wittmann A, Menz S, Schwiertz A, Adam P, Stecher B, et al. Endotoxicity of lipopolysaccharide as a determinant of T-cell-mediated colitis induction in mice. Gastroenterology 2014; 146:765-75.

56. Gerritsen J, Smidt H, Rijkers GT, de Vos WM. Intestinal microbiota in human health and disease: the impact of probiotics. Genes & nutrition 2011; 6:209-40.

57. Rodriguez JM, Murphy K, Stanton C, Ross RP, Kober OI, Juge N, Avershina E, Rudi K, Narbad A, Jenmalm MC, et al. The composition of the gut microbiota throughout life, with an emphasis on early life. Microb Ecol Health Dis 2015; 26:26050.

58. Coyne MJ, Zitomersky NL, McGuire AM, Earl AM, Comstock LE. Evidence of extensive DNA transfer between bacteroidales species within the human gut. mBio 2014; 5:e01305-14.

59. Wu S, Powell J, Mathioudakis N, Kane S, Fernandez E, Sears CL. Bacteroides fragilis enterotoxin induces intestinal epithelial cell secretion of interleukin-8 through mitogen-activated protein kinases and a tyrosine kinase-regulated nuclear factor-kappaB pathway. Infect Immun 2004; 72:5832-9.

60. Nguyen M, Vedantam G. Mobile genetic elements in the genus Bacteroides, and their mechanism(s) of dissemination. Mobile genetic elements 2011; 1:187-96.

61. Rogers TE, Pudlo NA, Koropatkin NM, Bell JS, Moya Balasch M, Jasker K, Martens EC. Dynamic responses of Bacteroides thetaiotaomicron during growth on glycan mixtures. Molecular microbiology 2013; 88:876-90.

62. Salyers AA, Shoemaker NB, Li LY. In the driver's seat: the Bacteroides conjugative transposons and the elements they mobilize. J Bacteriol 1995; 177:5727-31.

63. Frost LS, Leplae R, Summers AO, Toussaint A. Mobile genetic elements: the agents of open source evolution. Nat Rev Microbiol 2005; 3:722-32.

64. Bacic M, Parker AC, Stagg J, Whitley HP, Wells WG, Jacob LA, Smith CJ. Genetic and structural analysis of the Bacteroides conjugative transposon CTn341. J Bacteriol 2005; 187:2858-69.

65. Darmon E, Leach DR. Bacterial genome instability. Microbiol Mol Biol Rev 2014; 78:1-39.

66. De Palmenaer D, Siguier P, Mahillon J. IS4 family goes genomic. BMC evolutionary biology 2008; 8:18.

67. Vigil-Stenman T, Larsson J, Nylander JA, Bergman B. Local hopping mobile DNA implicated in pseudogene formation and reductive evolution in an obligate cyanobacteria-plant symbiosis. BMC genomics 2015; 16:193.

68. Pal C, Papp B. From passengers to drivers: Impact of bacterial transposable elements on evolvability. Mobile genetic elements 2013; 3:e23617.

69. Everitt RG, Didelot X, Batty EM, Miller RR, Knox K, Young BC, Bowden R, Auton A, Votintseva A, Larner-Svensson H, et al. Mobile elements drive recombination hotspots in the core genome of Staphylococcus aureus. Nature communications 2014; 5:3956.

70. Wang J, Shoemaker NB, Wang GR, Salyers AA. Characterization of a Bacteroides mobilizable transposon, NBU2, which carries a functional lincomycin resistance gene. J Bacteriol 2000; 182:3559-71.

71. Shoemaker NB, Wang GR, Salyers AA. The Bacteroides mobilizable insertion element, NBU1, integrates into the 3' end of a Leu-tRNA gene and has an integrase that is a member of the lambda integrase family. J Bacteriol 1996; 178:3594-600.

72. Xu J, Mahowald MA, Ley RE, Lozupone CA, Hamady M, Martens EC, Henrissat B, Coutinho PM, Minx P, Latreille P, et al. Evolution of Symbiotic Bacteria in the Distal Human Intestine. PLoSBiol 2007; 5:e156.

73. Shen Y, Giardino Torchia ML, Lawson GW, Karp CL, Ashwell JD, Mazmanian SK. Outer membrane vesicles of a human commensal mediate immune regulation and disease protection. Cell host & microbe 2012; 12:509-20.

74. Wang Y, Telesford KM, Ochoa-Reparaz J, Haque-Begum S, Christy M, Kasper EJ, Wang L, Wu Y, Robson SC, Kasper DL, et al. An intestinal commensal symbiosis factor controls neuroinflammation via TLR2-mediated CD39 signalling. Nature communications 2014; 5:4432.

75. Chatzidaki-Livanis M, Weinacht KG, Comstock LE. Trans locus inhibitors limit concomitant polysaccharide synthesis in the human gut symbiont Bacteroides fragilis. Proc Natl Acad Sci U S A 2010; 107:11976-80.

76. Muller M, Fink K, Geisel J, Kahl F, Jilge B, Reimann J, Mach N, Autenrieth IB, Frick JS. Intestinal colonization of IL-2 deficient mice with non-colitogenic *B. vulgatus* prevents DC maturation and T-cell polarization. PLoSONE 2008; 3:e2376.

77. Steimle A, Frick JS. Molecular Mechanisms of Induction of Tolerant and Tolerogenic Intestinal Dendritic Cells in Mice. Journal of immunology research 2016; 2016:1958650.

78. Jiang F, Meng D, Weng M, Zhu W, Wu W, Kasper D, Walker WA. The symbiotic bacterial surface factor polysaccharide A on Bacteroides fragilis inhibits IL-1beta-induced inflammation in human fetal enterocytes via toll receptors 2 and 4. PLoS One 2017; 12:e0172738.

79. Hansson GC. Role of mucus layers in gut infection and inflammation. Current opinion in microbiology 2012; 15:57-62.

80. Canas MA, Gimenez R, Fabrega MJ, Toloza L, Baldoma L, Badia J. Outer Membrane Vesicles from the Probiotic Escherichia coli Nissle 1917 and the Commensal ECOR12 Enter Intestinal Epithelial Cells via Clathrin-Dependent Endocytosis and Elicit Differential Effects on DNA Damage. PLoS One 2016; 11:e0160374.

81. Schwechheimer C, Kuehn MJ. Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions. Nat Rev Microbiol 2015; 13:605-19.

82. Kaser A, Zeissig S, Blumberg RS. Inflammatory bowel disease. Annu Rev Immunol 2010; 28:573-621.

83. Spor A, Koren O, Ley R. Unravelling the effects of the environment and host genotype on the gut microbiome. Nat Rev Microbiol 2011; 9:279-90.

84. Winter SE, Lopez CA, Baumler AJ. The dynamics of gut-associated microbial communities during inflammation. EMBO reports 2013; 14:319-27.

85. Rivera-Chavez F, Baumler AJ. The Pyromaniac Inside You: Salmonella Metabolism in the Host Gut. Annu Rev Microbiol 2015; 69:31-48.

86. Rakitina DV, Manolov AI, Kanygina AV, Garushyants SK, Baikova JP, Alexeev DG, Ladygina VG, Kostryukova ES, Larin AK, Semashko TA, et al. Genome analysis of E. coli isolated from Crohn's disease patients. BMC genomics 2017; 18:544.

87. Unden G, Bongaerts J. Alternative respiratory pathways of Escherichia coli: energetics and transcriptional regulation in response to electron acceptors. Biochim Biophys Acta 1997; 1320:217-34.

88. Staib L, Fuchs TM. Regulation of fucose and 1,2-propanediol utilization by Salmonella enterica serovar Typhimurium. Frontiers in microbiology 2015; 6:1116.

89. Garsin DA. Ethanolamine utilization in bacterial pathogens: roles and regulation. Nat Rev Microbiol 2010; 8:290-5.

90. Kaval KG, Garsin DA. Ethanolamine Utilization in Bacteria. mBio 2018; 9.

91. Kaval KG, Singh KV, Cruz MR, DebRoy S, Winkler WC, Murray BE, Garsin DA. Loss of Ethanolamine Utilization in Enterococcus faecalis Increases Gastrointestinal Tract Colonization. mBio 2018; 9.

92. Porwollik S, Wong RM, McClelland M. Evolutionary genomics of Salmonella: gene acquisitions revealed by microarray analysis. Proc Natl Acad Sci U S A 2002; 99:8956-61.

93. Gordienko EN, Kazanov MD, Gelfand MS. Evolution of pan-genomes of Escherichia coli, Shigella spp., and Salmonella enterica. J Bacteriol 2013; 195:2786-92.

94. Naftali T, Reshef L, Kovacs A, Porat R, Amir I, Konikoff FM, Gophna U. Distinct Microbiotas are Associated with Ileum-Restricted and Colon-Involving Crohn's Disease. Inflamm Bowel Dis 2016; 22:293-302.

95. Pearson WR. An introduction to sequence similarity ("homology") searching. Current protocols in bioinformatics 2013; Chapter 3:Unit3 1.

96. Lewis AC, Jones NS, Porter MA, Deane CM. What evidence is there for the homology of protein-protein interactions? PLoS computational biology 2012; 8:e1002645.

97. Mazmanian SK, Round JL, Kasper DL. A microbial symbiosis factor prevents intestinal inflammatory disease. Nature 2008; 453:620-5.

98. Zanoni I, Ostuni R, Marek LR, Barresi S, Barbalat R, Barton GM, Granucci F, Kagan JC. CD14 controls the LPS-induced endocytosis of Toll-like receptor 4. Cell 2011; 147:868-80.

99. Mak P, Zdybicka-Barabas A, Cytrynska M. A different repertoire of Galleria mellonella antimicrobial peptides in larvae challenged with bacteria and fungi. Developmental and comparative immunology 2010; 34:1129-36.

100. Hart AL, Al-Hassi HO, Rigby RJ, Bell SJ, Emmanuel AV, Knight SC, Kamm MA, Stagg AJ. Characteristics of intestinal dendritic cells in inflammatory bowel diseases. Gastroenterology 2005; 129:50-65.

101. Ng SC, Kamm MA, Stagg AJ, Knight SC. Intestinal dendritic cells: their role in bacterial recognition, lymphocyte homing, and intestinal inflammation. Inflamm Bowel Dis 2010; 16:1787-807.

102. Mogensen TH. Pathogen recognition and inflammatory signaling in innate immune defenses. Clinical microbiology reviews 2009; 22:240-73, Table of Contents.

103. Aderem A, Ulevitch RJ. Toll-like receptors in the induction of the innate immune response. Nature 2000; 406:782-7.

104. Xiao X, Yang L, Pang X, Zhang R, Zhu Y, Wang P, Gao G, Cheng G. A Mesh-Duox pathway regulates homeostasis in the insect gut. Nature microbiology 2017; 2:17020.

105. Kaneko T, Goldman WE, Mellroth P, Steiner H, Fukase K, Kusumoto S, Harley W, Fox A, Golenbock D, Silverman N. Monomeric and polymeric gram-negative peptidoglycan but not purified LPS stimulate the Drosophila IMD pathway. Immunity 2004; 20:637-49.

106. Muralidharan S, Mandrekar P. Cellular stress response and innate immune signaling: integrating pathways in host defense and inflammation. J Leukoc Biol 2013; 94:1167-84.

107. Sotolongo J, Ruiz J, Fukata M. The role of innate immunity in the host defense against intestinal bacterial pathogens. Current infectious disease reports 2012; 14:15-23.

108. Ha EM, Lee KA, Seo YY, Kim SH, Lim JH, Oh BH, Kim J, Lee WJ. Coordination of multiple dual oxidase-regulatory pathways in responses to commensal and infectious microbes in drosophila gut. Nat Immunol 2009; 10:949-57.

109. Nathan C, Cunningham-Bussel A. Beyond oxidative stress: an immunologist's guide to reactive oxygen species. Nat Rev Immunol 2013; 13:349-61.

110. Mone Y, Monnin D, Kremer N. The oxidative environment: a mediator of interspecies communication that drives symbiosis evolution. Proceedings Biological sciences / The Royal Society 2014; 281:20133112.

111. Ganz T. The role of antimicrobial peptides in innate immunity. Integrative and comparative biology 2003; 43:300-4.

112. Bolouri Moghaddam MR, Tonk M, Schreiber C, Salzig D, Czermak P, Vilcinskas A, Rahnamaeian M. The potential of the Galleria mellonella innate immune system is maximized by the co-presentation of diverse antimicrobial peptides. Biological chemistry 2016; 397:939-45.

113. Agier J, Efenberger M, Brzezinska-Blaszczyk E. Cathelicidin impact on inflammatory cells. Central-European journal of immunology 2015; 40:225-35.

114. Ostaff MJ, Stange EF, Wehkamp J. Antimicrobial peptides and gut microbiota in homeostasis and pathology. EMBO molecular medicine 2013; 5:1465-83.

115. Shiba H, Mouri Y, Komatsuzawa H, Ouhara K, Takeda K, Sugai M, Kinane DF, Kurihara H. Macrophage inflammatory protein-3alpha and beta-defensin-2 stimulate dentin sialophosphoprotein gene expression in human pulp cells. Biochem Biophys Res Commun 2003; 306:867-71.

116. Lin YL, Ip PP, Liao F. CCR6 Deficiency Impairs IgA Production and Dysregulates Antimicrobial Peptide Production, Altering the Intestinal Flora. Frontiers in immunology 2017; 8:805.

List of publications

Accepted

Lange A., Schäfer A., Bender A., Steimle A., Parusel R, Beier S, Frick J.S.: *Galleria mellonella*: a novel invertebrate model to distinguish intestinal symbionts from pathobionts. Frontiers in Immunology 9 (2114) (2018)

Lange A.*, Beier S.*, Huson D.H., Parusel R., Iglauer F., Frick J.-S.: Genome Sequence of *Galleria mellonella* (greater wax moth). Genome Announcement 6: e01220-17 (2018)

Maerz J.K., Steimle A., **Lange A**., Bender A., Fehrenbacher B., Frick J.-S.: Outer membrane vesicles blebbing contributes to *B. vulgatus* mpk-mediated immune response silencing. Gut Microbes 9: 1-12 (2018)

Parusel R., Steimle A., **Lange A**., Schäfer A., Maerz J.K., Bender A., Frick J.-S.: An important question: Which LPS do you use? Virulence 8(8): 1890-1893 (2017)

Steimle A., Gronbach K., Beifuss B., Schäfer A., Harmening R., Bender A., Maerz J.K., **Lange A**., Michaelis L., Maurer A., Menz S., McCoy K., Autenrieth I.B., Kalbacher H., Frick J.-S.: Symbiotic gut commensal bacteria act as host cathepsin S activity regulators. Journal of Autoimmunity 75: 82-95 (2016)

Lange A.*, Beier S.*, Steimle A., Autenrieth I.B., Huson D.H., Frick J.-S.: Extensive Mobilome-Driven Genome Diversification in Mouse Gut-Associated *Bacteroides vulgatus* mpk. Genome Biology and Evolution 8(4): 1197-207 (2016)

Wille T., Wagner C., Mittelstädt W., Blank K., Sommer E., Malengo G., Döhler D., **Lange A.**, Sourjik V., Hensel M., Gerlach R.G.: SiiA and SiiB are novel type I secretion system subunits controlling SPI4-mediated adhesion of *Salmonella enterica*. Cellular Microbiology 16(2):161-78 (2014)

Bender J.K., Wille T., Blank K., **Lange A.**, Gerlach R.G.: LPS structure and PhoQ activity are important for *Salmonella* Typhimurium virulence in the *Galleria mellonella* infection model. PLoS One 8(8): e73287 (2013)

In preparation

Lange A., Tschörner L., Beck C., Beier S., Schäfer A., Frick J.S.: Survival strategies of *E. coli* mpk in the inflamed intestine.

Lists of oral presentations

“Do the larvae of *G. mellonella* serve as an alternative model host to study gut commensals?”

4th International Conference on Model Hosts, Rhodos, Greece, September 2017

“Survival strategies of *Escherichia coli* mpk during intestinal inflammation”

1st Mini-Symposium of the GRK1708 „Stress Induced Metabolic Pathways“, July 2017

“Extensive mobilome-driven genome diversification in mouse gut-associated *Bacteroides vulgatus* mpk”

9. Seeon Conference “Microbiota, Probiotics and Host”, Seeon, Germany, June 2016

“Different survival strategies of gut commensal bacteria in IBD”

65. Jahrestagung der Deutschen Gesellschaft für Hygiene und Mikrobiologie (DGHM) Rostock, Germany, October 2013

List of poster presentations

"*Galleria mellonella*: a novel invertebrate model to distinguish intestinal symbionts from pathobionts"
56. Wissenschaftliche Tagung der Gesellschaft für Versuchstierkunde GV-SOLAS, September 2018

“Etablierung der Larve der großen Wachsmotte *Galleria mellonella* als alternatives Wirtsmodel mit Hilfe der vollständigen Genomsequenz“
55. Wissenschaftliche Tagung der Gesellschaft für Versuchstierkunde GV-SOLAS, September 2017

“How can external stimuli influence genome diversification in mouse gut-associated *Bacteroides vulgatus*?”
10th Seeon Conference “Microbiota, Probiotics and Host”, Seeon, Germany, July 2017

“Extensive mobilome-driven genome diversification in mouse gut-associated *Bacteroides vulgatus* mpk”
2nd Summer School “Microbes, host and infection”, Tübingen, Germany, July 2016

“Is *Galleria mellonella* a useful model to study bacterial influence innate immunity?”
Novel Concepts in Innate Immunity, Tübingen, Germany, September 2015

“Bacterial survival strategies during inflammation”
8th Seeon Conference “Microbiota, Probiotics and Host”, Seeon, Germany, July 2015

“Bacterial influence on gut homeostasis and inflammation”
4th Joint Conference of the Association for General and Applied Microbiology (VAAM) and the Society of Hygiene and Microbiology (DGHM), Dresden, Germany, October 2014

“Bacterial influence on gut homeostasis and inflammation”

7th Seeon Conference “Microbiota, Probiotics and Host”, Seeon, Germany, July 2014

“Different survival strategies of gut commensal bacteria in IBD”

1st International RTG Workshop on “Bacterial Survival Strategies”, September 2013

“Different survival strategies of gut commensal bacteria in IBD”

6th Seeon Conference “Microbiota, Probiotics and Host”, Seeon, Germany, June 2013

Acknowledgements

Als erstes möchte ich mich bei Prof. Dr. Julia-Stefanie Frick bedanken. Vielen Dank für all die Möglichkeiten und Chancen, die ich während meiner Zeit als Doktorandin in deiner Arbeitsgruppe bekommen habe. Danke dir auch für das Vertrauen eigenständig und unabhängig arbeiten zu können, aber auch dafür, dass du mir immer mit jeglicher Unterstützung oder Ideen zur Seite gestanden hast.

Bei Prof. Dr. Andreas Peschel möchte ich mich für die Übernahme des Zweitgutachtens dieser Arbeit danken.

Andi Schäfer vielen lieben Dank für deine exzellente Arbeit, Engagement, Knowhow und die Gespräche abseits der Arbeit.

Liebe Annika Bender, dir danke ich auch für deine Unterstützung und Zusammenarbeit, aber darüber hinaus für die positive Stimmung, die du verbreitet hast und die schönen Mittagspausen.

Alex Steimle, vielen vielen Dank für deine ständige Unterstützung, das viele wissenschaftliche Diskutieren und die vielen neuen Ideen, die dadurch entstanden sind. Vielen lieben Dank außerdem für das viele Korrekturlesen und konstruktive Kritisieren, das hat mir immer sehr weitergeholfen.

Ein besonderer Dank gilt auch allen anderen aktuellen Arbeitsgruppenmitgliedern Raphael, Jan, Lena, Thomas und Vanessa für die schöne und angenehme Arbeitsatmosphäre. Vielen Dank auch an alle ehemaligen AG-Mitgliedern für ihre Unterstützung, vor allem bedanke ich mich aber bei Biggi Beifuss, Sarah Menz, Isabell Flade, Robin Harmening und Christian Beck. Leonie Tschörner, meiner ehemaligen Masterstudentin, möchte ich für ihre wirklich sehr gute Arbeit und ihr Engagement danken.

Sina Beier vielen Dank für die unkomplizierte gute Zusammenarbeit, und für die bioinformatische Expertise über all die Jahre.

Isi, danke für die Mensagänge und alles was dazugehört!

Meinen Eltern danke ich für die bedingungslose Unterstützung und den Rückhalt vor, während und nach meines Studiums und meiner Doktorandenzeit.

Dani, danke dir für deinen immerwährenden Beistand in allen Lebenslagen und für und den Glauben daran, dass ich alles schaffen kann.

Appendix: publications

Accepted publication a

Lange A.*, Beier S.*, Steimle A., Autenrieth I.B., Huson D.H., Frick J.S.: Extensive Mobilome-Driven Genome Diversification in Mouse Gut-Associated *Bacteroides vulgatus* mpk. Genome Biology and Evolution 8(4): 1197-207 (2016)

Extensive Mobilome-Driven Genome Diversification in Mouse Gut-Associated *Bacteroides vulgatus* mpk

Anna Lange[1,†], Sina Beier[2,†], Alex Steimle[1], Ingo B. Autenrieth[1], Daniel H. Huson[2,*], and Julia-Stefanie Frick[1,*]

[1]Interfacultary Institute for Microbiology and Infection Medicine, Department for Medical Microbiology and Hygiene, University of Tübingen, Tübingen, Germany

[2]Algorithms in Bioinformatics, ZBIT Center for Bioinformatics, University of Tübingen, Tübingen, Germany

[†]These authors contributed equally to this work.

*Corresponding author: E-mail: julia-stefanie.frick@med.uni-tuebingen.de; daniel.huson@uni-tuebingen.de.

Accepted: March 19, 2016

Data deposition: This project has been deposited at GenBank under the accession: CP013020.

Abstract

Like many other *Bacteroides* species, *Bacteroides vulgatus* strain mpk, a mouse fecal isolate which was shown to promote intestinal homeostasis, utilizes a variety of mobile elements for genome evolution. Based on sequences collected by Pacific Biosciences SMRT sequencing technology, we discuss the challenges of assembling and studying a bacterial genome of high plasticity. Additionally, we conducted comparative genomics comparing this commensal strain with the *B. vulgatus* type strain ATCC 8482 as well as multiple other *Bacteroides* and *Parabacteroides* strains to reveal the most important differences and identify the unique features of *B. vulgatus* mpk. The genome of *B. vulgatus* mpk harbors a large and diverse set of mobile element proteins compared with other sequenced *Bacteroides* strains. We found evidence of a number of different horizontal gene transfer events and a genome landscape that has been extensively altered by different mobilization events. A CRISPR/Cas system could be identified that provides a possible mechanism for preventing the integration of invading external DNA. We propose that the high genome plasticity and the introduced genome instabilities of *B. vulgatus* mpk arising from the various mobilization events might play an important role not only in its adaptation to the challenging intestinal environment in general, but also in its ability to interact with the gut microbiota.

Key words: Bacteroides vulgatus, mobile elements, genome plasticity, horizontal gene transfer, CRISPR/Cas.

Introduction

The mammalian gut represents a complex and densely populated ecosystem in which two commensal bacterial phyla are predominant: *Firmicutes* and *Bacteroidetes* (Ottman et al. 2012; Kamada et al. 2013; Coyne et al. 2014). The latter are represented by at least 25 different species in the human gut (Coyne et al. 2014). In the mouse gut species diversity of *Bacteroides* might be similar to the human gut as the NCBI-NT (National Center for Biotechnology Information - Nucleotide) database includes similar numbers of *Bacteroides* species that are associated with the mouse host.

Although some *Bacteroides* sp. are obligate pathogens (Thomas et al. 2011), others have potential health benefits for their hosts, for example, *Bacteroides fragilis*, which is able to suppress intestinal inflammatory responses (Mazmanian et al. 2008), or *Bacteroides vulgatus* mpk, which has been shown to prevent *Escherichia coli*-induced colitis in gnotobiotic interleukin-2-deficient ($IL2^{-/-}$) mice (Waidmann et al. 2003) by induction of host anti-inflammatory immune responses (Waidmann et al. 2003; Bohn et al. 2006; Muller et al. 2008). The pool of different *Bacteroides* sp. in the gut ecosystem provides a diverse collection of phenotypic strain variations with various fitness advantages. Besides the abovementioned immunomodulatory functions, *Bacteroides* species have the potential to contain enterotoxins (e.g., in *B. fragilis*), different gene clusters to degrade a variety of dietary polysaccharides (Coyne et al. 2014), and many mobile genetic elements (Nguyen and Vedantam 2011).

In competitive environments like the gastrointestinal tract, commensal microbes are constantly forced to adapt and

survive upon different challenges like competition for nutrients or phage and pathogen attacks. Bacteria which are able to constantly reshape their genome architecture provide improved adaptation to their microenvironment, thus exhibiting a selective survival advantage (Xu et al. 2007). Recently, some evidence was provided that intestinal *Bacteroides* has exchanged DNA among each other within the human gut microbiota. Further it was proposed that some of the genes that might have been acquired contribute to bacterial fitness (Coyne et al. 2014). However, it is of high interest to prove whether the DNA exchange in human and mouse gut microbial communities is similar or significantly different because mice are model organisms to study gut microbiota.

Although *B. vulgatus* is a common constituent in the mouse and human gut microbiota, only few findings on genome composition and horizontal gene transfer (HGT) have been reported (Xu et al. 2007).

Here we report on the high quality genome draft of *B. vulgatus* strain mpk—further referred to as *B. vulgatus* mpk—which was isolated from mouse feces. Hitherto, only one complete genome reference was available for the *B. vulgatus* type strain ATCC 8482, a human isolate (Xu et al. 2007). Besides, several fragmented sequencing projects of other *Bacteroides* strains are available in the databases. We suspected that the high fragmentation rates of these available short-read assemblies are caused by the large amount of mobile elements, which are frequently found in *Bacteroides* genomes, and therefore increasing complexity of the assembly (Kingsford et al. 2010). Thus we decided to use long-read sequencing in order to produce a draft genome of *B. vulgatus* mpk which is suitable for further analysis, including comparative genomics. We achieved this aim (fig. 1) by using Pacific Biosciences (PacBio) SMRT sequencing, error correction using the PBCR pipeline (Koren et al. 2012), and assembly with Celera Assembler (Myers et al. 2000). Here we report on the results of *B. vulgatus* mpk whole-genome sequencing and provide a detailed analysis of the extraordinary repertoire of mobile genetic elements and how these different mobile structures might contribute to HGT and genome evolution. We show that *B. vulgatus* mpk harbors in fact a large number of repetitive sequences in the form of mobile elements like conjugative transposons, insertion sequences (ISs), a transposable bacteriophage, a CRISPR/Cas system, and numerous integrases and transposases. Further we provide evidence for different externally and internally driven genome rearrangements. We consider *B. vulgatus* mpk to be a potent competitor of the mouse microbiota due to its repertoire of mobile genetic elements and the different genome rearrangements we found.

Materials and Methods

Bacteroides vulgatus mpk and ATCC 8482 Growth Conditions

Bacteroides vulgatus mpk was initially isolated from the fecal material of a healthy mouse (Waidmann et al. 2003). *Bacteroides vulgatus* ATCC 8482 (DSM-1447) type strain, which was isolated originally from human feces, was purchased from DSMZ (German Collection of Microorganisms and Cell Cultures, Braunschweig). Both strains were grown on brain heart infusion (BHI) broth or agar plates at 37 °C anaerobically in a growth chamber.

Genomic DNA Isolation and PacBio Sequencing

The genomic DNA was extracted from bacteria grown on BHI agar with Genomic Tip 100/G (Qiagen, Hilden, Germany) according to manufacturer's instructions. DNA concentration was determined fluorometrically in the Tecan Infinite 200 PRO (Tecan, Mainz, Germany) using a Quant-It DNA assay kit (Life Technologies, Darmstadt, Germany). For library preparation at least 300 ng/µl DNA was used. DNA was used to construct a 10-kb insert size library according to PacBio standard protocols. Sequencing of the libraries was performed on five SMRT cells using PacBio RSII with P4-C2 chemistry achieving a genome coverage of 330-fold. PacBio sequencing was performed at Eurofins (Ebersberg, Germany).

Draft Genome Assembly

Quality assessment of the sequences was done using FastQC (Andrews 2010). The reads were self-corrected using the PBCR pipeline (Koren et al. 2012) and downsampled to 25-fold coverage, keeping only the longest corrected reads. Those reads were assembled using Celera Assembler v8.1 (Myers et al. 2000), resulting in 33 contigs. By mapping the full set of corrected reads back to the assembly, 25 small contigs were identified having a 10-fold or less average coverage and those were excluded from further analysis. The remaining 8 contigs were closely checked for misassemblies and manually broken into 15 contigs. Remapping the corrected reads to the 15 curated contigs showed that more than 96% of the data was accounted for. The 15 curated contigs were ordered using the Mauve ContigMover (Rissman et al. 2009) guided by the *B. vulgatus* type strain ATCC 8482 and then manually curated into a 5.1-Mbp-long draft genome. The Guanine-Cytosine content is 42.2%. The sequence is available on NCBI under the accession CP013020.

Functional Annotation

Automated annotation of the draft genome sequence was done using the RAST server (Aziz et al. 2008), additionally using BaSys (Van Domselaar et al. 2005) and xBase

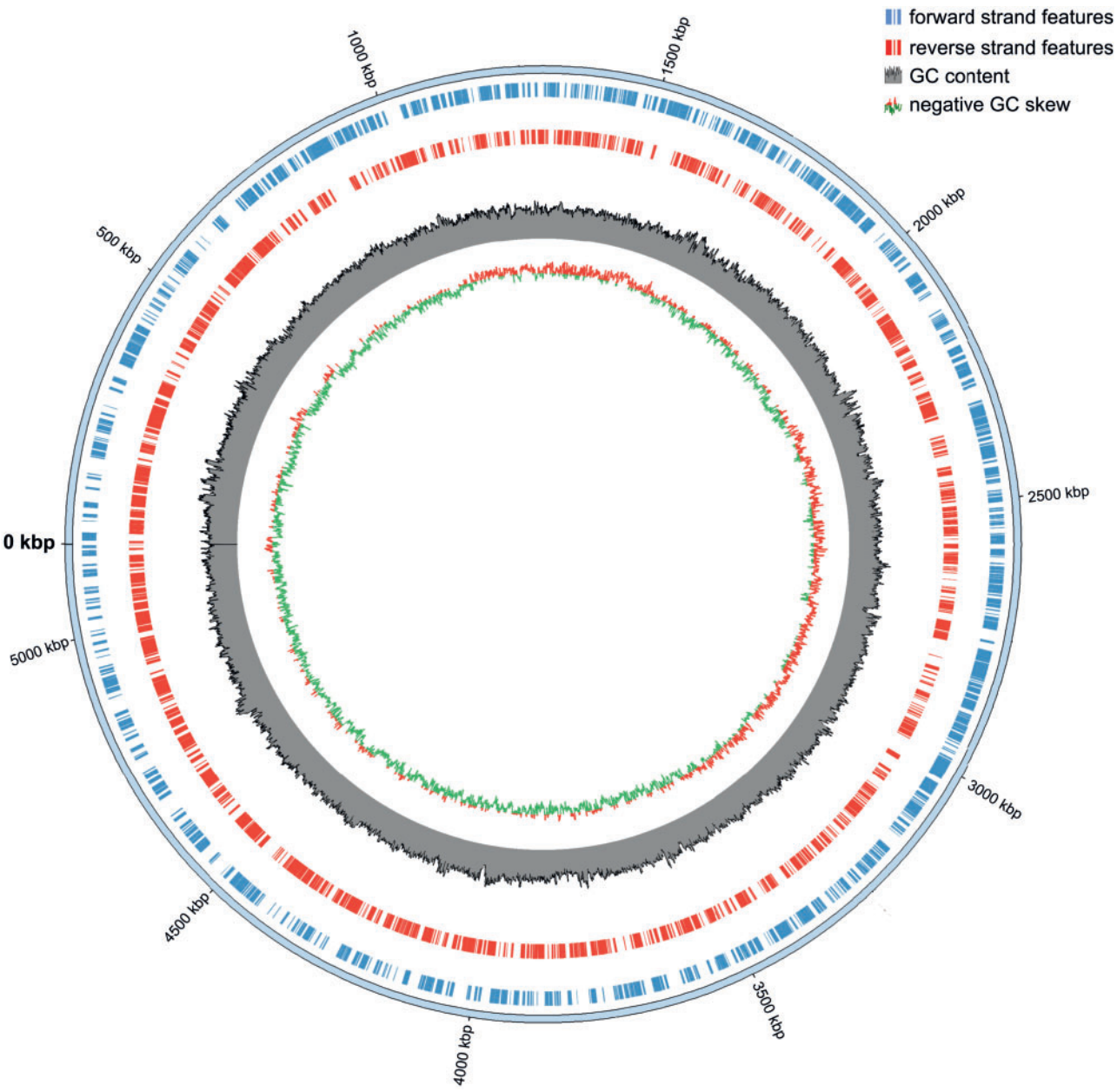


Fig. 1.—Circular representation of the *Bacteroides vulgatus* mpk genome. The forward strand features are shown in blue, reverse strand features in red, Guanine-Cytosine (GC) content in gray, and positive and negative GC skew in red-green.

(Chaudhuri et al. 2008), using the published genome of *B. vulgatus* ATCC 8482 as a reference. The three obtained annotations were merged, selecting the correct annotation by criteria determined by a previous assessment and comparison of the methods. The highest weight in the merging process was given to RAST annotations. The final annotation was curated removing genes following a selection by strong (<150 bp, annotated as hypothetical protein, no protein homology in comparison with selected related strains) and weak criteria (suspicious start codons, overlapping another well-annotated gene, different reading direction compared with all neighboring genes). An annotation was removed if it matched either a total of four criteria or all three of the strong criteria.

Comparative Genomics

Phylogenetic tree construction was based on a multiple sequence alignment of 15 *Bacteroides* species (supplementary table S1, Supplementary Material online) and *Parabacteroides distasonis*, using partial 16S ribosomal Ribonucleic acid (rRNA)

gene sequences acquired from GenBank. The tree was calculated using maximum likelihood with 100 bootstrap replications. The resulting tree was rerooted using *P. distasonis* as an outgroup. For comparative analysis, nine related strains from the order *Bacteroidales* were chosen for protein homology search. Protein sequences for each of those strains were compared using pairwise BLASTP and filtering for matches covering at least 60% of the longer protein sequence. Those potential homologs were filtered for 50%, 90%, and 98% protein sequence identity to increase the specificity of the analysis. Another pairwise BLASTP comparison of the full set of all predicted *B. vulgatus* mpk protein sequences against itself was done to identify paralogs. By identifying colocalized groups of homologous proteins, synteny regions were detected to identify structural differences and be able to screen the genomes for potential HGT (supplementary table S2, Supplementary Material online). Multiple sequence alignments were done using ClustalΩ (Sievers et al. 2011). Visualization of the results used AnnotationSketch (Steinbiss et al. 2009) and Circos (Krzywinski et al. 2009) and RNA structure prediction used the RNAfold server (Lorenz et al. 2011).

Antibiotic Sensitivity Tests

Bacterial suspensions of *B. vulgatus* mpk and ATCC 8482 at the same optical density (McFarland 1) were spread on fresh BHI agar plates and pads containing different antibiotics were put on the plates. Antibiotic susceptibility tests were performed according to the EUCAST criteria. After anaerobic incubation for 2 days at 37 °C, susceptibility or resistance toward the respective antibiotic was evaluated.

Results and Discussion

General Characteristics and Phylogenetic Placement of *Bacteroides vulgatus mpk*

The genome of *B. vulgatus* mpk is 5,165,891 bp in size and does not contain any extrachromosomal structures (fig. 1). Like other members of the *Bacteroides* genus, it encodes a high number of starch utilization systems (sus) for sugar degradation, for example, one SusR-regulated cluster and several SusC- and SusD-like clusters to degrade different sugars (fucose, rhamnose, galactose) (supplementary table S3, Supplementary Material online) (Xu et al. 2003; Martens et al. 2009). The genome contains nine different loci coding for capsular polysaccharides (e.g., glycosyltransferases; supplementary table S4, Supplementary Material online), which are common traits for *Bacteroides* sp. and are the most polymorphic genomic regions (Xu et al. 2007). Furthermore *B. vulgatus* mpk harbors an abundance of structures to mobilize DNA-like IS elements, conjugative transposons, a transposable bacteriophage, various transposases and integrases, and other unassigned mobile element proteins (table 1). Additionally, a CRISPR/Cas system was found. Mobile DNA structures make up 8% of the 4,233 protein-coding genes. *Bacteroides vulgatus* mpk harbors more mobile element genes than *B. vulgatus* ATCC 8482, which encodes 16 integrases, 89 transposases, 1–3 conjugative transposons, 22 IS elements, and 19 mobilization proteins (table 1).

Table 1
Genome Content of *Bacteroides vulgatus* mpk

Feature	
Genome length	5,165,891 bp
GC content	42.24%
Plasmids	None
Protein-coding genes	4,233
tRNAs	79
rRNAs	21 genes in 7 operons
IS elements	IS21-like 16
	IS4-like 14
Conjugative transposons	1 (+ 2 truncated)
Complete Mu phages	1
Transposases (including putative)	92
Integrases	33
Mobilizable proteins	3 pairs BmgA/B
	(and 11 orphan proteins)
Mobile element proteins	17
CRISPR/Cas systems	1 type I-C/Dvulg

For a phylogenetic placement, we compared *B. vulgatus* mpk with different *Bacteroides* isolates from human and porcine fecal samples as well as from a dog oral cavity isolate. As a distant relative, *P. distasonis* (formerly *Bacteroides distasonis*) was used for a phylogenetic placement to emphasize the different classification of the *Bacteroides* and *Parabacteroides* genus (fig. 2). *Bacteroides vulgatus* mpk is very closely related to *B. vulgatus* ATCC 8482 and *Bacteroides dorei* on the basis of 16s rRNA.

An analysis of gene synteny between *B. vulgatus* and close *B. dorei* strains (HS1_L_1_B_010 and HS1_L_3_B_079) indicates that *B. vulgatus* mpk shares multiple regions with *B. dorei* strains including the CRISPR/Cas system, which are missing in *B. vulgatus* ATCC 8482 (supplementary table S2, Supplementary Material online). This leads to the assumption that the relation between *B. vulgatus* mpk and *B. dorei* might be even closer than to *B. vulgatus* ATCC 8482 due to the evolutionary diversification of these species.

External- and Internal-Driven Genome Rearrangements

For *Bacteroides* species it was shown that conjugative transposons play an important role for HGT and hence for genome plasticity (Salyers et al. 1995; Coyne et al. 2014). However, they are also characterized by a high internal mobility of transposable elements as discussed below. Comparative genomics by protein homology showed that most of the sequences specific for *B. vulgatus* mpk compared with *B. vulgatus* ATCC 8482 are comprised of mobile elements. This also

holds true for the comparison of *B. vulgatus* with multiple other *Bacteroides* species.

Horizontally Transferred Genome Evolution

The most prominent mobile element of *B. vulgatus* mpk is the complete conjugative transposon. By protein homology analysis, the *B. vulgatus* mpk complete conjugative transposon region was revealed to have the closest relation to conjugative transposons found in *Bacteroides xylanisolvens* XB1A as well as in *B. dorei* isolate HS1_L_1_B_010 and *B. fragilis* YCH46. It has high similarity to the *Bacteroides* conjugative transposon type CTn341 (Bacic et al. 2005), which was originally isolated from a clinical human *B. vulgatus* isolate.

The *B. vulgatus* type strain ATCC 8482 also includes such conjugative transposon proteins, although it has a conjugative transposon which lacks the important TraP and TraD homologs and another conjugative transposon which additionally lacks TraB and TraE (supplementary table S5, Supplementary Material online). In comparison with the *B. fragilis* CTn341 sequence (AY515263.1), the conjugative transposon region contains an insertion of 11 proteins between the RteA and BmhA region and lacks a protein with significant homology to TetQ (supplementary table S6, Supplementary Material online). This protein is also absent in *B. vulgatus* ATCC 8482, but can be found in both of the compared *B. dorei* genomes. The inserted sequence includes transcriptional regulators as well as one glyxoxalase family protein (*BvMPK_0111*) and a β-lactamase gene *BvMPK_0116*. The insertion is absent from any of the compared *Bacteroides* species, with only *BvMPK_102* to *BvMPK_106* having low homology to proteins in *B. dorei*, *B. xylanisolvens*, and *B. fragilis* YCH46. The inserted proteins between *BvMPK_0101* and *BvMPK_0117* have no paralogs with significant similarity in the *B. vulgatus* mpk genome. The analysis of different antibiotic resistance genes suggests that both *B. vulgatus* mpk and ATCC 8482 encode the same genes for β-lactamases except *BvMPK_0116*, as its homolog is absent in ATCC 8482 (table 2). Testing of different antibiotics revealed that both strains exhibit similar resistance mechanisms against certain antibiotics like different aminoglycoside antibiotics (amikacin, gentamicin, or tobramycin) and some penicillins (penicillin, ampicillin, oxacillin) (supplementary table S7, Supplementary Material online). Only *B. vulgatus* mpk is resistant against different second-, third-, and fourth-generation cephalosporins. This indicates an additional resistance mechanism in *B. vulgatus* mpk mediated by *BvMPK_0116* which might have been acquired by HGT.

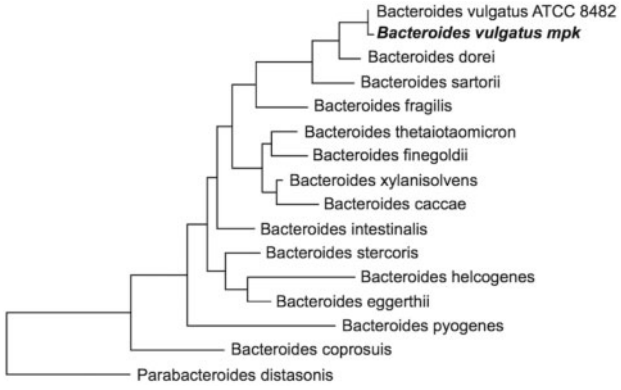


Fig. 2.—Phylogenetic tree of different *Bacteroides* strains. Phylogenetic placement of *Bacteroides vulgatus* mpk, 14 other *Bacteroides* species, and *Parabacteroides distasonis*, based on alignment of 1,488 aligned characters of 16S rRNA and calculated using maximum likelihood. Rooting was performed according to known taxonomy.

We also found evidence of a potential HGT of a genomic island with more than 40 kb in *B. vulgatus* mpk between *BvMPK_2233* and *BvMPK_2278*. The region may have been transferred among *B. vulgatus* mpk, *B. dorei* isolate HS1_L_3_B_079, and *Bacterioides thetaiotaomicron* VPI-5482, because those strains are the only significant hits in the NCBI-NT database filtered for all bacteria and a query coverage over 55%. As no other *B. vulgatus* or *B. dorei* strains include more than 53% percent of this region, the transfer might have occurred in a common ancestor of *B. vulgatus* mpk, *B. dorei*, or *B. thetaiotaomicron* and then got partially lost in all other subsequent strains. The island harbors different mobile element proteins and the encoded genes have a mosaic-like distribution. This random gene arrangement is a quite common characteristic for such genomic islands

Table 2
Genes Coding for β-Lactamases in *Bacteroides vulgatus* mpk and ATCC 8482

Annotation	PFAM Domain	*Bacteroides vulgatus* mpk	*Bacteroides vulgatus* ATCC 8482
AmpG protein, beta-lactamase induction signal transducer	PF07690.11 (MFS_1)	*BvMPK_0055*	*BVU_0613*
Beta-lactamase	PF13354.1 (Beta-lactamase2)	*BvMPK_0116*	Not found
Metal-dependent hydrolases of the beta-lactamase superfamily I	PF12706.2 (Lactamase_B_2)	*BvMPK_1767*	*BVU_2240*
Putative metallo-beta-lactamase	PF00753.22 (Lactamase_B)	*BvMPK_2668*	*BVU_3020*
Metallo-beta-lactamase family protein	PF00753.22 (Lactamase_B)	*BvMPK_2756*	*BVU_3071*
TPR repeat-containing protein	PF08238.7 (Sel1)	*BvMPK_2840*	*BVU_3175*
Beta-lactamase	PF13354.1 (Beta-lactamase2)	*BvMPK_2895*	*BVU_3257*

acquiring genes from possibly different donors (Darmon and Leach 2014).

It is surprising how similar—even on nucleotide level—that particular genomic island is in both mouse *B. vulgatus* mpk and human *B. thetaiotaomicron*. This could point toward a possible transmission route of *Bacteroides* sp. from humans to mice or vice versa, for example, in animal facilities.

The chromosome of *B. vulgatus* mpk contains a region encoding a transposable bacteriophage (Mu phage) absent in all other sequenced *Bacteroides* strains. Their integration into the genome occurs randomly, that is why they might influence downstream genes and operons and are able to inactivate genes. Mu phages are also able to drive further genomic rearrangements and promote mobility of other phages or stimulate recombination events between other transposable elements (Darmon and Leach 2014). Besides the mutative character of such phages, it is also reported that integration can occur simultaneously with the uptake of, for example, toxins which are prevalent in gut viral metagenomic data sets (Minot et al. 2011).

Most HGT events are suggested to be neutral or detrimental and are therefore lost rapidly. Such events just remain within a population if it brings an advantage for the bacteria (Darmon and Leach 2014). We therefore propose that the integration of the additional β-lactamase gene and the transfer of the genomic island should provide a fitness benefit for *B. vulgatus* mpk.

CRISPR/Cas Systems Provide a Potential to Prevent Integration of Invading External Mobile Elements

We have identified a complete type I-C CRISPR/Cas system (fig. 3A), comparable with the one identified in *Bacillus halodurans* C-125 (Nam et al. 2012). The same system is also found in *B. dorei* and in some of the incomplete *B. vulgatus* assemblies, but not in *B. vulgatus* ATCC 8482.

CRISPR/Cas systems are found in about 45% of the sequenced bacterial genomes (Nam et al. 2012). Due to their high prevalence they might be distributed frequently through HGT (Rath et al. 2015). The occurrence of the CRISPR/Cas system in *B. vulgatus* mpk might also be the result of a horizontal transfer event as a mobile element pair is located directly upstream of the CRISPR region (fig. 3A). The CRISPR/Cas system is a sophisticated adaptive immunity tool of both bacteria and archaea (Barrangou and Marraffini 2014; van der Oost et al. 2014). It offers specific immunization against invading mobile genetic elements and is able to integrate nucleic acid fragments into the CRISPR region (Jore et al. 2012). In the original automated annotation of the *B. vulgatus* mpk genes, the CRISPR/Cas seemed to be incomplete, as there were only 13 CRISPR repeats and it lacked the Cas2 protein. These components are essential to match any known functional type of CRISPR/Cas system. As the *B. vulgatus* mpk CRISPR/Cas system resembled the system type I-C (Nam et al. 2012), we manually checked the presence of the Cas2 gene between the repeat region and the Cas1 gene. By sequence comparison, we found a protein which is an exact match to a *Bacteroides* Cas2 (WP_005852931.1) and manually included this annotation.

The I-C/DVULG CRISPR-Cas system is special in having a Cas5d protein, which is an endoribonuclease important for processing pre-crRNA (pre-CRISPR RNA) into its mature form (Nam et al. 2012). This is guided by the stem-loop-containing secondary structure formed by the repeat regions, with Cas5d cleaving at the 3′ end of the hairpin (Nam et al. 2012). The CRISPR region of *B. vulgatus* mpk is made up of thirteen 32 nt repeats and 12 spacer sequences of 33–34 nt. The predicted secondary structure for *B. vulgatus* mpk CRISPR repeats is a

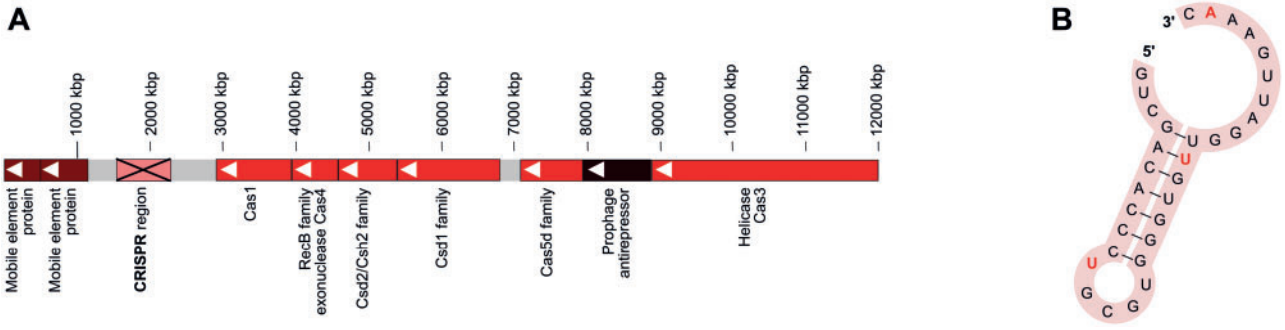


Fig. 3.—The CRISPR/Cas system of *Bacteroides vulgatus* mpk. (*A*) The Cas system is a complete type I-C/DVULG system as postulated in *Bacillus halodurans* C-125. The Cas2 annotation was manually added, as it was missed by the automated annotation pipelines. The prerequisite parts of the CRISPR/Cas type I-C system are shown in the operon: The helicase Cas3, Cas5d which is involved in interacting with pre-crRNA, Csd1 family protein, Csd2/Csh2 familiy protein, Cas4 belonging to the family of RecB exonuclases, and Cas1. The operon is followed downstream by the CRISPR region, and two mobile element proteins.(*B*) The CRISPR repeat secondary structure was predicted from the repeat consensus RNA sequence obtained from the CRISPR region. It shows the conserved hairpin region which is necessary for recognition of the pre-crRNA by Cas5d to enable cleaving to obtain functional crRNA. Red bold letters indicate the nucleotide changes compared with *Bacillus halodurans*.

7-bp stem-loop structure, which has also been described for *Bacillus halodurans* (Nam et al. 2012) (fig. 3B). In comparison with the *Bacillus halodurans* CRISPR-repeat sequence, *Bacillus vulgatus* mpk CRISPR region has two changes inside the stem structure: The G4 is changed to A, and the U8 is changed to C (Nam et al. 2012). In the loop region, the U11 is changed to G and A13 changed to G in the presented system. It was shown that the deletion of the last two nucleotides at the 5′ end at the stem loop had little effect on the functionality of Cas5d (Nam et al. 2012). The last nucleotide of the stem loop is changed in *B. vulgatus* from a U32 to C, while all other bases of the 5′ tail remain conserved. These observations point toward a substrate specificity of Cas5d for the pre-crRNA structure in *B. vulgatus* mpk and provides theoretical evidence for general CRISPR/Cas system functionality. The system therefore might help *B. vulgatus* mpk to prevent invading mobile elements to integrate their DNA into the genome.

The specific function of the CRISPR/Cas type I-C is not well studied. The frequent integration of external DNA in *B. vulgatus* mpk shows that the strain is challenged by external sequences and thus could profit from a mechanism to reduce the challenge to retain stability. However, we can only refer to current literature which did not identify certain function of the CRISPR/Cas type I-C. Most studies about CRISPR/Cas systems are dealing with the system's potential to protect from integration of invading genetic material. Additionally, these systems are reported to be involved in other processes than immunity-like regulation of virulence and DNA repair (Rath et al. 2015) as the system could even influence genome evolution by targeting the host chromosome and creating mutants with broad genome rearrangements (Rath et al. 2015). It remains unclear when *B. vulgatus* mpk acquired the CRISPR/Cas system in its evolutionary process and whether the high prevalence of HGT events can be influenced or regulated by the CRISPR/Cas system.

Internal Genome Evolution

Besides studying the synteny of homologous proteins between *B. vulgatus* mpk and the other selected strains to identify potential HGT events, we also investigated paralogy in *B. vulgatus* mpk to see if mobile elements were transferred within the genome. With that paralogy analysis we could identify two different types of IS elements occurring multiple times in the genome: IS4-like and IS21-like elements (supplementary table S8, Supplementary Material online). The latter type is composed of two open reading frames (ORFs). Those sequences are able to copy themselves and jump randomly into different genomic regions (De Palmenaer et al. 2008). Integrated IS elements can lead to insertion or deletion of genes, it can alternate gene expression, or cause DNA inversions (Darmon and Leach 2014). IS4-like elements encoded by just a single ORF are found in both *B. vulgatus* mpk and ATCC 8482 with about the same number of paralogs. There are 14 copies in *B. vulgatus* mpk and 18 in *B. vulgatus* ATCC 8482. Three of the IS21-like elements (*BvMPK_1852/53*, *BvMPK_1710/11*, *BvMPK_1044/45*) are paralogs having almost identical IS21-like ATP-binding protein and transposon pairs. *BvMPK_1044/45* might be a result of a duplication and transition event. Upstream of the ATP-binding protein we found two proteins (TraG and TraH), which are also part of the conjugative transposon upstream of the *BvMPK_1852/53* pair. So it might be possible that TraG and TraH were copied along with *BvMPK_1044/45*. In *B. vulgatus* ATCC 8482, two homologous pairs of these IS21-like elements can be found (*BVU_0971/72*, *BVU_3472/73*). The other 13 IS21-like gene pairs (*BvMPK_0302/03*, *BvMPK_0856/57*, *BvMPK_1364/65*, *BvMPK_1880/81*, *BvMPK_2015/16*, *BvMPK_2101/02*, *BvMPK_2109/10*, *BvMPK_2229/30*, *BvMPK_2704/05*, *BvMPK_2762/63*, *BvMPK_3201/02*, *BvMPK_3251/52*, *BvMPK_4277/78*) are paralogs with a different ATP-binding protein and transposon. These paralogous gene pairs share very high sequence similarity. Unlike the previous group of IS21-like elements, it is not possible to identify copies that might have been carried with the IS element proteins as they are always located with different genes. In *B. vulgatus* ATCC 8482, this pair is only found once (*BVU_3194/95*). Interestingly, it is reported that IS elements occur more frequently in genomes under low evolutionary pressure and are major drivers for reductive evolution (Vigil-Stenman et al. 2015). IS element clusters could be correlated with areas of recombination and gene losses, for example, in *Shigella flexneri* (Vigil-Stenman et al. 2015).

We further studied other paralogous proteins in the genome and found groups of hypothetical proteins that are all direct neighbors of a transposase and share different levels of protein homology. We checked *B. vulgatus* ATCC 8482 for similar proteins to determine whether they are present with a comparable frequency. We only found six incidences of a comparable hypothetical protein/transposase group in *B. vulgatus* ATCC 8482, while in *B. vulgatus* mpk we found 18 of such pairs. By aligning all 24 hypothetical proteins found in *B. vulgatus* mpk and *B. vulgatus* ATCC 8482 using ClustalΩ, we found that they can be divided into 4 groups. Subsequently, we generated separate multiple sequence alignments for the four groups (fig. 4A).

The group 1 proteins were originally annotated as 1-acyl-sn-glycerol-3-phosphate acyltransferase. This annotation was assigned by RAST based on FIGfam similarity. There is another protein (*BvMPK_1744*) with the annotation of 1-acyl-sn-glycerol-3-phosphate acyltransferase, which is not accompanied by a transposase, but located in a lipoprotein-rich locus. As the other proteins show no significant similarity to this protein, we compared the proteins of group 1 with the NCBI-NR database to verify the annotation. All significant BLASTP hits to group 1 proteins are annotated as hypothetical proteins, with the only exception of a protein from *B. vulgatus* dnLKV7 that is in fact

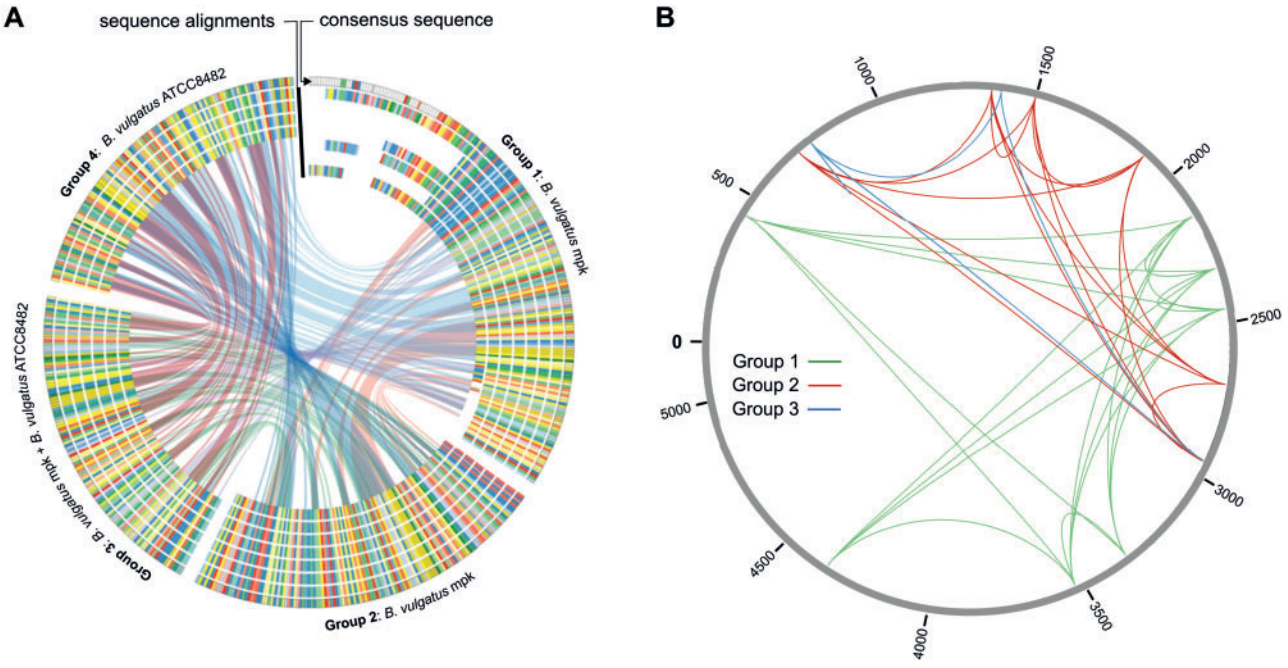


Fig. 4.—Analysis of grouped protein paralogs found in *Bacteroides vulgatus* mpk and *B. vulgatus* ATCC 8482. (*A*) Alignments of hypothetical protein groups found multiple times in the genomes of *B. vulgatus* mpk and *B. vulgatus* ATCC 8482. The similarity of protein used for alignments is 90% at minimum. Group 1 includes seven hypothetical proteins from *B. vulgatus* mpk. Group 2 includes seven closely related hypothetical proteins from *B. vulgatus* mpk. Group 3 includes four proteins from *B. vulgatus* mpk as well as two proteins from *B. vulgatus* ATCC 8482. Group 4 is made up by four proteins found only in *B. vulgatus* ATCC 8482. The outer circle represents the consensus sequence of each group, and the inner circle shows the multiple sequence alignments of the proteins. Regions of exact matches between the consensus sequences are drawn as links between the groups. **(*B*)** Possible movements of the transposons adjacent to group 1 (green), group 2 (red), and group 3 (blue) proteins across the *B. vulgatus* mpk genome.

annotated as 1-acylglycerol-3-phosphate *O*-acyltransferase. As the proteins from group 1 do not match either of the two FIGfams associated with this function (FIG005243 and FIG135282) and the dnLKV7 annotation was done automatically and is uncurated, we classified this as a misannotation and changed the functional annotation to "hypothetical protein." We conclude that *BvMPK_1744* is the only correctly annotated 1-acyl-sn-glycerol-3-phosphate acyltransferase protein in *B. vulgatus* mpk.

We also visualized the distribution of transposons colocated with the paralogous proteins to follow-up their integration across the *B. vulgatus* mpk genome (fig. 4B). The transposons seem to avoid to integrate into certain regions and seem to prefer locations to jump in. Interestingly, the transposons colocating with group 2 and group 3 proteins are randomly located in a certain part of the genome and seem to avoid the other part. We propose that the abovementioned internal transitions influence *B. vulgatus* mpk's genetic repertoire. But we cannot conclude if they caused any positive or negative mutations. The dissemination of mobile elements is expected to produce high mutation rates for large deletions with little counteractive selection, and a massive genome reduction might be the consequence. Accordingly, most changes can rather be put down to genetic drift than adaptation (Moran and Plague 2004).

Area of Mobilization Hotspots in the Genome

Having identified so many mobile elements, paralogous proteins, and genome reformations among the *B. vulgatus* mpk genome, it was prudent to determine whether mobile elements are equally distributed over the genome or if they are concentrated at particular regions in the genome like recently proposed for *Staphylococcus aureus* (Everitt et al. 2014). Correlation of protein paralogy and the location of mobile elements in the genome (fig. 5) reveal that there is in fact an area with a concentration of mobile elements. We identified a region with highly frequent transition sites and a high density of mobile elements. The region is located between 1,125 and 3,300 kbp on the chromosome (fig. 5). It was demonstrated in *Bacteroides* that integration of different transposase proteins was located at the 3′ end of a Leu-tRNA or Ser-tRNA (Shoemaker et al. 1996; Wang et al. 2000). Therefore we screened the genome for tRNA sequences to identify tRNAs associated with mobile elements (supplementary table S9, Supplementary Material online). Out of 79 tRNA sequences, 17 had an associated integrase or transposase. Eleven of those tRNA-mobile element pairs were located in the 1,125–3,300 kbp region. With this we suggest a connection between the increased occurrence of tRNA-mobile element pairs and the colocalization of transition

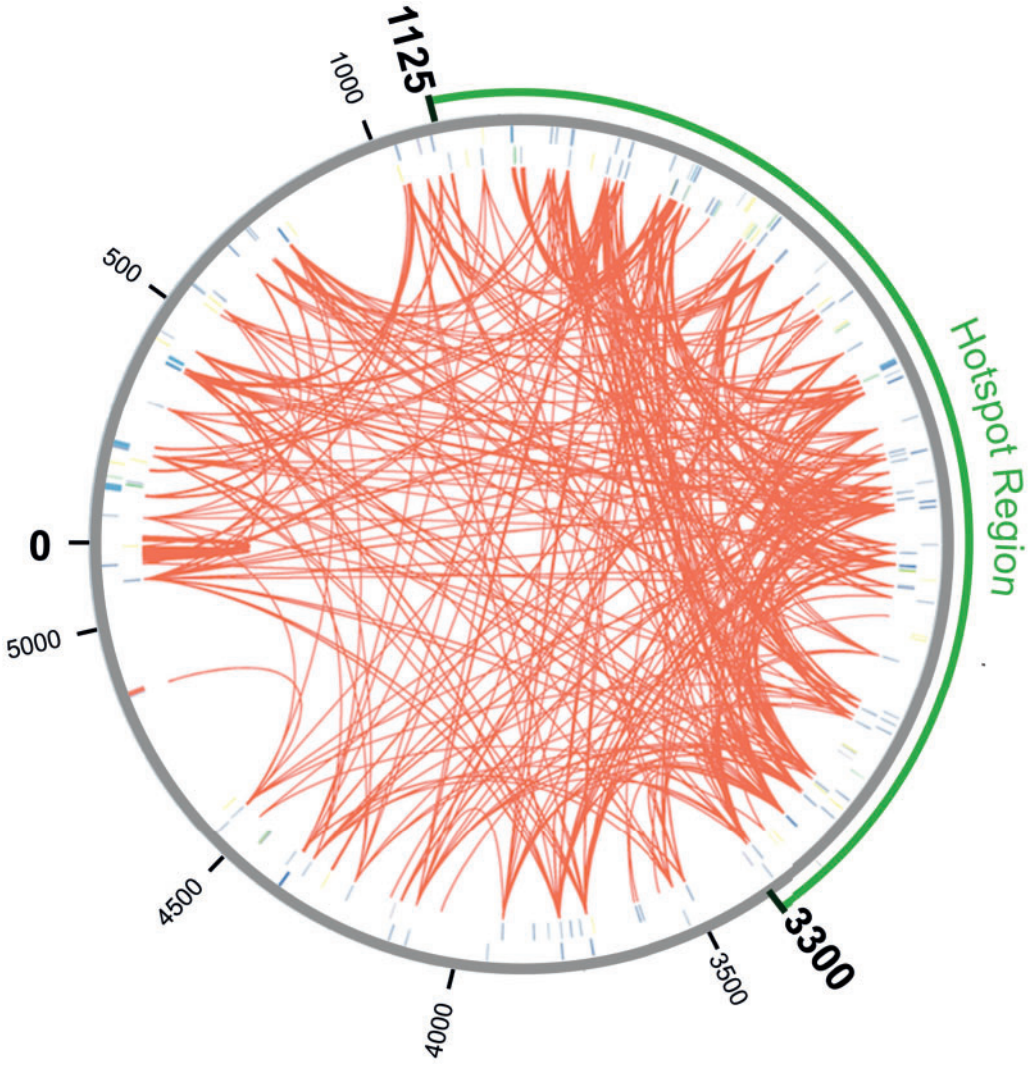


Fig. 5.—A circular view of the *Bacteroides vulgatus* mpk genome with protein paralogy highlights and display of mobile elements. Red lines indicate paralogy based on at least 90% protein identity. Different mobile elements are shown as lines on the forward (inner circle) and reverse (outer circle) strand: Conjugative transposon proteins, CRISPR/Cas-related proteins, integrases, mobilization proteins, transposases, and other mobile elements. The 1,125–3,300 kbp region with highest density of protein paralogy, occurrence of mobile elements, and the incidence of tRNAs associated with mobile elements was highlighted within a green hotspot area.

events. Thus we were able to identify a potential hotspot region for genome transitions and mobile elements. Such regions are areas with high genome instability and are reported to function as a driver for bacterial evolution and adaptation (Darmon and Leach 2014).

Conclusions

Here we provide the second draft genome sequence for *B. vulgatus* mpk generated by PacBio SMRT technology. SMRT sequencing of bacterial genomes for assembly has been proposed as a new standard (Roberts et al. 2013). We and others (Faino et al. 2015; Rhoads and Au 2015) propose it to be a successful technique to sequence genomes with a high content of repetitive sequences. Analysis of the *B. vulgatus* mpk genome revealed a huge variety of mobile elements and different internally and externally driven genome rearrangements have contributed to shape the genome. We found high copy numbers of certain transposase-hypothetical protein pairs in both *B. vulgatus* mpk and *B. vulgatus* ATCC 8482 and propose that these pairs might have evolved simultaneously into four different groups by mostly retaining the full transposase and transposed protein sequence. The finding that these transposase-protein pairs occur in both *B. vulgatus* mpk and *B. vulgatus* ATCC 8482 and that they were spread over both genomes might suggest that both species could use

that strategy to shape their genomes and are at the same time retaining their genome stability.

With this study we suggest that mouse *B. vulgatus* mpk and the human isolate ATCC 8482 might be able to constantly adapt to their environment and benefit from their mobilome, the extensive genome evolution, and the introduced genome instability.

The data presented here might help to improve the understanding of *Bacteroides* sp. genome diversification on a molecular level especially because *Bacteroidetes* represents a highly abundant phylum in the mammalian gut microbiome. The genome databases include only few complete genomes of these abundant members of the gut microbiota and contain only many fragmented genome projects without being very well annotated or studied. The genome data we generated, the comparison with *Bacteroides* genomes isolated from humans, and the insights derived from human gut microbiota studies are so far the only way to generate hypotheses for potential HGT between the mouse commensal strain *B. vulgatus* mpk and other members of the *Bacteroides* genus.

Supplementary Material

Supplementary tables S1–S9 and figure S1 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

This work was supported by the DFG (DFG FR 2087/6-1, DFG FR 2087/8-1, CRC685, SPP1656), the DFG research training group 1708, the Bundesministerium für Bildung und Forschung, and the German Center for Infection Research.

Literature Cited

Andrews S. 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data. [Internet]. [cited 2015 May 6]. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Aziz RK, et al. 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9:75. doi: 10.1186/1471-2164-9-75.

Bacic M, et al. 2005. Genetic and structural analysis of the Bacteroides conjugative transposon CTn341. J Bacteriol. 187:2858–2869.

Barrangou R, Marraffini LA. 2014. CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity. Mol Cell. 54:234–244.

Bohn E, et al. 2006. Host gene expression in the colon of gnotobiotic interleukin-2-deficient mice colonized with commensal colitogenic or noncolitogenic bacterial strains: common patterns and bacteria strain specific signatures. Inflamm Bowel Dis. 12:853–862.

Chaudhuri RR, et al. 2008. xBASE2: a comprehensive resource for comparative bacterial genomics. Nucleic Acids Res. 36:D543–D546.

Coyne MJ, Zitomersky NL, McGuire AM, Earl AM, Comstock LE. 2014. Evidence of extensive DNA transfer between bacteroidales species within the human gut. MBio 5:e01305–e01314.

Darmon E, Leach DR. 2014. Bacterial genome instability. Microbiol Mol Biol Rev. 78:1–39.

De Palmenaer D, Siguier P, Mahillon J. 2008. IS4 family goes genomic. BMC Evol Biol. 8:18.

Everitt RG, et al. 2014. Mobile elements drive recombination hotspots in the core genome of *Staphylococcus aureus*. Nat Commun. 5:3956.

Faino L, et al. 2015. Single-molecule real-time sequencing combined with optical mapping yields completely finished fungal genome. MBio 6: e00936–15.

Jore MM, Brouns SJ, van der Oost J. 2012. RNA in defense: CRISPRs protect prokaryotes against mobile genetic elements. Cold Spring Harb Perspect Biol. 4: a003657.

Kamada N, Chen GY, Inohara N, Nunez G. 2013. Control of pathogens and pathobionts by the gut microbiota. Nat Immunol. 14:685–690.

Kingsford C, Schatz MC, Pop M. 2010. Assembly complexity of prokaryotic genomes using short reads. BMC Bioinformatics 11:21.

Koren S, et al. 2012. Hybrid error correction and de novo assembly of single-molecule sequencing reads. Nat Biotechnol. 30:693–700.

Krzywinski M, et al. 2009. Circos: an information aesthetic for comparative genomics. Genome Res. 19:1639–1645.

Lorenz R, et al. 2011. ViennaRNA Package 2.0. Algorithms Mol Biol. 6:26.

Martens EC, Koropatkin NM, Smith TJ, Gordon JI. 2009. Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm. J Biol Chem. 284:24673–24677.

Mazmanian SK, Round JL, Kasper DL. 2008. A microbial symbiosis factor prevents intestinal inflammatory disease. Nature 453:620–625.

Minot S, et al. 2011. The human gut virome: inter-individual variation and dynamic response to diet. Genome Res. 21:1616–1625.

Moran NA, Plague GR. 2004. Genomic changes following host restriction in bacteria. Curr Opin Genet Dev. 14:627–633.

Muller M, et al. 2008. Intestinal colonization of IL-2 deficient mice with non-colitogenic *B. vulgatus* prevents DC maturation and T-cell polarization. PLoS One 3:e2376.

Myers EW, et al. 2000. A whole-genome assembly of *Drosophila*. Science 287:2196–2204.

Nam KH, et al. 2012. Cas5d protein processes pre-crRNA and assembles into a cascade-like interference complex in subtype I-C/Dvulg CRISPR-Cas system. Structure 20:1574–1584.

Nguyen M, Vedantam G. 2011. Mobile genetic elements in the genus Bacteroides, and their mechanism(s) of dissemination. Mob Genet Elements 1:187–196.

Ottman N, Smidt H, de Vos WM, Belzer C. 2012. The function of our microbiota: who is out there and what do they do? Front Cell Infect Microbiol. 2:104.

Rath D, Amlinger L, Rath A, Lundgren M. 2015. The CRISPR-Cas immune system: biology, mechanisms and applications. Biochimie. 117:119–128.

Rhoads A, Au KF. 2015. PacBio sequencing and its applications. Genomics Proteomics Bioinformatics 13:278–289.

Rissman AI, et al. 2009. Reordering contigs of draft genomes using the Mauve aligner. Bioinformatics 25:2071–2073.

Roberts RJ, Carneiro MO, Schatz MC. 2013. The advantages of SMRT sequencing. Genome Biol. 14:405.

Salyers AA, Shoemaker NB, Li LY. 1995. In the driver's seat: the Bacteroides conjugative transposons and the elements they mobilize. J Bacteriol. 177:5727–5731.

Shoemaker NB, Wang GR, Salyers AA. 1996. The Bacteroides mobilizable insertion element, NBU1, integrates into the 3′ end of a Leu-tRNA gene and has an integrase that is a member of the lambda integrase family. J Bacteriol. 178:3594–3600.

Sievers F, et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 7:539.

Steinbiss S, Gremme G, Scharfer C, Mader M, Kurtz S. 2009. AnnotationSketch: a genome annotation drawing library. Bioinformatics 25:533–534.

Thomas F, Hehemann JH, Rebuffet E, Czjzek M, Michel G. 2011. Environmental and gut bacteroidetes: the food connection. Front Microbiol. 2:93.

van der Oost J, Westra ER, Jackson RN, Wiedenheft B. 2014. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. Nat Rev Microbiol. 12:479–492.
Van Domselaar GH, et al. 2005. BASys: a web server for automated bacterial genome annotation. Nucleic Acids Res. 33:W455–W459.
Vigil-Stenman T, Larsson J, Nylander JA, Bergman B. 2015. Local hopping mobile DNA implicated in pseudogene formation and reductive evolution in an obligate cyanobacteria-plant symbiosis. BMC Genomics 16:193.
Waidmann M, et al. 2003. *Bacteroides vulgatus* protects against *Escherichia coli*-induced colitis in gnotobiotic interleukin-2-deficient mice. Gastroenterology 125:162–177.
Wang J, Shoemaker NB, Wang GR, Salyers AA. 2000. Characterization of a Bacteroides mobilizable transposon, NBU2, which carries a functional lincomycin resistance gene. J Bacteriol. 182:3559–3571.
Xu J, et al. 2003. A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. Science 299:2074–2076.
Xu J, et al. 2007. Evolution of symbiotic bacteria in the distal human intestine. PLoS Biol. 5:e156.

Associate editor: David Bryant

Supplementary Material

Table S1: List of strains used for 16S phylogenetic tree construction

strain	NR database accession
Bacteroides caccae JCM 9498	NR_112932.1
Bacteroides coprosuis JCM 13475	NR_112934.1
Bacteroides dorei 175	NR_041351.1
Bacteroides finegoldii 199	NR_041313.1
Bacteroides fragilis JCM 11019	NR_112936.1
Bacteroides helcogenes JCM 6297	NR_112937.1
Bacteroides intestinalis 341	NR_041307.1
Bacteroides pyogenes JCM 10003	NR_041280.1
Bacteroides sartorii JCM 17136	NR_113195.1
Bacteroides stercoris JCM 9496	NR_112943.1
Bacteroides thetaiotaomicron JCM 5827	NR_112944.1
Bacteroides vulgatus JCM 5826	NR_112946.1
Bacteroides xylanisolvens XB1A	NR_112947.1
Parabacteroides distasonis JCM 5825	NR_041342.1
Bacteroides eggerthii JCM 12986	NR_112935.1

Table S2: Gene synteny between different *Bacteroides* genomes

Excel table online:

https://academic.oup.com/gbe/article/8/4/1197/2574103#supplementary-data

Table S3: SusC/SusD homologs, adjacent genes and ECF-sigma/ anti-sigma factors (if present)

SusC homolog	SusD homolog	Downstream gene	ECF-sigma factor	anti-sigma factor
BvMPK_0003	(BvMPK_0002)			
(BvMPK_0026)	BvMPK_0025	Putative arylsulfate sulfotransferase		
(BvMPK_0044)	BvMPK_0043	Polysaccharidase	BvMPK_0051	BvMPK_0052
BvMPK_0186	(BvMPK_0187)	haloacid dehalogenase-like hydrolase		
(BvMPK_0224)	(BvMPK_0223)			
BvMPK_0239	(BvMPK_0240)	Beta-xylosidase		
(BvMPK_0254) (BvMPK_0252)	(BvMPK_0253) (BvMPK_0251)	Beta-galactosidase	BvMPK_0256	BvMPK_0255
(BvMPK_0263)	(BvMPK_0264)	Arylsulfatase	BvMPK_0261	BvMPK_0262
(BvMPK_0496)	(BvMPK_0495)	Dipeptidyl aminopeptidase/ acylaminoacyl-peptidase	BvMPK_0498	BvMPK_0497
(BvMPK_0508)	(BvMPK_0507)	Aldose 1-epimerase		
BvMPK_0608	BvMPK_0609	Glycerophosphoryl diester phosphodiesterase		
BvMPK_0619	(BvMPK_0620)	Xylanase		
BvMPK_0717	BvMPK_0718	hypothetical protein		
(BvMPK_0754)	(BvMPK_0755)	Choline-sulfatase		
BvMPK_0795	(BvMPK_0796)	Collagenase precursor		
BvMPK_0835	BvMPK_0832	hypothetical protein		
BvMPK_0844	BvMPK_0843	Lipoprotein		
(BvMPK_0871)	(BvMPK_0872)	Alpha-galactosidase		
(BvMPK_0896)	(BvMPK_0897)	Choline-sulfatase	BvMPK_0894	BvMPK_0895
(BvMPK_1024)	(BvMPK_1025)	Mucin-desulfating sulfatase		
BvMPK_1211	(BvMPK_1210)	putative disulphide-isomerase	BvMPK_1213	BvMPK_1212
(BvMPK_1268)	(BvMPK_1269)	hypothetical protein	BvMPK_1266	BvMPK_1267
BvMPK_1610	BvMPK_1611	Outer membrane protein SusF		
(BvMPK_1780)	(BvMPK_1779)	glycoside hydrolase family protein		

(BvMPK_1860) (BvMPK_1863)	(BvMPK_1861) (BvMPK_1862)	Arylsulfatase	BvMPK_1858	BvMPK_1859
BvMPK_2067	BvMPK_2068	thiol:disulfite interchange protein		
(BvMPK_2105) (BvMPK_2106)	(BvMPK_2107)	putative secreted protein	BvMPK_2103	BvMPK_2104
BvMPK_2250	BvMPK_2251	hypothetical protein		
BvMPK_2560 & BvMPK_2561	BvMPK_2562	Histidine Acid Phosphatase	BvMPK_2558	
BvMPK_3665	(BvMPK_3664)	Sialidase		
(BvMPK_3744)	BvMPK_3743	hypothetical protein		
(BvMPK_3855)	(BvMPK_3854)	Pectic Acid Lyase	BvMPK_3857	BvMPK_3856
(BvMPK_3872)	(BvMPK_3871)	hypothetical protein		
(BvMPK_3874)	(BvMPK_3875)	hypothetical protein		
(BvMPK_3920) (BvMPK_3919)	BvMPK_3918	L-fucose mutarotase, type 2		
(BvMPK_4052)	(BvMPK_4051)	glycoside hydrolase family protein	BvMPK_4054	BvMPK_4053
(BvMPK_4144)	(BvMPK_4145) (BvMPK_4146)	Beta-galactosidase	BvMPK_4142	BvMPK_4143
BvMPK_4235 & BvMPK_4236	BvMPK_4234	Alpha-xylosidase		
(BvMPK_4241)	(BvMPK_4242)	glycoside hydrolase family protein		
(BvMPK_4268)	(BvMPK_4267)	hypothetical protein		
BvMPK_4250 (BvMPK_4248)	BvMPK_4249 (BvMPK_4247)	putative chitobiase		
BvMPK_4288	BvMPK_4287	Mucin-desulfating sulfatase		
BvMPK_4311	(BvMPK_4310) (BvMPK_4309)	hypothetical protein		
BvMPK_4321	(BvMPK_4322)	hypothetical protein		

Table S4: Capsular polysaccharide (CPS) loci associated with transcription start sites

Locus 1	
BvMPK_1713	hypothetical protein
BvMPK_1714	hypothetical protein
BvMPK_1715	putative N-acetylmuramoyl-L-alanine amidase
BvMPK_1716	transposase
BvMPK_1717	hypothetical protein
BvMPK_1718	hypothetical protein
BvMPK_1719	putative transmembrane protein
BvMPK_1720	oxidoreductase of aldo/keto reductase family, subgroup 1
BvMPK_1721	hypothetical protein
BvMPK_1722	GCN5-Related N-Acetyltransferase
BvMPK_1723	hypothetical protein
BvMPK_1724	Putative amino acid activating enzyme
BvMPK_1725	3-oxoacyl-[acyl-carrier-protein] synthase, KASIII
BvMPK_1726	putative acyl carrier protein
BvMPK_1727	quinone oxidoreductase
BvMPK_1728	putative transcriptional regulatory protein
Locus 2	
BvMPK_1974	putative transcriptional regulator UpxY-like protein
BvMPK_1975	hypothetical protein
BvMPK_1976	putative LPS biosynthesis related polysaccharidetransporter/flippase
BvMPK_1977	hypothetical protein
BvMPK_1978	hypothetical protein
BvMPK_1979	Glycosyl Transferase Family Protein
BvMPK_1980	hypothetical protein
BvMPK_1981	related to glycosyltransferase
BvMPK_1982	glycosyl transferase family protein
BvMPK_1983	Radical SAM domain heme biosynthesis protein
BvMPK_1984	Alpha-D-GlcNAc alpha-1,2-L-rhamnosyltransferase

BvMPK_1985	hypothetical protein
Locus 3	
BvMPK_2022	Transcription antitermination protein UpdY
BvMPK_2023	hypothetical protein
BvMPK_2024	Polysaccharide export outer membrane protein
BvMPK_2025	Tyrosine-protein kinase Wzc
BvMPK_2026	Tyrosine-protein kinase Wzc
BvMPK_2027	putative flippase
BvMPK_2028	quinone oxidoreductase
BvMPK_2029	hypothetical protein
BvMPK_2030	Galactoside O-acetyltransferase
BvMPK_2031	glycosyl transferase family protein
BvMPK_2032	glycosyl transferase family protein
BvMPK_2033	beta-glycosidase
BvMPK_2034	quinone oxidoreductase
BvMPK_2035	hypothetical protein
BvMPK_2036	N-acetylmannosaminyltransferase
BvMPK_2037	hypothetical protein
BvMPK_2038	Glycosyltransferase
BvMPK_2039	hypothetical protein
BvMPK_2040	hypothetical protein
BvMPK_2041	hypothetical protein
BvMPK_2042	N-acetylmuramoyl-L-alanine amidase
BvMPK_2043	hypothetical protein
BvMPK_2044	hypothetical protein
BvMPK_2045	hypothetical protein
BvMPK_2046	hypothetical protein
Locus 4	
BvMPK_2338	hypothetical protein

BvMPK_2339	Mannose-1-phosphate guanylyltransferase (GDP)
BvMPK_2340	GDP-L-fucose synthetase
BvMPK_2341	GDP-mannose 4,6-dehydratase
BvMPK_2342	Colanic acid biosynthesis glycosyl transferase WcaE
BvMPK_2343	glycosyl transferase family protein
BvMPK_2344	Putative colanic acid biosynthesis acetyltransferasewcaF
BvMPK_2345	Glycosyltransferase
BvMPK_2346	hypothetical protein
BvMPK_2347	glycosyl transferase family protein
BvMPK_2348	Glycosyltransferase
BvMPK_2349	Uncharacterized acetyltransferase MJ1064
BvMPK_2350	putative glycosyltransferase
BvMPK_2351	hypothetical protein
BvMPK_2352	F420H2:quinone oxidoreductase
BvMPK_2353	putative flippase
BvMPK_2354	hypothetical protein
BvMPK_2355	DNA Polymerase Beta Domain Protein Region
BvMPK_2356	Transcription antitermination protein UpdY
BvMPK_2357	hypothetical protein
BvMPK_2358	hypothetical protein
BvMPK_2359	hypothetical protein
BvMPK_2360	N-acetylmuramoyl-L-alanine amidase
BvMPK_2361	Tyrosine-protein kinase Wzc
BvMPK_2362	Tyrosine-protein kinase Wzc
BvMPK_2363	Polysaccharide export outer membrane protein
BvMPK_2364	Glycosyltransferase
Locus 5	
BvMPK_2580	putative transcriptional regulatory protein
BvMPK_2581	hypothetical protein

BvMPK_2582	ISPsy5, transposase
BvMPK_2583	Polysaccharide Biosynthesis Protein
BvMPK_2584	4-hydroxy-2-oxovalerate aldolase
BvMPK_2585	Acylneuraminate cytidylyltransferase
BvMPK_2586	2-dehydro-3-deoxygluconate kinase
BvMPK_2587	2-dehydro-3-deoxygluconate kinase
BvMPK_2588	glycosyl transferase family protein
BvMPK_2589	glycosyl transferase family protein
BvMPK_2590	exopolysaccharide polymerization protein
BvMPK_2591	glycosyl transferase family protein
BvMPK_2592	Glycosyltransferase
BvMPK_2593	UDP-galactopyranose mutase
BvMPK_2594	UDP-glucose 4-epimerase
BvMPK_2595	hypothetical protein
BvMPK_2596	Mobile element protein
BvMPK_2597	N-acetylmuramoyl-L-alanine amidase
BvMPK_2598	hypothetical protein
BvMPK_2599	hypothetical protein
BvMPK_2600	hypothetical protein
BvMPK_2601	hypothetical protein
Locus 6	
BvMPK_3010	hypothetical protein
BvMPK_3011	N-acetylmannosaminyltransferase
BvMPK_3012	glycosyl transferase family protein
BvMPK_3013	hypothetical protein
BvMPK_3014	putative glycosyltransferase protein
BvMPK_3015	Beta-1,3-glucosyltransferase
BvMPK_3016	hypothetical protein
BvMPK_3017	Alpha-1,2-fucosyltransferase

BvMPK_3018	Glycosyl transferase, family 2
BvMPK_3019	Transposase
BvMPK_3020	hypothetical protein
BvMPK_3021	putative LPS biosynthesis related polysaccharidetransporter/flippase
BvMPK_3022	hypothetical protein
BvMPK_3023	hypothetical protein
BvMPK_3024	Capsular polysaccharide transcription antiterminationprotein,UpxY family
Locus 7	
BvMPK_3051	hypothetical protein
BvMPK_3052	hypothetical protein
BvMPK_3053	hypothetical protein
BvMPK_3054	N-acetylmuramoyl-L-alanine amidase
BvMPK_3055	putative mannosyltransferase
BvMPK_3056	hypothetical protein
BvMPK_3057	Glycosyltransferase
BvMPK_3058	Glycosyltransferase
BvMPK_3059	Glycosyltransferase
BvMPK_3060	Hemolysin hemolytic protein
BvMPK_3061	3-oxoacyl-[acyl-carrier protein] reductase
BvMPK_3062	Long-chain-fatty-acid--CoA ligase
BvMPK_3063	putative acyl carrier protein
BvMPK_3064	3-oxoacyl-[acyl-carrier-protein] synthase, KASIII
BvMPK_3065	3-oxoacyl-[acyl-carrier protein] reductase
BvMPK_3066	putative acyl carrier protein
BvMPK_3067	Acetyltransferase (isoleucine patch superfamily)
BvMPK_3068	putative aminotransferase
BvMPK_3069	Lipopolysaccharide biosynthesis protein WzxC
BvMPK_3070	Teichuronic acid biosynthesis protein tuaB
BvMPK_3071	putative transcriptional regulator UpxY-like protein

Locus 8	
BvMPK_3427	hypothetical protein
BvMPK_3428	hypothetical protein
BvMPK_3429	hypothetical protein
BvMPK_3430	hypothetical protein
BvMPK_3431	N-acetylmuramoyl-L-alanine amidase
BvMPK_3432	hypothetical protein
BvMPK_3433	DNA Primase/Helicase
BvMPK_3434	hypothetical protein
BvMPK_3435	putative nucleic acid-binding protein
BvMPK_3436	hypothetical protein
BvMPK_3437	hypothetical protein
BvMPK_3438	N-acetylmannosaminyltransferase
BvMPK_3439	glycosyl transferase family protein
BvMPK_3440	Alpha-1 2-Fucosyltransferase
BvMPK_3441	Putative acetyltransferase
BvMPK_3442	Glycosyltransferase
BvMPK_3443	hypothetical protein
BvMPK_3444	hypothetical protein
BvMPK_3445	Putative glycosyltransferase
BvMPK_3446	Putative glycosyltransferase epsH
BvMPK_3447	O-antigen flippase Wzx
BvMPK_3448	hypothetical protein
BvMPK_3449	putative nucleic acid-binding protein
BvMPK_3450	putative transcriptional regulator UpxY-like protein
Locus 9	
BvMPK_3543	hypothetical protein
BvMPK_3544	hypothetical protein
BvMPK_3545	hypothetical protein

BvMPK_3546	hypothetical protein
BvMPK_3547	glycosyl transferase family protein
BvMPK_3548	glycosyl transferase family protein
BvMPK_3549	hypothetical protein
BvMPK_3550	capsular polysaccharide biosynthesis protein
BvMPK_3551	hypothetical protein
BvMPK_3552	Mobile element protein
BvMPK_3553	putative flippase
BvMPK_3554	oxidoreductase of aldo/keto reductase family, subgroup 1
BvMPK_3555	3-oxoacyl-[acyl-carrier-protein] synthase, KASIII
BvMPK_3556	3-oxoacyl-[acyl-carrier-protein] synthase, KASIII
BvMPK_3557	3-oxoacyl-[acyl-carrier-protein] reductase
BvMPK_3558	putative acyl carrier protein
BvMPK_3559	Galactoside O-acetyltransferase
BvMPK_3560	hypothetical protein
BvMPK_3561	hypothetical protein
BvMPK_3562	hypothetical protein
BvMPK_3563	hypothetical protein
BvMPK_3564	hypothetical protein
BvMPK_3565	DNA-binding protein, histone-like family
BvMPK_3566	N-acetylmuramoyl-L-alanine amidase
BvMPK_3567	putative transcriptional regulator UpxY-like protein

Table S5: Clusters of synteny in *B. vulgatus* ATCC 8482 in comparison to *B. fragilis* CTn341

ctn341	Cluster 1	Cluster 1 similarity	Cluster 2	Cluster 2 similarity	Cluster 3	Cluster 3 similarity	Cluster 4	Cluster 4 similarity
Ctn001								
Ctn002							BVU_3399	51.45
Ctn003							BVU_3400	36.64
Ctn004								
yahA								
lys	BVU_0659	54.07	BVU_1573	54.81				
traQ	BVU_0660	48.97	BVU_1572	48.28	BVU_2146	30.50		
traP								
traO	BVU_0661	49.48	BVU_1571	53.12	BVU_2145	39.58		
traN	BVU_0662	69.82	BVU_1570	68.60	BVU_2144	54.74	BVU_3381	22.83
traM	BVU_0663	45.52	BVU_1569	46.01	BVU_2142	33.33		
traL								
traK	BVU_0665	73.91	BVU_1567	75.85	BVU_2140	67.31	BVU_3378	30.96
traJ	BVU_0666	68.60	BVU_1566	68.90	BVU_2139	63.66	BVU_3377	25.50
traI	BVU_0667	66.03	BVU_1565	65.37				
traH	BVU_0668	66.67	BVU_1564	68.14				
traG	BVU_0669	67.46	BVU_1563	67.70			BVU_3371	26.57
traF	BVU_0670	62.96	BVU_1562	60.75				

traE			BVU_1561	85.90				
traD								
traB			BVU_1559	25.66				
traA	BVU_0676	42.00	BVU_1558	42.00			BVU_3366	31.11
mobA	BVU_0677	45.71	BVU_1557	45.71				
mobB	BVU_0678	43.49	BVU_1556	43.49			BVU_3365	22.85
mobC	BVU_0679	73.19	BVU_1555	73.19				
Cnt029								
rteC								
Ctn031	BVU_0680	28.57	BVU_1554	28.57				
rteB	BVU_0681	42.43	BVU_1553	42.43				
rteA	BVU_0682	29.77	BVU_1552	29.77				
tetQ								
bmhA								
Ctn038								
exc	BVU_0709	38.02	BVU_1523	38.26				
Ctn040	BVU_0707	33.13	BVU_1525	32.73				
Ctn041	BVU_0719	34.25	BVU_1516	30.77				
Ctn042								
Ctn043	BVU_0718	36.05	BVU_1515	34.88				

rhuM								
int	BVU_0706	24.66	BVU_1527	24.93				

Table S6: Clusters of synteny in *B. vulgatus* mpk in comparison to *B. fragilis* CTn341

ctn341	Cluster 1	Cluster 1 similarity	Cluster 2	Cluster 2 similarity	Cluster 3	Cluster 3 similarity	Cluster 4	Cluster 4 similarity	Cluster 5	Cluster 5 similarity
Ctn001	BvMPK_0069	90.36								
Ctn002	BvMPK_0070	90.71			BvMPK_0316	76.92	BvMPK_0864	77.46	BvMPK_1832	77.86
Ctn003	BvMPK_0071	84.54			BvMPK_0315	43.65	BvMPK_0863	43.42	BvMPK_1833	64.79
Ctn004					BvMPK_0312	79.37	BvMPK_0860	77.78	BvMPK_1836	58.11
yahA	BvMPK_0074	96.08			BvMPK_0311	91.18	BvMPK_0859	91.18	BvMPK_1838	59.80
lys	BvMPK_0075	91.76	BvMPK_0169	55.88	BvMPK_0309	80.59			BvMPK_1840	45.12
traQ	BvMPK_0076	90.13	BvMPK_0170	47.83	BvMPK_0330	48.94			BvMPK_1841	53.90
traP	BvMPK_0077	79.45							BvMPK_1842	56.55
traO	BvMPK_0078	90.05	BvMPK_0171	52.94	BvMPK_0329	51.05			BvMPK_1843	56.91
traN	BvMPK_0079	93.41	BvMPK_0172	69.82	BvMPK_0328	56.59			BvMPK_1844	77.95
traM	BvMPK_0080	92.24	BvMPK_0173	47.41	BvMPK_0327	42.67			BvMPK_1845	51.41
traL	BvMPK_0081	80.00			BvMPK_0326	26.76			BvMPK_1846	32.50
traK	BvMPK_0082	99.52	BvMPK_0175	75.85	BvMPK_0325	73.91			BvMPK_1847	81.16
traJ	BvMPK_0083	91.32	BvMPK_0176	69.23	BvMPK_0324	71.34			BvMPK_1848	73.95
traI	BvMPK_0084	95.22	BvMPK_0177	64.85	BvMPK_0323	58.08			BvMPK_1849	55.98
traH	BvMPK_0085	92.77			BvMPK_0322	33.33			BvMPK_1850	58.87
traG	BvMPK_0086	96.74	BvMPK_0178	68.23						
traF	BvMPK_0087	97.27	BvMPK_0179	63.21						

traE	BvMPK_0088	92.23	BvMPK_0180	86.42						
traD	BvMPK_0089	59.76								
traB	BvMPK_0091	28.95								
traA	BvMPK_0092	47.30								
mobA	BvMPK_0093	90.85								
mobB	BvMPK_0094	81.38								
mobC	BvMPK_0095	93.75								
Cnt029	BvMPK_0102	96.28								
rteC	BvMPK_0103	84.88								
Ctn031	BvMPK_0104	70.68								
rteB	BvMPK_0105	84.99								
rteA	BvMPK_0106	68.02								
tetQ										
bmhA	BvMPK_0118	92.86								
Ctn038	BvMPK_0120	80.00					BvMPK_0854	91.93		
exc	BvMPK_0121	89.58					BvMPK_0853	93.53		
Ctn040	BvMPK_0122	81.53					BvMPK_0852	91.97		
Ctn041	BvMPK_0123	86.21					BvMPK_0851	98.28		
Ctn042	BvMPK_0124	59.50					BvMPK_0850	59.17		
Ctn043	BvMPK_0125	91.75					BvMPK_0849	100.00		

rhuM	BvMPK_0128	74.17					BvMPK_0847	93.33		
int	BvMPK_0129	95.86					BvMPK_0846	97.08		

Table S7: List of antibiotic sensitivity tests

Agent class	Antimicrobial agent	*B. vulgatus* MPK	*B. vulgatus* ATCC 8482
Beta-lactams	Cefuroxim (2nd Ceph)	resistant	susceptible
	Cefotaxim (3rd Ceph)	resistant	susceptible
	Cefdopoxin (3rd Ceph)	resistant	susceptible
	Caftazidim (3rd Ceph)	resistant	susceptible
	Cefepim (4th Ceph)	resistant	susceptible
	Penicillin (Pen)	resistant	resistant
	Piperacillin (Pen)	susceptible	susceptible
	Oxacillin (Pen)	resistant	resistant
	Aztreonam (Mono)	resistant	susceptible
	Imipenem (Carb)	susceptible	susceptible
	Meropenem (Carb)	susceptible	susceptible
Aminoglycosides	Amikacin	resistant	resistant
	Gentamycin	resistant	resistant
	Tobramycin	resistant	resistant
Quinolones	Ciprofloxacin	resistant	resistant
	Levofloxacin	susceptible	susceptible
	Moxifloxacin	susceptible	susceptible
Others	Co-trimoxazole	resistant	resistant
	Fosfomycin	resistant	resistant
	Tetracycline	susceptible	susceptible
	Erythromycin	susceptible	susceptible
	Rifampicin	susceptible	susceptible
	Chloramphenicol	susceptible	susceptible
	Colistin	resistant	resistant

Ceph, cephalosporin; Pen, Penicillins; Mono, Monobactams; Carb, Carbapenems

Table S8: Overview on IS4- and IS21-like elements

IS 4-like elements	**Comments**
BvMPK_1146	
BvMPK_1220	
BvMPK_1226	
BvMPK_1316	
BvMPK_1547	
BvMPK_2006	
BvMPK_2202	
BvMPK_2538	
BvMPK_2666	
BvMPK_2675	
BvMPK_2765	
BvMPK_2844	
BvMPK_3343	
BvMPK_3517	
IS21-like element pairs type 1	
BvMPK_1044/45	
BvMPK_1710/11	
BvMPK_1852/53	
IS21-like element pairs type 2	
BvMPK_0302/3	truncated Transposase
BvMPK_0856/57	truncated Transposase
BvMPK_1364/65	
BvMPK_1880/81	
BvMPK_2015/16	truncated Transposase
BvMPK_2101/02	
BvMPK_2109/10	truncated transposase and ATP-binding protein
BvMPK_2229/30	
BvMPK_2704/05	truncated Transposase

BvMPK_2762/63	
BvMPK_3201/02	
BvMPK_3251/52	
BvMPK_4277/78	

Table S9: tRNAs / tRNA clusters and associated mobile elements in *B. vulgatus* mpk

Locus Tag	**tRNA**	**Mobile elements / rRNA cluster**
BvMPK_0141, BvMPK_0142	tRNA-Pro-CGG, tRNA-Phe-GAA	BvMPK_0143 (integrase)
BvMPK_0193	tRNA-Lys-TTT	
BvMPK_0352, BvMPK_0353	tRNA-Ala-TGC, tRNA-Ile-GAT	in rRNA cluster
BvMPK_0437, BvMPK_0438	tRNA-Ala-TGC, tRNA-Ile-GAT	in rRNA cluster
BvMPK_0447	tRNA-Phe-GAA	BvMPK_0448 (integrase)
BvMPK_0581	tRNA-Trp-CCA	
BvMPK_0583, BvMPK_0584, BvMPK_0585, BvMPK_0586	tRNA-Thr-GGT, tRNA-Gly-TCC, tRNA-Tyr-GTA, tRNA-Thr-TGT	
BvMPK_0599, BvMPK_0600	tRNA-Asn-GTT, tRNA-Asn-GTT	BvMPK_0601 (transposase)
BvMPK_0636, BvMPK_0637	tRNA-Leu-TAA, tRNA-Gly-GCC	BvMPK_0634 and BvMPK_0635 (transposases)
BvMPK_0696	tRNA-Arg-CCT	
BvMPK_0928	tRNA-Gln-TTG	
BvMPK_1118	tRNA-Pro-GGG	BvMPK_1116 (transposase) and BvMPK_1118 (hypothetical protein)
BvMPK_1142	tRNA-Gln-CTG	BvMPK_1143 (transposase)
BvMPK_1168	tRNA-Ser-GGA	
BvMPK_1237, BvMPK_1238	tRNA-Ile-GAT, tRNA-Ala-TGC	in rRNA cluster
BvMPK_1324	tRNA-Trp-CCA	

BvMPK_1333	tRNA-Ser-TGA	BvMPK_1334 (integrase), BvMPK_1335 and BvMPK_1336 (transposases)
BvMPK_1545	tRNA-Leu-CAA	BvMPK_1547 to BvMPK_1549 (transposases)
BvMPK_1595	tRNA-Glu-TTC	
BvMPK_1686	tRNA-Met	
BvMPK_1759	tRNA-Ser-GGA	
BvMPK_1795	tRNA-His-GTG	BvMPK_1794 (transposase)
BvMPK_2072	tRNA-Glu-TTC	BvMPK_2070 (transposase) and BvMPK_2071 (hypothetical protein)
BvMPK_2093	tRNA-Ser-TGA	BvMPK_2094 and BvMPK_2095 (transposases)
BvMPK_2302, BvMPK_2303	tRNA-Gly-TCC, tRNA-Tyr-GTA	BvMPK_2300 (transposase) and BvMPK_2301 (hypothetical protein)
BvMPK_2324	tRNA-Thr-TGT	
BvMPK_2391	tRNA-Lys-TTT	
BvMPK_2409	tRNA-Arg-TCT	BvMPK_2410 (transposase)
BvMPK_2519	tRNA-Glu-CTC	BvMPK_2518 and BvMPK_2520 (transposases)
BvMPK_2522, BvMPK_2523	tRNA-Ser-GCT, tRNA-Cys-GCA	BvMPK_2520 (transposase), BvMPK_2521 (hypothetical protein)
BvMPK_2616	tRNA-Asp-GTC	BvMPK_2619 (mobilization protein)
BvMPK_2680	tRNA-Ala-GGC	BvMPK_2675 (transposase), BvMPK_2681 (integrase)
BvMPK_2753, BvMPK_2754, BvMPK_2755	tRNA-Lys-CTT, tRNA-Lys-CTT, tRNA-Lys-CTT	BvMPK_2762 (transposase)
BvMPK_2797	tRNA-Gly-CCC	BvMPK_2798 (integrase)
BvMPK_2805	tRNA-Gly-CCC	BvMPK_2808 (integrase)
BvMPK_2843	tRNA-Pro-TGG	BvMPK_2844 (transposase)

BvMPK_2942, BvMPK_2943, BvMPK_2944	tRNA-Arg-ACG, tRNA-Arg-ACG, tRNA-Arg-ACG	
BvMPK_3081, BvMPK_3082, BvMPK_3083, BvMPK_3084, BvMPK_3085, BvMPK_3086	tRNA-Gly-GCC, tRNA-Leu-CAG, tRNA-Gly-GCC, tRNA-Leu-GAG, tRNA-Leu-CAG, tRNA-Gly-GCC	
BvMPK_3122, BvMPK_3123	tRNA-Ala-TGC, tRNA-Ile-GAT	in rRNA cluster
BvMPK_3358, BvMPK_3359	tRNA-Ala-TGC, tRNA-Ile-GAT	in rRNA cluster
BvMPK_3365, BvMPK_3366, BvMPK_3367	tRNA-Asp-GTC, tRNA-Asp-GTC. tRNA-Asp-GTC	
BvMPK_3539	tRNA-Thr-CGT	BvMPK_3542 (transposase)
BvMPK_3666	tRNA-Arg-CCG	
BvMPK_3835, BvMPK_3836	tRNA-Met-CAT, tRNA-Met-CAT	
BvMPK_3914	tRNA-Leu-TAG	
BvMPK_3930, BvMPK_3931	tRNA-Ala-TGC, tRNA-Ile-GAT	in rRNA cluster
BvMPK_4095	tRNA-Thr-TGT	
BvMPK_4141	tRNA-Met-CAT	
BvMPK_4161, BvMPK_4162, BvMPK_4163	tRNA-Val-TAC, tRNA-Val-TAC, tRNA-Val-TAC	
BvMPK_4202, BvMPK_4203	tRNA-Ala-TGC, tRNA-Ile-GAT	in rRNA cluster

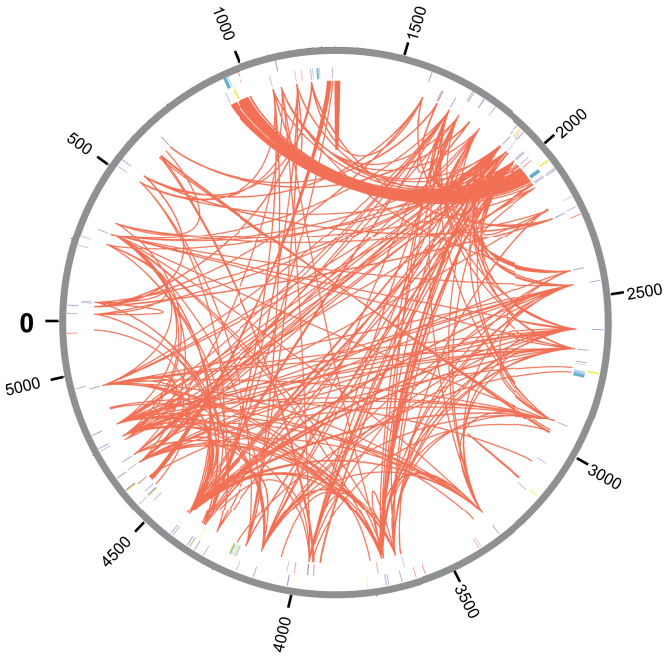


Figure S1: A circular view of the *Bacteroides vulgatus* ATC8482 genome with protein paralogy highlights and display of mobile elements. Red lines indicate paralogy based on at least 90% protein identity. Mobile elements are shown as lines on the forward (inner circle) and reverse (outer circle) strand: Conjugative transposon proteins, CRISPR/Cas-related proteins, integrases, mobilization proteins, transposases, and other mobile elements.

Accepted publication b

Maerz J.K., Steimle A., **Lange A.**, Bender A., Fehrenbacher B., Frick J.S.: Outer membrane vesicles blebbing contributes to *B. vulgatus* mpk-mediated immune response silencing. Gut Microbes 9: 1-12 (2018)

Gut Microbes

ISSN: 1949-0976 (Print) 1949-0984 (Online) Journal homepage: http://www.tandfonline.com/loi/kgmi20

Outer membrane vesicles blebbing contributes to *B. vulgatus* mpk-mediated immune response silencing

Jan Kevin Maerz, Alex Steimle, Anna Lange, Annika Bender, Birgit Fehrenbacher & Julia-Stefanie Frick

To cite this article: Jan Kevin Maerz, Alex Steimle, Anna Lange, Annika Bender, Birgit Fehrenbacher & Julia-Stefanie Frick (2018) Outer membrane vesicles blebbing contributes to *B. vulgatus* mpk-mediated immune response silencing, Gut Microbes, 9:1, 1-12, DOI: 10.1080/19490976.2017.1344810

To link to this article: https://doi.org/10.1080/19490976.2017.1344810

Accepted author version posted online: 07 Jul 2017.
Published online: 13 Jul 2017.

GUT MICROBES
2018, VOL. 9, NO. 1, 1–12
https://doi.org/10.1080/19490976.2017.1344810

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER/REPORT

Outer membrane vesicles blebbing contributes to *B. vulgatus* mpk-mediated immune response silencing

Jan Kevin Maerz[a], Alex Steimle[a], Anna Lange[a], Annika Bender[a], Birgit Fehrenbacher[b], and Julia-Stefanie Frick[a]

[a]Institute of Medical Microbiology and Hygiene, University of Tübingen, Tübingen, Germany; [b]University Department of Dermatology, University of Tübingen, Tübingen, Germany

ABSTRACT

The Gram negative intestinal symbiont *Bacteroides vulgatus* mpk is able to prevent from induction of colonic inflammation in *Rag1*$^{-/-}$ mice and promotes immune balance in *Il2*$^{-/-}$ mice. These inflammation-silencing effects are associated with *B. vulgatus* mpk-mediated induction of semi-mature dendritic cells, especially in the colonic lamina propria (cLP). However the beneficial interaction of bacteria with host immune cells is limited due to the existence of a large mucus layer covering the intestinal epithelium. How can intestinal bacteria overcome this physical barrier and contact the host immune system?

One mechanism is the production of outer membrane vesicles (OMVs) via ubiquitous blebbing of the outer membrane. These proteoliposomes have the ability to traverse the mucus layer. Hence, OMVs play an important role in immunomodulation and the maintenance of a balanced gut microbiota. Here we demonstrate that the stimulation of bone marrow derived dendritic cells (BMDCs) with isolated OMVs originated from *B. vulgatus* mpk leads to the induction of a tolerant semi-mature phenotype. Thereby, microbe- associated molecular patterns (MAMPs) delivered by OMVs are crucial for the interaction and the resulting maturation of immune cells. Additional to the binding to host TLR4, a yet unknown ligand to TLR2 is indispensable for the conversion of immature BMDCs into a semi-mature state. Thus, crossing the epithelial mucus layer and directly contact host cells, OMV mediate cross-tolerance via the transport of various Toll-like receptor antigens. These features make OMVs to a key attribute of *B. vulgatus* mpk for a vigorous acellular prevention and treatment of systemic diseases.

ARTICLE HISTORY
Received 21 April 2017
Revised 8 June 2017
Accepted 15 June 2017

KEYWORDS
Bacteroides vulgatus; dendritic cells; immune response; outer membrane vesicles

Introduction

The intestinal microbiota provides important features which are considered to be beneficial for the host organism, such as (1) the maintenance of the structural integrity of the intestinal mucosal barrier, (2) the development of epithelial cell function, (3) host nutrient metabolism, (4) modulation of the immune system and (5) prevention of bacterial overgrowth.[1,2] Hereby, a certain microbial diversity is considered to be required for the maintenance of the immune homeostasis and crucial for the optimal functionality and interaction with the host. Concerning the interplay between microbiota and the host immune system, a disturbed or altered composition of the gut microbiota, called "dysbiosis," is associated with the pathogenesis of both intestinal and systemic immunological disorders.[3] Such a dysbiosis of the intestinal microbiota, can also influence the disease onset and progress in a variety of mouse models for autoimmune diseases (AID) such as rheumatoid arthritis (RA), type 1 diabetes (T1D) and inflammatory bowel disease (IBD).[4-6] This highlights the importance of a deeper knowledge of the interaction between the host immune system and the intestinal microbes. Based on this cross talk, specific modulation of the intestinal microbiota composition is considered to be used as therapy for the treatment of microbiota-associated AID.[7,8] Hereby, intestinal commensals are decisive for the induction (pathobionts) or prevention (symbionts) of a pathological immune response in a certain predisposed host.[9] In this context, members of the genus *Bacteroides* were already demonstrated to exhibit symbiotic properties concerning the modulation of the intestinal immune system.[10]

CONTACT Prof. Dr. Julia-Stefanie Frick julia-stefanie.frick@med.uni-tuebingen.de Institute of Medical Microbiology and Hygiene, University of Tübingen, Elfriede-Aulhorn-Str.6, D-72076 Tübingen, Germany.

Accordingly, *Bacteroides vulgatus* mpk was already depicted to mediate inflammation-silencing effects finally preventing from colitis induction in various moue models for IBD.[1-13] These immune regulatory properties of *B. vulgatus* mpk are traced back to the induction of a tolerant dendritic cell (DC) phenotype in the colonic lamina propria (cLP) termed semi-mature.[12-14] However, direct physical interaction between intestinal commensals and cLP DCs is limited due to the presence of a thick and almost sterile mucus layer covering the intestinal epithelium. Therefore, we were interested in how this symbiotic commensal manages to enhance the communication with host cells to modulate the immune system.

Instead of direct physically interacting with target cells, bacteria can communicate with more distant host environments via secretion of signaling molecules, such as toxins, quorum sensing molecules and DNA.[15] These mechanisms involving non-viable bacteria provide distinct advantages for bacteria, since secreted soluble material is smaller and non-viable.[16,17] Therefore, this material can gain access to environments usually being inaccessible for living bacteria. In this context, the biogenesis and release of bacterial outer membrane vesicles (OMVs) by Gram-negative bacteria represent a secretion pathway combining the long-distance traveling ability of small soluble bacterial products with mimicking the characteristics of the whole bacterial cell by facilitating the transport of insoluble molecules.[18,19] OMV release therefore constitutes a critical mechanism for intra- and inter-kingdom communication.[20,21] Derived from the outer membrane of bacteria, OMVs harbour a broad variety of microbe-associated molecular patterns (MAMPs) which interact with pattern recognition receptors (PRRs) of host target cells, leading to innate and adaptive immune responses.[22,23] It was shown that the application of OMVs critically modulates the course of disease in animal models of intestinal inflammation by either a direct or an indirect interaction with the intestinal immune system.[24,25] Hickey et al. even showed that OMVs of *B. thetaiotaomicron* can be found in macrophages of the intestinal mucosa. This implicates a passaging through not only the mucus layer but also through the intestinal epithelial cell layer.

In this study, we provide evidence that the symbiotic commensal *B. vulgatus* mpk produces OMVs. This generation of OMVs contribute to the mediation of the immune-system silencing properties provoked by this strain via the induction of DC semi-maturation. Thereby, both TLR2- and MD-2/TLR4-mediated signaling is crucial for this OMV-induced generation of semi-mature DCs, indicating that TLR2 and TLR4 ligands carried by OMVs are indispensable for the manifestation of this tolerant DC phenotype. Therefore, B. vulgatus mpk-caused OMV blebbing seems to contribute to host-microbe communication at intestinal sites which are hardly accessible for the living bacterium.

Results

Bacteroides vulgatus mpk outer membrane vesicles induce tolerance in CD11c$^+$ cells

It is well known that components of the intestinal microbiota have widespread effects on the mucosal immune system in the intestine.[14,26] Hereby, commensal bacteria exhibit important functions for the priming of immune cells underlying the mucosal epithelial barrier. However, physical contact between bacteria and such immune cells is restricted due to the presence of a large mucus layer covering the intestinal epithelium. Almost all bacteria produce outer membrane vesicles (OMVs) through bulging of the outer bacterial membrane.[27] These vesicles are able to cross the mucin layer, finally increasing the probability of interaction of mucosal immune cells with bacterial surface structures such as MAMPs.[28] The mechanism of OMV production is a key characteristic of Gram negative bacteria and commensal-derived OMVs seem to be important for the modulation of the host immune system.[29]

In previous work, we could demonstrate that *B. vulgatus* mpk modulate the immune response *in vivo* and *in vitro* in an inflammation-silencing manner. This is mainly mediated by the induction of tolerant CD11c$^+$ cells in the colonic lamina propria, which seem to be responsible for the maintenance of homeostatic conditions in the intestine.[12,13] Several molecular mechanisms for tolerance induction have already been identified in CD11c$^+$ bone marrow derived dendritic cells (BMDCs).[30] Until now, not all immune system-modulating surface structures of *B. vulgatus* mpk were already identified. Therefore, we were interested if this symbiotic commensal produces OMVs and if this vesicle production contributes to the observed immuno-

modulatory properties of this commensal strain. The secretion of OMVs can be a response due to environmental stress of the prokaryotic cell.[28,31] Therefore, we incubated the bacteria under different culture conditions and for different periods of time, to elucidate, if this strain produces OMVs and if this is dependent on the culturing conditions. We could demonstrate that *B. vulgatus* produces vesicles ubiquitarily under non-detrimental conditions through bulging of the outer membrane as demonstrated in TEM pictures (Fig. 1a). To gain information on the immune system modulating features of OMVs, $CD11c^+$ BMDCs were generated as described and stimulated with *B. vulgatus* mpk-derived OMVs (OMV_{BV}) for 24 h (Fig. 1b). After end of incubation time, BMDCs were checked for surface expression of MHC-II, CD40, CD80 and CD86 by flow cytometry (Fig. 1c, gating strategy) since these surface markers indicate the maturation status of $CD11c^+$ BMDCs.[11,12,14,32] PBS-stimulated BMDCs were used as a negative control (mock), representing completely immature $CD11c^+$ cells. BMDCs stimulated with pathobiotic *E. coli* mpk were used as positive controls, representing mature $CD11c^+$ cells, since $CD11c^+$ BMDC stimulation with this strain was demonstrated to induce complete maturation.[11] *B. vulgatus* mpk was already demonstrated to induce $CD11c^+$ cell semi-maturation as indicated by intermediate expression of MHC-II, CD40, CD80 and CD86 which was significantly lower compared with the expression of these respective surface proteins of mature $CD11c^+$ BMDCs.[11] We could demonstrate that OMV_{BV} provide the same properties like live *B. vulgatus* mpk when used at a concentration of 50 ng mL^{-1}, since we could not detect any difference in the expression of MHC-II, CD40, CD80 and CD86 between $CD11c^+$ cells stimulated with either OMV_{BV} or live *B. vulgatus* mpk, respectively (Fig. 1d). Additionally and as confirmed for live *B. vulgatus* mpk, $CD11c^+$ cells stimulated with OMV_{BV} provided significantly lower surface expression of these proteins compared with *E. coli* mpk stimulated $CD11c^+$ BMDCs. However, the surface expression of MHC-II and T cell co-stimulatory surface proteins of $CD11c^+$ BMDCs represents only a first hint on the maturation of these cells. A key feature of semi-mature $CD11c^+$ cells is tolerance toward a secondary bacterial stimulus after priming with *B. vulgatus* mpk.[11] Therefore, we primed $CD11c^+$ BMDCs with either OMV_{BV}, live *B. vulgatus* mpk, *E. coli* mpk or PBS (mock), incubated these cells for 24 h, changed cell culture medium and challenged them with either PBS or *E. coli* mpk (Fig. 1e). We analyzed the surface expression of MHC-II and the secretion of the pro-inflammatory cytokines TNF and IL-6. Immature PBS-stimulated $CD11c^+$ cells showed a significantly increased surface expression of MHC-II as well as secretion of TNF and IL-6 after challenge with *E. coli* mpk. Already mature *E. coli* mpk primed $CD11c^+$ cells maintained their high expression level of MHC-II and IL-6 after challenge with *E. coli* mpk. Due to the high secretion of TNF during the first 24 h, indicating a maturation of the BMDCs, secreted TNF was not detectable in the supernatant of mature $CD11c^+$ cells upon a second challenge with *E.coli* mpk after exchange of media.[33,34] This is in line with the literature reporting on a non-responsiveness of mature BMDCs to additional stimuli.[11,32,35] Semi-mature $CD11c^+$ cells stimulated with either live *B. vulgatus* mpk or OMV_{BV} were non-responsive toward subsequent challenge with *E. coli* mpk concerning the surface expression of MHC-II and the secretion of TNF in the supernatant (Fig. 1f). Furthermore, to provide evidence for a direct interaction of *B. vulgatus*-derived OMVs with antigen-presenting cells, we performed an internalisation timeline assay using $CD11c^+$ BMDCs in combination with different concentration of FITC-labeled vesicles. OMV_{BV} were taken up by cultured BMDCs, leading to an extensive increase of $CD11c^+OMV_{BV}{}^+$ cell population after 30 minutes as shown in Figure 1g. As shown by Shen et al., the rapid internalization of vesicles can be achieved in an actin-dependent manner directing the expression of surface markers and inflammatory cytokine secretion in BMDCs.[24]

OMV_{BV}-mediated immunomodulatory properties are sensed via host TLR2 and TLR4 receptor complexes

We demonstrated that *B. vulgatus* mpk and its outer membrane vesicles interact with innate immune cells and modulate the host immune system in an anti-inflammatory sense. Though, the exact mechanisms and involved pattern recognition receptors for signal transduction were not yet accurately identified. Therefore, we wanted to describe the responsible receptor on host target cells for the interaction with microbial ligands on bacterial cell surface and outer membrane vesicles, respectively. In this context, Toll-like receptors (TLRs) are the most pivotal pattern-recognition

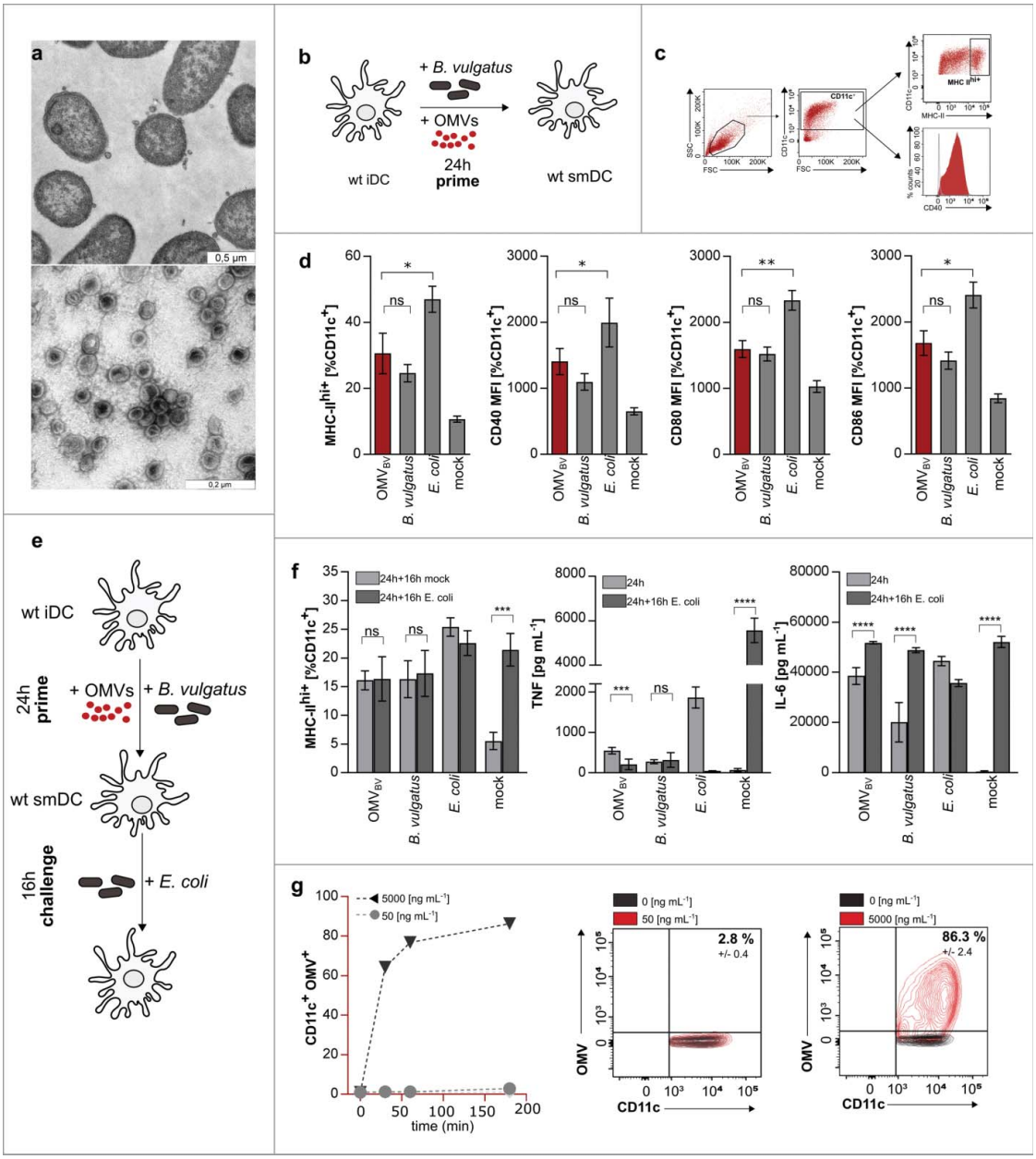


Figure 1. Induction of dendritic cell semi-maturation via *B. vulgatus* OMVs (OMV_{BV}). (a) Evidence for vesicle production derived from the outer membrane of cultivated *Bacteroides vulgatus* cell. Black arrows indicate single secreted vesicles of fixed bacterial cells after toluidine blue staining captured with Transmission Electron Microscope (upper picture). Native OMV_{BV} after isolation and purification stained negative with uranyl acetate (lower picture). (b) Experimental setting for the analysis of surface maturation marker expression on dendritic cells after stimulation with OMVs. BMDCs are primed for 24 h with either PBS (mock), *B. vulgatus* or *E. coli* as control and additionally with 50 ng mL^{-1} OMV_{BV}. (c) Gating strategy for the determination of MHC class II high positive (MHC-II $^{hi+}$), CD40, CD80 or CD86 BMDCs. CD11c-negative cells were excluded and the proportion of MHC-II $^{hi+}$ (dot plot), CD40, CD80 or CD86 (histogram) DCs within the population of $CD11c^+$ cells was determined as shown. Dashed line shows negative control. (d) MHC class II $^{hi+}$, CD40, CD80 or CD86 population $CD11c^+$ BMDCs that were primed with PBS (mock), *B. vulgatus* mpk, 50 ng mL^{-1} OMV_{BV} or *E. coli* mpk for 24 h (n = 4). (e) Experimental setting for investigation of dendritic cell tolerance after induction of semi-maturation. BMDCs are primed for 24 h with either PBS (mock) to preserve an immature phenotype, *B. vulgatus* mpk or 50 ng mL^{-1} OMV_{BV} to induce semi-maturation or *E. coli* to induce BMDC maturation. After medium change, these cells were secondarily challenged for 16 h with either PBS (mock) as controls or *E. coli* to proof non-responsiveness of tolerant cells. (f) Expression of analyzed maturation markers MHC class II, TNF and IL-6 of $CD11c^+$ BMDCs after priming with PBS (mock), *B. vulgatus* mpk, 50 ng mL^{-1} OMV_{BV} or *E. coli* mpk for 16 h and subsequently challenge with either PBS (mock) or *E. coli* mpk for 16 h (n = 4). (g) Flow cytometry analysis of OMV_{BV} internalization by BMDCs. Different concentrations of vesicles were labeled with fluorescein isothiocyanate (FITC) and incubated with cultured DCs for various times. Percentages show $CD11c^+OMV_{BV}^+$ cell populations. Unlabeled OMVs served as negative control (black population in dot plot graph). All statistical analyses were performed using student's t test. Error bars represent SD.

receptors (PRRs) for the detection of MAMPs.[36-39] The most abundant MAMPs exhibiting strong immunogenic properties are lipopolysaccharides (LPS), which are recognized by CD14 and the TLR4/MD-2 receptor complex.[14,40-42] Besides TLR4 agonists, TLR2 ligands like bacterial lipopeptides, play a prominent role regarding target activation of antigen presenting cells.[43-45] In general, TLR4- as well as TLR2- agonists can be integrated into OMVs and therefore be transported by these vesicles.[46] To characterize the *B. vulgatus* mpk- and OMV_{BV}- associated ligands which might be relevant for the induction of semi-mature BMDCs, we used Human Embryonic Kidney cells (HEK) overexpressing mouse TLR2 or the mouse CD14/TLR4/MD-2 receptor complex. The resulting IL-8 secretion into cell supernatant was detected to determine receptor activation. PBS treated cells served as negative control, whereas Pam_3CSK_4 was used as TLR2 ligand. As demonstrated in Fig. 2a, OMV_{BV} induced both activation of the TLR2 receptor as well as CD14/TLR4/MD-2 receptor complex. To validate these results, we used $CD11c^+$ BMDCs derived from mice deficient for Toll-like receptor 2 ($Tlr2^{-/-}$), TLR4 ($Tlr4^{-/-}$) and for both receptors ($Tlr2^{-/-}xTlr4^{-/-}$). $CD11c^+$ BMDCs of *wt* and $Tlr2^{-/-}$, $Tlr4^{-/-}$ and $Tlr2^{-/-}xTlr4^{-/-}$ mice were generated as described and primed with OMV_{BV} for 24 h (Fig. 2b). $CD11c^+$ cells were harvested and analyzed for surface expression of MHC-II and CD40 (Fig. 1c, gating strategy). Additionally, the amount of secreted pro-inflammatory TNF and IL-6 was detected. The deficiency of both TLRs on the cell surface leads to a complete loss of antigen recognition independent of the used stimulus as indicated by the low expression of MHC-II and an abolished secretion of TNF of $Tlr2^{-/-}xTLR4^{-/-}$ $CD11c^+$ BMDCs. In accordance with the findings generated with HEK cells (Fig. 2a), the activation of the TLR4/MD-2 receptor complex via *B. vulgatus* mpk- secreted OMVs is decisive for the main outcoming signal. In addition, the expression of MHC-II, the costimulatory proteins and the secretion of the inflammatory cytokines in *Tlr2−/−* stimulated BMDCs is nearly as strong as in the *wt* cells after challenging with *E.coli* mpk, *B. vulgatus* mpk or high concentration of OMV_{BV}. Nevertheless, we also observed a dose-dependent increase of all analyzed proteins upon stimulation of $Tlr4^{-/-}$ $CD11c^+$ BMDCs with purified OMV_{BV}, pointing out that the TLR2 receptor is competent to recognize agonistic structures present in OMV_{BV}. Furthermore, TLR2-dependent signaling is also involved in vesicle- induced BMDC activation and maturation (Fig. 2c). This observation prompted us to elucidate whether the contact of host immune cells with a bacterial TLR4 agonist such as LPS alone is sufficient for the activation and differentiation of tolerant BMDCs or whether an additional TLR2 activation is accessory required for adequate induction of BMDC semi-maturation. Therefore, BMDCs were generated from $Tlr2^{-/-}$ mice, primed with OMV_{BV} for 24 h and challenged with *E. coli* mpk for additional 16 h (Fig. 2d). As mentioned before, this experimental setting is required to check for tolerance in semi-mature BMDCs. We could verify that a TLR2-dependent signaling is necessary for tolerance induction in BMDCs, since $Tlr2^{-/-}$ failed to become tolerant upon priming with *B. vulgatus* mpk or OMV_{BV} as demonstrated by the high expression of MHC class II and the enhanced secretion of TNFα upon *E. coli* mpk challenge (Fig. 2e).

Discussion

In this study, we demonstrated that the Gram negative symbiotic commensal *B. vulgatus* mpk produces outer membrane vesicles (OMV). This OMV blebbing mechanism seems to contribute to the immune response modulating properties of this strain which was demonstrated in several previous studies. *B. vulgatus* mpk-derived OMVs (OMV_{BV}) induced tolerance in $CD11c^+$ BMDCs in a TLR4 and TLR2 dependent manner. However, the defined vesicle-associated receptor ligands were not identified. Nevertheless, simultaneous activation of both receptors was required to induce effective DC tolerance. In line with Shen et al. and Hickey et al.[24,25,27] this feature indicates an additional mechanism how *B. vulgatus* mpk as a symbiotic commensal, modulates the immune system at mucosal interfaces. As Bacteria are separated from the host intestinal mucosa by a thick mucus layer that is thought to impede a direct contact between luminal bacteria and the intestinal mucosa,[47] the shedding of OMVs by *B. vulgatus* is of special importance since OMVs can diffuse through the mucin layer.[25,48] This underlines the role of OMVs for the modulation of intestinal immune cells. These characteristics make OMV_{BV} to an important attribute of *B. vulgatus* mpk for an effective passive protection and treatment of systemic diseases.

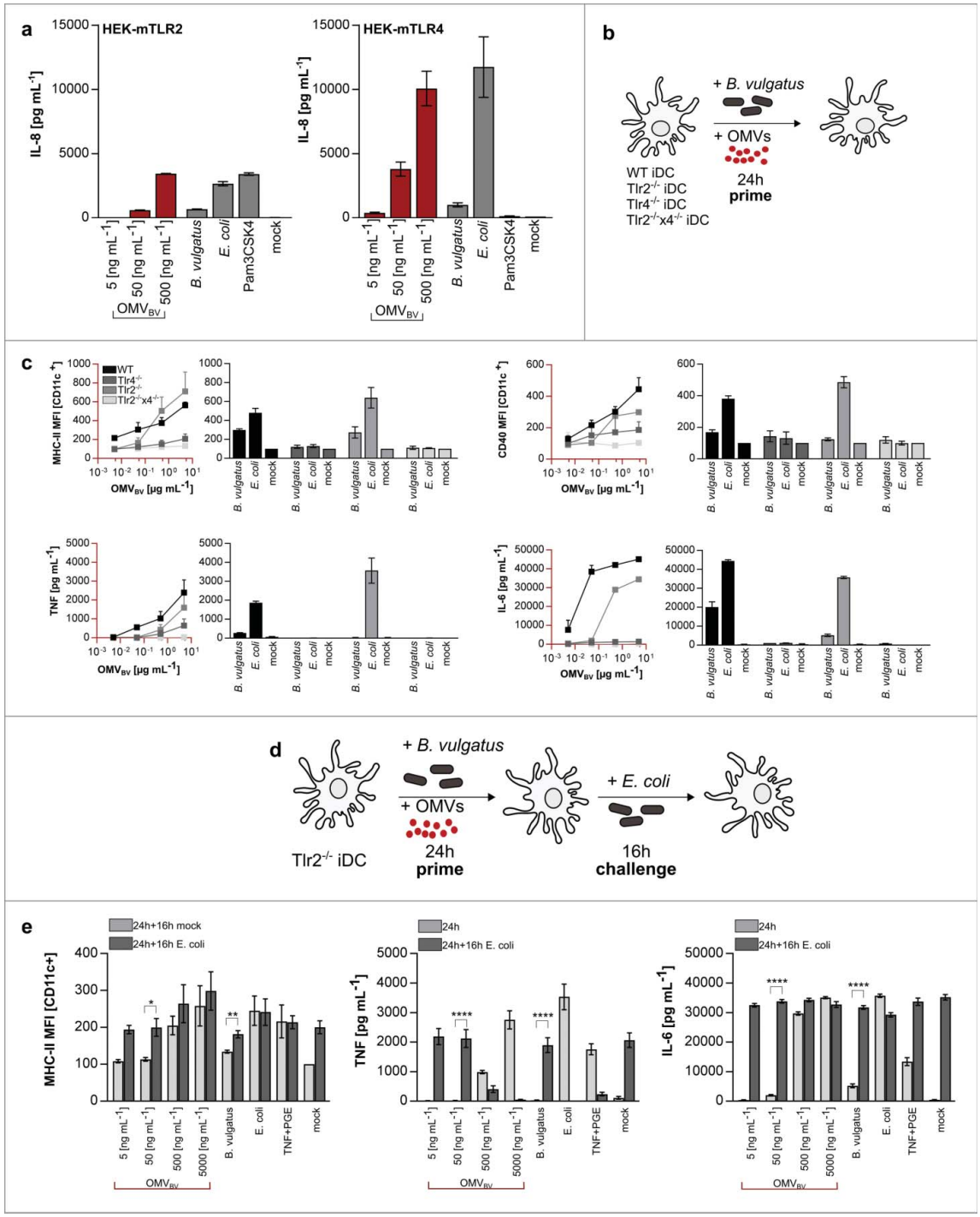


Figure 2. Induction of dendritic cell tolerance via OMV_{BV} is TLR2 and TLR4 dependent. (a) Stimulation of Human embryonic kidney (HEK) cells overexpressing murine TLR 2 or 4. HEK cells are primed with PBS (mock), *B. vulgatus* mpk, various concentrations of OMV_{BV} or *E. coli* mpk for 24 h. Pam3CSK4 served as TLR2 specific control. (b) Experimental setting for the analysis of surface maturation marker expression on dendritic cells after stimulation with OMVs. *wt, Tlr2*$^{-/-}$, *Tlr4*$^{-/-}$ and *Tlr2*$^{-/-}$*xTlr4*$^{-/-}$ BMDCs are primed for 24 h with either PBS (mock), *B. vulgatus* or *E. coli* as control and additionally with increasing concentrations of OMV_{BV}. (c) MHC class II^{hi+} or CD40 population (CD11c^{+}) (normalized to mock) and secreted concentration of TNF or IL-6 of *wt* and TLR-deficient BMDCs that were primed with PBS (mock), *B. vulgatus* mpk, different concentrations of OMV_{BV} or *E. coli* mpk for 24 h (n = 4). (d) Experimental setting for investigation of dendritic cell tolerance after induction of semi-maturation in Tlr2$^{-/-}$ DCs.BMDCs are primed for 24 h with either PBS (mock) to preserve an immature phenotype, *B. vulgatus* or OMV_{BV} to induce semi-maturation or *E. coli* to induce BMDC maturation. After medium change, these cells were secondarily challenged for 16 h with either PBS (mock) as controls or *E. coli* to proof non-responsiveness of tolerant cells. (e) Expression of analyzed maturation markers MHC class II, TNF and IL-6 of CD11c^{+} BMDCs of TLR2$^{-/-}$ mice after priming with PBS (mock), *B. vulgatus* mpk, OMV_{BV} or *E. coli* mpk for 16 h and subsequently challenge with either PBS (mock) or *E. coli* mpk for 16 h (n = 4). All statistical analyses were performed using student's t test. Error bars represent SD.

Gram negative bacteria are able to deliver a chemically diverse spectrum of cargo protected from the external environment over long distances by the help of outer membrane vesicles.[49] Though, OMVs can cross physical barriers which are impermeable for the whole bacterium such as the intestinal mucus layer, gain access to the intestinal epithelium and therefore to the underlying immune cells present in the intestinal lamina propria.[24,25,28,48] This makes OMV blebbing to an important immune modulating mechanism. Considering pathogens, vesicles of *Salmonella enterica* or *Helicobacter pylori* harbour virulence factors to enhance bacterial survival and the ability to colonize host mucosal tissues.[50,51] This leads to a stringent immune reaction and local inflammation.[52] However, OMVs derived from commensal symbiotic bacteria were demonstrated to be involved in maintaining or restoring immune homeostasis in the intestine.[53,54] For example, *Bacteroides fragilis*, representing an intestinal commensal, produces OMVs and interacts with the host immune system by delivering a capsular polysaccharide (PSA) to dendritic cells.[21,55] The release of PSA in OMVs induces immune-modulatory effects, activates regulatory T cells and prevents from experimental colitis.[24] However, a general role for OMVs in mediating immune responses, accounting for all OMV-producing commensal, was not yet demonstrated.[31,56]

Findings using live *B. vulgatus* mpk cells indicate that this symbiotic bacterium prevents from disease induction in different mouse models for experimental colitis.[12,13] In this context, *B. vulgatus* mpk promotes the maintenance of the immune equilibrium via induction of tolerant semi-mature dendritic cells in the intestine and the resulting regulation of host Cathepsin S activity and secretion in these antigen-presenting cells.[32] In this study, we demonstrated that OMVs derived from *B. vulgatus* mpk contribute to these observed effects since they interact with bone marrow derived dendritic cells, leading to a semi-mature phenotype characterized by a low expression of T cell activation and maturation markers as well as pro-inflammatory cytokines.

Innate immune cells recognize conserved microbial ligands through pattern recognition receptors (PRRs) leading to immunologic responses. For example, Polysaccharide A (PSA) from *Bacteroides fragilis* is recognized by TLR2. In consequence, the binding of PSA to TLR2 leads to the induction of $FoxP3^{+}$ regulatory T cells which affect both the development and homeostasis of the host immune system.[57,58] Our study now reveals that OMV_{BV} also contain TLR2 activating ligands which are, in consequence, necessary for the induction of tolerant BMDCs. Since these tolerance-inducing effects are abolished in TLR2- deficient BMDCs, a TLR2-dependent intracellular signaling is indispensable for OMV_{BV}-mediated induction of tolerance in BMDCs. These observations indicate that both receptors, MD-2/TLR4 and TLR2, are required for the induction of DC tolerance. In addition to the LPS in the outer membrane, OMVs of *B. vulgatus* mpk contain and a yet unidentified TLR2 agonist, as it is e.g. the case for PSA in vesicles secreted by *B. fragilis*. Interestingly, the genome of *B. vulgatus* mpk contains 9 different loci coding for capsular polysaccharides (e.g., glycosyltransferases and polysaccharide export outer membrane proteins), indicating the presence of a TLR2-interacting polysaccharide on the outer membrane of the bacterium cell and on vesicles derived from *B. vulgatus* mpk.[46] But the comparison of the various loci in *B. fragilis* and *B. vulgatus* revealed different arrangements of the genes necessary for the biosynthesis of capsular polysaccharides (data not shown). This implicates, that polysaccharides present in *B. vulgatus* mpk has a different structure as compared with the capsular polysaccharide produce by *B. fragilis*.

The presence of a TLR4- and TLR2- ligand on the outer membrane indicates that the mechanism by which OMV_{BV} induce tolerance in DCs is not originated in classical endotoxin tolerance. Usually, low doses of one TLR agonist desensitize the host to subsequent stimulation with a related agonist or endotoxin.[59-61] However, the addition of a Toll-like receptor (TLR) ligand can also induces tolerance to subsequent challenges with other stimuli that signal through one or more different TLRs (cross-tolerance). Since outer membrane vesicles harbor various different TLR ligands, OMV_{BV} might even mediate TLR-cross-tolerance. The simultaneous interaction of different MAMPs with different receptors might therefore result in complementary, synergistic or antagonistic effects that modulate innate and adaptive immunity.[62-64]

In summary, outer membrane vesicles derived from the symbiotic intestinal commensal of *B. vulgatus* mpk play a pivotal role in modulating and regulating immune responses of the host. With their combined

qualities of crossing physically barriers and the mediation of endotoxin tolerance and even cross-tolerance in dendritic cells via the delivery of different microbial ligands to immune cells, the production of OMV represents an important key feature of this symbiotic strain, to prevent from intestinal inflammation in the host. Due to these properties, OMV_{BV} might even represent a potential therapeutical tool to hamper or recover from inflammatory bowel diseases (IBD) and other systemic inflammatory disorders.

Materials and methods

Cultivation of bone marrow derived dendritic cells (BMDCs)

Bone marrow cells were isolated and cultivated for differentiation as described previously.[65] At day 7 after isolation, resulting CD11c positive, bone marrow derived dendritic cells (BMDCs) were harvested and used for stimulation.

Cultivation of human embryonic kidney cells (HEK cells)

Before stimulation, cryoconserved cell stocks were thawed and cultured in DMEM media supplemented with 10% FCS, 1% Penicillin/Streptomycin up to passage 7.

Stimulation of bone marrow derived dendritic cells

Two × 10^6 BMDCs were stimulated with either PBS or *B. vulgatus* mpk or *E. coli* mpk at a MOI of 1 or various concentrations of *B. vulgatus*-derived OMVs (5, 50, 500, 5000 ng mL^{-1}). Cells were stimulated for a maximum of 24 hours. If a second challenge was used, media were changed and cells were restimulated with either PBS or *E. coli* mpk MOI 1 for a maximum of 16 hours. Additionally, exogenous TGF (Sigma-Aldrich) 20 ng mL^{-1} and prostaglandin E (Sigma-Aldrich) 1 μM were used as positive and Toll-like receptor- independent control.

Stimulation of human embryonic kidney cells

Two × 10^4 HEK cells were stimulated with either PBS or *B. vulgatus* mpk or *E. coli* mpk at a MOI of 1 or various concentrations of *B. vulgatus*-derived OMVs (5, 50, 500 ng mL^{-1}). Cells were stimulated for a maximum of 24 hours. FSL-1 (10, 100 ng mL^{-1}) and Pam3CSK4 (30, 300 ng mL^{-1}) (both InvivoGen) were used as positive and TLR2-specific agonist.

Flow cytometrical analysis

Multi-color FCM analyses were performed on a FACS LSRII (BD Biosciences). All fluorochrome-coupled antibodies were purchased from BD Biosciences. Data were analyzed using the FlowJo software (Tree Star Inc., USA).

Cytokine analysis by ELISA

For analysis of IL6, TNF and IL1ß concentrations in cell culture supernatants, ELISA Kits purchased from BD Bioscience were used according to the manufacturer's instructions.

Bacteria cultivation

The bacteria used for stimulation of the murine dendritic cells were *Escherichia coli* mpk and *Bacteroides vulgatus* mpk. The *E. coli* mpk strain was grown in Luria-Bertani (LB) medium under aerobic conditions at 37°C. *Bacteroides vulgatus* was grown in Brain-Heart-Infusion (BHI) medium and anaerobic conditions at 37°C.

Detection of secreted OMVs

Part of the cultured bacteria suspension was sedimented for 2 h and fixed with Paraformaldehyde-Glutaraldehyde Solution (Karnovsky's Fixative) in a rate of 7:2 (A:B) for 1 h at room temperature after removal of excess fluids. Following, for the stabilization of proteins and lipids samples were post-fixed in osmium tetroxide. Fixed cells are then embedded in agar and washed in 0.1 M cacodylate buffer. Addition of uranyl acetate for 1 h enhanced membrane stabilization and improved overall contrast. Dehydration was performed with an ETOH dilution row, followed by transfer in epoxide solution. The polymerization of the agarose pad was achieved via incubation for 20 h at 45°C and plus 48 h at 60°C. For visualization, the semi-thin sections were treated with toluidine blue staining. Images were recorded with Zeiss Libra 120 Transmission Electron Microscope.

Isolation protocol and microscopy of purified OMVs

Cultured *B. vulgatus* mpk suspension was slightly vortexed and centrifuged for 10 min at 10,000 x g. Supernatant was vacuum sterile filtrated and vesicles were pelletized for 2 h at 38,000 x g. Pellets were resuspended in sterile PBS and mixed with Iodixanol (OptiPrep) density gradient solution to prepare a 40% working solution. Subsequently, a dilutions serial with decreasing densities of 40%, 35%, 30%, 25%, 20% and 5% OptiPrep was assembled and stacked up in an appropriate ultracentrifugation tube. The samples were fractionated for 3 h at 280,000 x g and the developed, vesicle containing layers were collected. The isolated vesicles were natively held via the grid-on-drop technic and stained for negative imaging via TEM with uranyl acetate.

FITC-labeling of isolated vesicles

Isolated and purified OMVs were incubated in the dark for 1 h with 1 mg mL^{-1} Fluorescein isothiocyanate (FITC) at room temperature. Suspension was washed twice with HEPES buffer and centrifuged for 30 min at 52,000 x g. The pellet containing fluorescence-labeled OMVs was resuspended in sterile PBS.

Determination of OMVs concentration for stimulation experiments

To ensure a standardized application of vesicles and invariant stimulation of BMDCs and HEK cells, the exact concentration of every OMV batch was calculated and defined by the Pierce™ BCA Protein Assay Kit using the manufacturer's instructions.

Mice

C57BL/6J mice were purchased from Charles River Laboratories. Toll-like receptor 2 and 4 deficient (*Tlr2*$^{-/-}$, *Tlr4*$^{-/-}$) mice were provided by Jackson Laboratory. All animals were kept and bred under SPF conditions. For isolation of bone marrow, only female mice aged 6–12 weeks were used. Animal experiments were reviewed and approved by the responsible institutional review committee and the local authorities.

Statistics

Statistical analyses were performed using the unpaired student's t test. For comparing *in vitro* results, samples were considered to be biologically independent if the samples were generated from BMDCs from different mice. Differences were considered to be statistically significant if $p < 0.05$. Error bars, if shown, represent ± standard deviation (SD).

Disclosure of potential conflicts of Interest

No potential conflicts of interest were disclosed.

Funding

Work by J.S.F was supported by the DFG (DFG FR 2087/6–1, DFG FR 2087/8–1, CRC685, SPP1656), the DFG research training group 1708, the Bundesministerium für Bildung und Forschung (BMBF) and the German Center for Infection Research (DZIF).

References

[1] Frick JS and Autenrieth IB, The Gut Microflora and Its Variety of Roles in Health and Disease. Between Pathogenicity and Commensalism, U Dobrindt, JH Hacker, and C Svanborg Eds. Springer: Berlin Heidelberg 2012: p. 273-289

[2] Jandhyala SM, Talukdar R, Subramanyam C, Vuyyuru H, Sasikala M, Nageshwar Reddy D. Role of the normal gut microbiota. World J Gastroenterol 2015; 21(29):8787-803; PMID:26269668; https://doi.org/10.3748/wjg.v21.i29.8787

[3] Carding S, Verbeke K, Vipond DT, Corfe BM, Owen LJ. Dysbiosis of the gut microbiota in disease. Microb Ecol Health Dis 2015; 26:26191; PMID:25651997

[4] Manichanh C, Borruel N, Casellas F, Guarner F. The gut microbiota in IBD. Nat Rev Gastroenterol Hepatol 2012; 9(10):599-608; PMID:22907164; https://doi.org/10.1038/nrgastro.2012.152

[5] Gao J, Ma X, Gu W, Fu M, An J, Xing Y, Gao T, Li W, Liu Y. Novel functions of murine B1 cells: active phagocytic and microbicidal abilities. Eur J Immunol 2012; 42 (4):982-92; https://doi.org/10.1002/eji.201141519

[6] Scher JU, Sczesnak A, Longman RS, Segata N, Ubeda C, Bielski C, Rostron T, Cerundolo V, Pamer EG, Abramson SB, et al. Expansion of intestinal Prevotella copri correlates with enhanced susceptibility to arthritis. Elife 2013; 2:e01202; https://doi.org/10.7554/eLife.01202

[7] Ianiro G, Bibbò S, Gasbarrini A, Cammarota G. Therapeutic modulation of gut microbiota: current clinical applications and future perspectives. Curr Drug Targets 2014; 15(8):762-70; https://doi.org/10.2174/1389450115666140606111402

[8] Gallo A, Passaro G, Gasbarrini A, Landolfi R, Montalto M. Modulation of microbiota as treatment for intestinal inflammatory disorders: An uptodate. World J

Gastroenterol 2016; 22(32):7186-202; https://doi.org/10.3748/wjg.v22.i32.7186

[9] Hornef M. Pathogens, Commensal Symbionts, and Pathobionts: Discovery and Functional Effects on the Host. ILAR J 2015; 56(2):159-62; PMID:26323625; https://doi.org/10.1093/ilar/ilv007

[10] Perez-Lopez A, Behnsen J, Nuccio SP, Raffatellu M. Mucosal immunity to pathogenic intestinal bacteria. Nat Rev Immunol 2016; 16(3):135-48; https://doi.org/10.1038/nri.2015.17

[11] Frick JS, Zahir N, Müller M, Kahl F, Bechtold O, Lutz MB, Kirschning CJ, Reimann J, Jilge B, Bohn E, et al. Colitogenic and non-colitogenic commensal bacteria differentially trigger DC maturation and Th cell polarization: an important role for IL-6. Eur J Immunol 2006; 36 (6):1537-47; https://doi.org/10.1002/eji.200635840

[12] Muller M, Fink K, Geisel J, Kahl F, Jilge B, Reimann J, Mach N, Autenrieth IB, Frick JS. Intestinal colonization of IL-2 deficient mice with non-colitogenic B. vulgatus prevents DC maturation and T-cell polarization. PLoS One 2008; 3(6):e2376

[13] Waidmann M, Bechtold O, Frick JS, Lehr HA, Schubert S, Dobrindt U, Loeffler J, Bohn E, Autenrieth IB. Bacteroides vulgatus protects against Escherichia coli-induced colitis in gnotobiotic interleukin-2-deficient mice. Gastroenterology 2003; 125(1):162-177; https://doi.org/10.1016/S0016-5085(03)00672-3

[14] Gronbach K, Flade I, Holst O, Lindner B, Ruscheweyh HJ, Wittmann A, Menz S, Schwiertz A, Adam P, Stecher B, et al. Endotoxicity of lipopolysaccharide as a determinant of T-cell-mediated colitis induction in mice. Gastroenterology 2014; 146(3):765-75; https://doi.org/10.1053/j.gastro.2013.11.033

[15] Haurat MF, Aduse-Opoku J, Rangarajan M, Dorobantu L, Gray MR, Curtis MA, Feldman MF. Selective sorting of cargo proteins into bacterial membrane vesicles. J Biol Chem 2011; 286(2):1269-76; https://doi.org/10.1074/jbc.M110.185744

[16] Jang SC, Kim SR, Yoon YJ, Park KS, Kim JH, Lee J, Kim OY, Choi EJ, Kim DK, Choi DS, et al. In vivo kinetic biodistribution of nano-sized outer membrane vesicles derived from bacteria. Small 2015; 11(4):456-61; https://doi.org/10.1002/smll.201401803

[17] Bomberger JM, Maceachran DP, Coutermarsh BA, Ye S, O'Toole GA, Stanton BA. Long-distance delivery of bacterial virulence factors by Pseudomonas aeruginosa outer membrane vesicles. PLoS Pathog 2009; 5(4):e1000382; https://doi.org/10.1371/journal.ppat.1000382

[18] Ratner D, Orning MP, Lien E. Bacterial secretion systems and regulation of inflammasome activation. J Leukoc Biol 2016; 101(1):165-181; PMID:27810946

[19] Dale C, Moran NA. Molecular interactions between bacterial symbionts and their hosts. Cell 2006; 126 (3):453-65; PMID:16901780; https://doi.org/10.1016/j.cell.2006.07.014

[20] Haurat MF, Elhenawy W, Feldman MF. Prokaryotic membrane vesicles: new insights on biogenesis and biological roles. Biol Chem 2015; 396(2):95-109; PMID:25178905; https://doi.org/10.1515/hsz-2014-0183

[21] Kuehn MJ, Kesty NC. Bacterial outer membrane vesicles and the host-pathogen interaction. Genes Dev 2005; 19 (22):2645-55; PMID:16291643; https://doi.org/10.1101/gad.1299905

[22] Mogensen TH. Pathogen recognition and inflammatory signaling in innate immune defenses. Clin Microbiol Rev 2009; 22(2):240-73, Table of Contents; PMID:19366914; https://doi.org/10.1128/CMR.00046-08

[23] Chu H and Mazmanian SK. Innate immune recognition of the microbiota promotes host-microbial symbiosis. Nat Immunol 2013; 14(7):668-75; PMID:23778794; https://doi.org/10.1038/ni.2635

[24] Shen Y, Giardino Torchia ML, Lawson GW, Karp CL, Ashwell JD, Mazmanian SK. Outer membrane vesicles of a human commensal mediate immune regulation and disease protection. Cell Host Microbe 2012; 12(4):509-20; https://doi.org/10.1016/j.chom.2012.08.004

[25] Hickey CA, Kuhn KA, Donermeyer DL, Porter NT, Jin C, Cameron EA, Jung H, Kaiko GE, Wegorzewska M, Malvin NP, et al. Colitogenic Bacteroides thetaiotaomicron Antigens Access Host Immune Cells in a Sulfatase-Dependent Manner via Outer Membrane Vesicles. Cell Host Microbe 2015; 17(5):672-80; https://doi.org/10.1016/j.chom.2015.04.002

[26] Duboc H, Rajca S, Rainteau D, Benarous D, Maubert MA, Quervain E, Thomas G, Barbu V, Humbert L, Despras G, et al. Connecting dysbiosis, bile-acid dysmetabolism and gut inflammation in inflammatory bowel diseases. Gut 2013; 62(4):531-9; https://doi.org/10.1136/gutjnl-2012-302578

[27] Kaparakis-Liaskos M, Ferrero RL. Immune modulation by bacterial outer membrane vesicles. Nat Rev Immunol 2015; 15(6):375-87; PMID:25976515; https://doi.org/10.1038/nri3837

[28] Schwechheimer C, Kuehn MJ. Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions. Nat Rev Microbiol 2015; 13(10):605-19; PMID:26373371; https://doi.org/10.1038/nrmicro3525

[29] Muraca M, Putignani L, Fierabracci A, Teti A, Perilongo G. Gut microbiota-derived outer membrane vesicles: under-recognized major players in health and disease? Discov Med 2015; 19(106):343-8; PMID:26105697

[30] Steimle A, Frick JS. Molecular Mechanisms of Induction of Tolerant and Tolerogenic Intestinal Dendritic Cells in Mice. J Immunol Res 2016; 2016:1958650; PMID:26981546; https://doi.org/10.1155/2016/1958650

[31] McBroom AJ, Kuehn MJ. Release of outer membrane vesicles by Gram-negative bacteria is a novel envelope stress response. Mol Microbiol 2007; 63(2):545-58; PMID:17163978; https://doi.org/10.1111/j.1365-2958.2006.05522.x

[32] Steimle A, Gronbach K, Beifuss B, Schäfer A, Harmening R, Bender A, Maerz JK, Lange A, Michaelis L, Maurer A, et al. Symbiotic gut commensal bacteria act as host

cathepsin S activity regulators. J Autoimmun 2016; 75:82-95; https://doi.org/10.1016/j.jaut.2016.07.009

[33] Lutz MB, Schuler G. Immature, semi-mature and fully mature dendritic cells: which signals induce tolerance or immunity? Trends Immunol 2002; 23(9):445-9; PMID:12200066; https://doi.org/10.1016/S1471-4906(02)02281-0

[34] Dudek AM, Martin S, Garg AD, Agostinis P. Immature, Semi-Mature, and Fully Mature Dendritic Cells: Toward a DC-Cancer Cells Interface That Augments Anticancer Immunity. Front Immunol 2013; 4:438; https://doi.org/10.3389/fimmu.2013.00438

[35] Gerlach AM, Steimle A, Krampen L, Wittmann A, Gronbach K, Geisel J, Autenrieth IB, Frick JS. Role of CD40 ligation in dendritic cell semimaturation. BMC Immunol 2012; 13:22; https://doi.org/10.1186/1471-2172-13-22

[36] Abreu MT. Toll-like receptor signalling in the intestinal epithelium: how bacterial recognition shapes intestinal function. Nat Rev Immunol 2010; 10(2):131-44; PMID:20098461; https://doi.org/10.1038/nri2707

[37] Janssens S, Beyaert R. Role of Toll-like receptors in pathogen recognition. Clin Microbiol Rev 2003; 16(4):637-46; PMID:14557290; https://doi.org/10.1128/CMR.16.4.637-646.2003

[38] Jimenez-Dalmaroni MJ, Gerswhin ME, Adamopoulos IE. The critical role of toll-like receptors–From microbial recognition to autoimmunity: A comprehensive review. Autoimmun Rev 2016; 15(1):1-8; PMID:26299984; https://doi.org/10.1016/j.autrev.2015.08.009

[39] Frosali S, Pagliari D, Gambassi G, Landolfi R, Pandolfi F, Cianci R. How the Intricate Interaction among Toll-Like Receptors, Microbiota, and Intestinal Immunity Can Influence Gastrointestinal Pathology. J Immunol Res 2015; 2015:489821; https://doi.org/10.1155/2015/489821

[40] Alexander C, Rietschel ET. Bacterial lipopolysaccharides and innate immunity. J Endotoxin Res 2001; 7(3):167-202; PMID:11581570

[41] Steimle A, Autenrieth IB, Frick JS. Structure and function: Lipid A modifications in commensals and pathogens. Int J Med Microbiol 2016; 306(5):290-301; PMID:27009633; https://doi.org/10.1016/j.ijmm.2016.03.001

[42] Molinaro A, Holst O, Di Lorenzo F, Callaghan M, Nurisso A, D'Errico G, Zamyatina A, Peri F, Berisio R, Jerala R, et al. Chemistry of lipid A: at the heart of innate immunity. Chemistry 2015; 21(2):500-19; https://doi.org/10.1002/chem.201403923

[43] Akira S, Takeda K. Toll-like receptor signalling. Nat Rev Immunol 2004; 4(7):499-511; PMID:15229469; https://doi.org/10.1038/nri1391

[44] Takeda K, Akira S. Toll-like receptors in innate immunity. Int Immunol 2005; 17(1):1-14; PMID:15585605; https://doi.org/10.1093/intimm/dxh186

[45] Liew FY, Patel M, Xu D. Toll-like receptor 2 signalling and inflammation. Ann Rheum Dis 2005; 64(Suppl 4):iv104-5; PMID:16239376

[46] Lange A, Beier S, Steimle A, Autenrieth IB, Huson DH, Frick JS. Extensive Mobilome-Driven Genome Diversification in Mouse Gut-Associated Bacteroides vulgatus mpk. Genome Biol Evol 2016; 8(4):1197-207; https://doi.org/10.1093/gbe/evw070

[47] Hansson GC. Role of mucus layers in gut infection and inflammation. Curr Opin Microbiol 2012; 15(1):57-62; PMID:22177113; https://doi.org/10.1016/j.mib.2011.11.002

[48] Canas MA, Giménez R, Fábrega MJ, Toloza L, Baldomà L, Badia J. Outer Membrane Vesicles from the Probiotic Escherichia coli Nissle 1917 and the Commensal ECOR12 Enter Intestinal Epithelial Cells via Clathrin-Dependent Endocytosis and Elicit Differential Effects on DNA Damage. PLoS One 2016; 11(8):e0160374; https://doi.org/10.1371/journal.pone.0160374

[49] Bonnington KE, Kuehn MJ. Protein selection and export via outer membrane vesicles. Biochim Biophys Acta 2014; 1843(8):1612-9; PMID:24370777; https://doi.org/10.1016/j.bbamcr.2013.12.011

[50] Olsen I, Amano A. Outer membrane vesicles - offensive weapons or good Samaritans? J Oral Microbiol 2015; 7:27468; PMID:25840612; https://doi.org/10.3402/jom.v7.27468

[51] Deatherage BL, Cookson BT. Membrane vesicle release in bacteria, eukaryotes, and archaea: a conserved yet underappreciated aspect of microbial life. Infect Immun 2012; 80(6):1948-57; PMID:22409932; https://doi.org/10.1128/IAI.06014-11

[52] Ercolini AM, Miller SD. The role of infections in autoimmune disease. Clin Exp Immunol 2009; 155(1):1-15; PMID:19076824; https://doi.org/10.1111/j.1365-2249.2008.03834.x

[53] Shi Y, Mu L. An expanding stage for commensal microbes in host immune regulation. Cell Mol Immunol 2017; 14(4):339-348; https://doi.org/10.1038/cmi.2016.64

[54] Lin L, Zhang J. Role of intestinal microbiota and metabolites on gut homeostasis and human diseases. BMC Immunol 2017; 18(1):2; PMID:28061847; https://doi.org/10.1186/s12865-016-0187-3

[55] Mazmanian SK, Round JL, Kasper DL. A microbial symbiosis factor prevents intestinal inflammatory disease. Nature 2008; 453(7195):620-625; PMID:18509436; https://doi.org/10.1038/nature07008

[56] Deatherage BL, Lara JC, Bergsbaken T, Rassoulian Barrett SL, Lara S, Cookson BT. Biogenesis of bacterial membrane vesicles. Mol Microbiol 2009; 72(6):1395-407; https://doi.org/10.1111/j.1365-2958.2009.06731.x

[57] Round JL, Mazmanian SK. Inducible Foxp3+ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. Proc Natl Acad Sci U S A 2010; 107(27):12204-9; PMID:20566854; https://doi.org/10.1073/pnas.0909122107

[58] Wang Q, McLoughlin RM, Cobb BA, Charrel-Dennis M, Zaleski KJ, Golenbock D, Tzianabos AO, Kasper DL. A bacterial carbohydrate links innate and adaptive responses through Toll-like receptor 2. J Exp Med 2006;

203(13):2853-63; PMID:17178920; https://doi.org/10.1084/jem.20062008

[59] Brooke MS. Conversion of immunological paralysis to immunity by endotoxin. Nature 1965; 206(984):635-6; PMID:4378605; https://doi.org/10.1038/206635a0

[60] Greisman SE, Young EJ, Workman JB, Ollodart RM, Hornick RB. Mechanisms of endotoxin tolerance. The role of the spleen. J Clin Invest 1975; 56(6):1597-1607; https://doi.org/10.1172/JCI108242

[61] Broad A, Jones DE, Kirby JA. Toll-like receptor (TLR) response tolerance: a key physiological "damage limitation" effect and an important potential opportunity for therapy. Curr Med Chem 2006; 13 (21):2487-502; PMID:17017906; https://doi.org/10.2174/092986706778201675

[62] Sato S, Nomura F, Kawai T, Takeuchi O, Mühlradt PF, Takeda K, Akira S. Synergy and cross-tolerance between toll-like receptor (TLR) 2- and TLR4-mediated signaling pathways. J Immunol 2000; 165(12):7096-101; https://doi.org/10.4049/jimmunol.165.12.7096

[63] Hayashi T, Gray CS, Chan M, Tawatao RI, Ronacher L, McGargill MA, Datta SK, Carson DA, Corr M. Prevention of autoimmune disease by induction of tolerance to Toll-like receptor 7. Proc Natl Acad Sci U S A 2009; 106 (8):2764-9; https://doi.org/10.1073/pnas.0813037106

[64] Tan RS, Ho B, Leung BP, Ding JL. TLR cross-talk confers specificity to innate immunity. Int Rev Immunol 2014; 33(6):443-53; https://doi.org/10.3109/08830185.2014.921164

[65] Hevia A, Milani C, López P, Cuervo A, Arboleya S, Duranti S, Turroni F, González S, Suárez A, Gueimonde M, et al. Intestinal dysbiosis associated with systemic lupus erythematosus. MBio 2014; 5(5):e01548-14; https://doi.org/10.1128/mBio.01548-14

Manuscript in preparation, publication c

Lange A., Tschörner L., Beck C., Beier S., Schäfer A., Frick J.S.: Survival strategies of *E. coli* mpk in the inflamed intestine.

Survival strategies of *Escherichia* coli mpk in the inflamed intestine

Anna Lange[1], Leonie Tschörner[1], Andrea Schäfer[1], Christian Beck[1], Sina Beier[2], Julia-Stefanie Frick[1]*

[1]Interfacultary Institute for Microbiology and Infection Medicine, Department for Medical Microbiology and Hygiene, University of Tübingen, Germany

[2]Algorithms in Bioinformatics, ZBIT Center for Bioinformatics, University of Tübingen, Germany

*Corresponding author: E-mail: julia-stefanie.frick@med.uni-tuebingen.de

Introduction

Inflammatory bowel diseases (IBD) are characterized to be multifactorial (Kaser et al., 2010). Its exact etiology is still not well understood, but it is reported that the development and course of the disease involves host genetic susceptibility and the intestinal microbiota (Kaser et al., 2010; Vejborg et al., 2011). Several studies revealed a compositional shift of the intestinal microbiota in IBD patients, including a decrease in *Bacteroidetes* and *Firmicutes* and a higher abundance of *Enterobacteriaceae* such as *Escherichia coli* (Kaser et al., 2010; Spor et al., 2011). During intestinal homeostasis *E. coli* are generally poor gut colonizers but their growth is boosted by inflammatory conditions (Spees et al., 2013b). Inflammation lowers the colonization resistance of *E. coli* and the other resident microbes, and *E. coli* expansion is improved (Spees et al., 2013b). *E. coli* and *Salmonella*, both members of the *Enterobacteriaceae* family, share a multitude of genes suggesting that they have a close common evolutionary origin (Karberg et al., 2011). It was

reported that *Salmonella* gains a growth advantage in the inflamed intestine by the consumption of ethanolamine or 1,2-propanediol (Thiennimitr et al., 2011; Faber et al., 2017). The ability to consume ethanolamine is mediated by altered growth conditions in the intestine; *Salmonella* is able to occupy a specific nutritional niche generated by anaerobic respiration with tetrathionate as electron acceptor (Thiennimitr et al., 2011). For 1,2-propanediol utilization anaerobic respiration is not required (Faber et al., 2017).

The same pathways for ethanolamine and 1,2-propanediol utilization could be identified in our commensal *E. coli* mpk2 mouse fecal isolate. Furthermore, we could identify the gene clusters for nitrate respiration, which could enable *E. coli* to fill the respiration niche during inflammation like *Salmonella*.

Materials and methods

Bacterial strains

E. coli mpk and *B. vulgatus* mpk were isolated from mouse feces of a healthy individual (Waidmann et al., 2003; Lange et al., 2016). All *E. coli* strains used in this study are listed in Table 1. The strains were growth aerobically at 37 °C in Luria-Bertani (LB) medium (Becton Dickinson) or on LB agar (15 g/l agar) supplemented with the appropriate antibiotics (tetracycline 10 µg/ml, kanamycin 50 µg/ml, chloramphenicol 25 µg/ml). *B. vulgatus* mpk was cultivated anaerobically at 37 °C in brain heart infusion (BHI) medium (Oxoid) or on BHI agar (15 g/l agar).

E. coli mpk whole genome sequencing, draft assembly and annotation

E. coli mpk was initially isolated from mouse feces. High molecular weight DNA was isolated using Genomic-tip 100/G (Qiagen). Sequencing library was prepared using the TruSeq Nano kit (Illumina). The 750 bp insert library was sequenced on Illumina MiSeq creating 2x 300 bp PE-reads. Sequencing coverage was 150-fold. Reads

were quality assessed using FastQC (Andrews, 2010). The reads were assembled using SPAdes genome assembler (Bankevich et al., 2012) resulting in 92 contigs. Automated annotation was performed using the RAST server (Aziz et al., 2008).

Identification of 1,2-propanediol and ethanolamine utilization pathways and comparative genomics

We screened different genomes for availability of 1,2-propanediol and ethanolamine utilization pathways, and if all clusters are complete. Genome sequences or assemblies were obtained from NCBI database. We included in our analysis *E. coli* IBD isolates (*E. coli* LF82: NC_011993.1, *E. coli* UM146: NC_017632.1, *E. coli* NRG857c: NC_017634.1), commensal *E. coli* strains (*E. coli* K12 MG1655: NC_000913.3, E. coli Nissle 1917: NZ_CP007799.1, E. coli HS: NC_009800.1, *E. coli* ED1a: NC_011745.1, *E. coli* NCTC86: CP019778.1, *E. coli* SE11: NC_011415.1, *E. coli* SE15: NC_013654.1), pathogenic *E. coli* strains (*E. coli* 0127:H6 E2348/69: NC_011601.1, *E. coli* O104:H4 str. C227-11: NZ_CP011331.1, *E. coli* O104:H4 str. 2011C-3493: NC_018658.1), and *Salmonella* strains (*Salmonella* Typhimurium LT2: NC_003197.2, *Salmonella* Typhi: NC_004631.1). The nucleotide gene clusters of both pathways were extracted and a multiple sequence alignment was performed with Clustal Omega. The Newick phylogenetic tree data was displayed with Unipro UGENE.

E. coli mpk mutant construction

E. coli mpk ΔeutR was generated using LambdaRed recombination described by Datsenko and Wanner (Datsenko and Wanner, 2000). See supplemental information for cloning strategies and knockout generation.

All other knockouts were generated using suicide plasmids. Suicide plasmids were created to encode the desired deletion plus a 1000 bp flanking sequence on each side, which are homologous to the chromosomal destination locus. All plasmids and

primers are listed in Table 2 and 3. The plasmids were clones *via* Gibson cloning (Gibson et al., 2009). The Gibson reactions were transformed into *E. coli* 116 λpir+ and then transformed into *E. coli* β2168 Δnic35 that lacks the asd gene and is therefore only able to grow on medium supplemented with diaminoplimelic acid (DAP). *E. coli* β2168 Δ*nic35* was able to mate with *E. coli* mpk and mediated the transfer of the suicide plasmid. Recipient strains that had the suicide plasmid integrated into the chromosome were recovered on LB plates containing tetracyclin. Sucrose plates were used to select the strains after the second crossover. Deletions of the target genes were confirmed by PCR.

The knockout strains were transformed with low-copy pHP45Ω-cm and wild type *E. coli* mpk strain was transformed with pHP45Ω-km to enable the differentiation between intrinsic *E. coli* strains during in vitro and in vivo competition (Sassone-Corsi et al., 2016).

Aerobic growth measurements

E. coli mpk and all respective knockout strains (Table 1) were grown in different media for comparison. They were grown at 37 °C in M9 minimal medium (Na_2HPO_4 33.7 mM, KH_2PO_4 22.0 mM, NaCl 8.55 mM, NH_4Cl 9.35 mM, $MgSO_4$ 1 mM, $CaCl_2$ 0.3 mM, vitamins: cobalamin 200 nM, Thiamin 5 mg/L, trace metals: 0.1 µM $ZnSO_4$, 0.045 µM $FeSO_4$, 0.2 µM Na_2MoO_4, 2 µM $MnSO_4$, 0.1 µM $CuSO_4$, 3 µM $CoSO_4$, 0.1 µM $NiSO_4$) supplemented with 0.5% of different carbon sources (ethanolamine, 1,2-propanediol) or 15 mg/ml mouse mucus. Aerobic growth measurements were performed continuously in a plate reader (Tecan Infinite) using transparent round bottom 96-well plates.

Anaerobic competition assay

Anaerobic competition assays were performed to evaluate survival advantages of *E. coli* mpk and the knockout strains. Cultures were inoculated with a 1:1 mixture of *E.*

coli mpk pHP45Ω-km and a knockout strain containing pHP45Ω-cm. The strain mixtures were left for competition in M9 medium supplemented with 15 mg/ml mouse mucus. CFU of the kanamycin resistant *E. coli* mpk wild type and the chloramphenicol resistant knockout strains were determined after 24 h from the cultures.

Animal experiments

All protocols and experiments in this study were reviewed and approved by the responsible Institutional Review Committee and the local authorities.

Female C57BL/6N wild-type SPF mice aged 8-10 weeks were obtained from Charles River. Mice were initially administered with $2x10^8$ CFU of a 1:1 mixture of the *E. coli* mpk strains. Bacteria administration was performed orally via the drinking water. Two days after initial bacteria administration the mice received $2x10^8$ CFU of a 1:1 bacterial mixture in drinking water containing 3 % dextran sodium sulfate (DSS) (day 1). The drinking water containing the bacteria and DSS was replaced on day 2 and 4. On day 5 the drinking water was exchanged for drinking water containing only 3 % DSS. Weight loss and disease activity index (DAI) was determined between days 3-7 to evaluate disease progression. Scores for DAI are defined as follows: stool consistency: 0 (normal), 1 (loose stool), and 3 (diarrhea); blood in stool: 0 (no blood), 3 (positive); anus: 1 (inflamed), 2 (bloody), severe bleeding and swollen anus (3); behavior: 0 (normal), 2 (less activity, less stretching movements), 3 (hardly visible movement, poor posture, humpbacked); fur: 0 (normal), rough (3).

Fecal CFU were determined over the whole course of the experiment (day 1 and day 3-6). Mice were sacrificed after 7 days of DSS treatment. After preparation CFU were determined from ileal, caecal and colonic content.

Histopathology

Ileal and caecal tissues were fixed in 4 % paraformaldehyde and stained with hematoxylin and eosin (HE). The score is evaluated from the severity of inflammation (I): 0 (rare inflammatory cells in the lamina propria), 1 (increased number of inflammatory cells), 2 (confluence of inflammatory cells extending to the submucosa), 3 (transmural extension of the inflammatory cell infiltrate); and extent of injury (II): 0 (no mucosal damage), 1(discrete lymphoepithelial lesions), 2 (surface mucosal erosion), 3 (widespread mucosal ulceration and extension through deeper bowel wall structures). The severity of inflammation scores (I) and extent of injury (II) were added and a mean was calculated.

Results

Ethanolamine but not 1,2-propanediol utilization is ubiquitous in different *E. coli* strains

E. coli mpk, a mouse commensal fecal isolate, was evaluated to be nonpathogenic; however, it was demonstrated to cause intestinal inflammation in *Il2*$^{-/-}$ mono-associated gnotobiotic mice (Waidmann et al., 2003). Furthermore, commensal *E. coli* was found to accumulate during intestinal inflammation (Spees et al., 2013a; Rakitina et al., 2017).

To better characterize *E. coli* mpk we performed whole genome sequencing. Sequencing revealed that the *E. coli* mpk genome harbors an ethanolamine utilization gene cluster (*eutRKLCBAHGJENMDTQPS*), a 1,2-propanediol degradation cluster (*pocRpduFABCDEGHJKLMNOPQSTUV*) and contains a fucose and rhamnose degradation pathway, which enable the production of the fermentation end-product 1,2-propanediol. It further contains gene clusters for nitrate respiration (*narGHJI, narZYWV, napFDAGHBC*), which might enable the use of RNS as terminal

electron acceptors during anaerobic respiration. Additionally, *E. coli* contains a fumarate reductase as an alternative electron acceptor (Unden and Bongaerts, 1997). However, it does not contain a tetrathionate respiration pathway like *Salmonella* Typhimurium. Further *E. coli* mpk encodes a cobalamin synthesis operon which produces the necessary cofactor for propanediol dehydratase, the initial enzyme of 1,2-propanediol degradation pathway (Staib and Fuchs, 2015).

We could identify many metabolic pathways in nonpathogenic *E. coli* mpk that are also found in the enteropathogen *Salmonella* Typhimurium. In order to prove that these pathways only occur in *E. coli* mpk several other *E. coli* strains were screened for the presence of ethanolamine and 1,2-propanediol utilization pathways. We analyzed *E. coli* isolates from IBD patients (LF82, UM146, NRG857c), commensal *E. coli* strains (K12 MG1655, Nissle 1917, HS, ED1a, NCTC86, SE11, SE15), pathogenic *E. coli* strains (0127:H6 E2348/69, O104:H4 str. C227-11, O104:H4 str. 2011C-3493), and *Salmonella* strains (*Salmonella enterica* serovar Typhimurium LT2, *Salmonella enterica* serovar Typhi, *Salmonella bongori*). The nucleotide sequences of both pathways were extracted if they are present in the genome and a multiple sequence alignment was performed.

A cluster for ethanolamine utilization was identified in all analyzed *E. coli* strains; however the commensal *E. coli* MG1655 strain encoded an ethanolamine degradation pathway interrupted by a prophage. This insertion may have resulted in a dysfunctional pathway. *S.* Typhimurium and *S.* Typhi also harbor this ethanolamine cluster, whereas no such cluster was identified in *S. bongori*.

In order to analyze how closely related the ethanolamine utilization clusters are among different nonpathogenic and pathogenic strains of the *Enterobacteriaceae* family is, a multiple sequence alignment was performed, including the commensal *E. coli* strains, *E. coli* isolates from IBD patients, pathogenic *E. coli* strains, and *S.*

Typhimurium and *S.* Typhi. The comparison revealed a clustering of the tested strains into three groups (Figure 1A). The first group consists of the *Salmonella* and enterohemorrhagic *E. coli* (EHEC) strains and *E. coli* UM146, which was isolated from an IBD patient. *E. coli* mpk and other nonpathogenic *E. coli* strains belong into the second group. A third group consists of additional nonpathogenic *E. coli* strains. It includes *E. coli* NRG857c and LF82, which have been isolated from IBD patients, the probiotic strain *E. coli* Nissle 1917 and *E. coli* ED1a. *E. coli* SE15 cannot be assigned to one of the groups. It is rather located between the two groups consisting of nonpathogenic *E. coli* strains. Interestingly, the two IBD patient isolates *E. coli* LF82 and NRG857c are very closely related within a commensal *E. coli* group, and *E. coli* mpk is related to nonpathogenic strains. These findings are in accordance with reports that find commensal *E. coli* to accumulate during intestinal inflammation. The ethanolamine utilization cluster of *E. coli* UM146 another isolate from an IBD patient was identified to be closer related to pathogenic strains although being a commensal. This supports the pathobiotic character of commensals that act “pathogenic” during colitis by similar metabolic properties.

We could identify an 1,2-propanediol utilization cluster in the *S. enterica* serovars Typhimurium and Typhi; however an 1,2-propanediol utilization cluster was absent from the *S. bongori* genome. The 1,2-propanediol utilization gene cluster was identified in pathogenic *E. coli* O104:H4 str. 2011C-3493, and *E. coli* NRG857c and LF82, which have been isolated from IBD patients. This leads to the assumption that the presence of this cluster is not a ubiquitous feature such as the presence of an ethanolamine utilization cluster.

The 1,2-propanediol utilization clusters of the *Salmonella* serovars is closely related (Figure 1B). The 1,2-propanediol gene cluster in *E. coli* NRG857c and LF82 is very closely related. The cluster of *E. coli* mpk and EHEC strain 2011C-3493 cannot be

grouped with IBD patient-related strains or the *Salmonella* 1,2-propandiol utilization cluster.

Ethanolamine and 1,2-propanediol utilization contribute to *E. coli in vitro* competitiveness

In order to determine if either ethanolamine, 1,2-propanediol utilization or both pathways confer a growth advantage to *E. coli* mpk, we constructed different knockout strains. We constructed a strain lacking the *eutR* gene, which regulates expression of the ethanolamine utilization operon, a knockout which lacks *pduD*, which is part of the *pduCDE*-encoded 1,2-propanediol dehydratase, the initial degradation enzyme, and a double knockout strain with depleted ethanolamine and 1,2-propanediol degradation pathways. All *E. coli* strains showed no growth defects aerobically in rich medium (data not shown). However, the strains poorly grew in aerobic growth assays using minimal medium supplemented with 5 mM ethanolamine or 1,2-propanediol (Supplementary Figure S1A and B). In order to mimic the environmental conditions in the intestine and the available nutrients we used a minimal medium supplemented with scraped mucus from either the ileum, or from caecum and colon. Growth of all *E. coli* stains was assayed aerobically in a plate reader. Both minimal mucus media provided the strains enough substrates to grow. The ileum medium allowed for better growth compared to the caecum/colon medium (Supplementary Figure S1D and E). Interestingly, both mucus media showed similar growth characteristics of the wild type and knockout strains. The wild type strain started to grow earlier and *E. coli* Δ*eutR* Δ*pduD* later; both single knockout strains showed less effective growth than the wild type. In the ileum medium an intermediate growth phenotype of the single knockout strains can be observed, which started to grow later than the wild type but earlier than the double knockout strain.

As we could observe different growth characteristics in the mucus media, we decided to assess whether the mucus media allows anaerobic competitive growth analysis in vitro to evaluate fitness advantage of all knockouts compared to the wild type. A 1:1 mixture of *E. coli* wild type and a mutant strain was used to inoculate the mucus media. The competitive cultures were incubated anaerobically and 40mM potassium nitrate was added as electron acceptor. After for 24 h the competitive indices (CIs) were determined from the colony-forming units (CFUs) determined from the wild type (kanamycin resistant) or from the mutant (chloramphenicol resistant). CI is calculated by dividing the CFUs of the wild type by the CFUs of the mutant. The competition assays in ileum and caecum/colon medium suggest that aerobic conditions do not resemble the in vivo environment (Figure 2A and B, left panels), because all determined CIs were approximately 1. However, the anaerobic conditions seem to favor the competition (Figure 2A and B, right panels). The single knockouts exhibit no growth advantage weather the electron acceptor was added or not. The *E. coli* Δ*eutR* Δ*pduD* strain showed a significantly enhanced CI, when the culture was anaerobically incubated without electron acceptor. The CI was significantly higher compared to aerobic wt/ Δ*eutR* Δ*pduD* culture, and compared to the anaerobic wt/ Δ*eutR* and wt/ Δ*pduD* cultures. Similar results could be observed in the caecum/colon medium.

Ethanolamine and 1,2-propanediol utilization support *E. coli* mpk survival during acute intestinal inflammation

We orally administered C57BL/6 SPF mice with a 1:1 mixture of E. coli wild type/ *E. coli* wild type (reference group), or *E. coli* wild type/ *E. coli* Δ*eutR* Δ*pduD* (experimental group) in the drinking water and compared how the strains were maintained in the intestine i.e. the faeces, and evaluated the fitness of the strains for 7 days of acute DSS-induced inflammation. Stool was collected daily and CFUs of

the bacterial strains was determined, at the same time the progress and degree of intestinal inflammation was monitored by determination of the bodyweight and the disease activity index (DAI). Between day 0, the initial day of DSS administration, and day 3 an increase of the total bacterial CFUs was observed (Figure 3A and B), which correlates with the first signs of intestinal inflammation (Figure 3C). Afterwards at day 5 the bacterial load decreased and in the wt/ Δ*eutR* Δ*pduD*-administered group we could observe slightly lower CFUs of the knockout strain. This matches with the observation of the further increased DAI and the increased incidences of soft stool or diarrhea, which leads to a flushing out of bacteria. The effect of decreasing Δ*eutR* Δ*pduD* CFUs was enhanced at day 6, and at the end of the experiment the wt and mutant CFU populations were completely separated. On the other hand, the wt/ wt-administered group maintained similar CFU levels of both different labeled wt strains throughout the experiment. At the end of the experiment, we determined CFUs not only from colonic faeces, but also from ileal and caecal content. Interestingly, we could identify differences in the recovered CFUs from each strain, assuming varying competitive conditions in different intestinal sections. In the ileal contents the CI of wt/ wt was approximately 1, and CI of the wt/ Δ*eutR* Δ*pduD* was approximately 60, but the differences of the means were not significant (Figure 3D). Whereas, in the caecal and colonic content we could observe significant differences of CI means from the wt/ wt reference group and the wt/ Δ*eutR* Δ*pduD* group (Figure 3E and F). Therefore, we propose increasing competitive conditions from the ileum to caecum and colon among *E. coli* strains. *E. coli* wild type benefits from ethanolamine and 1,2-propanediol utilization in the inflamed intestine and E. coli Δ*eutR* Δ*pduD* was significantly less competitive.

Discussion

Many different bacterial genomes encode ethanolamine utilization clusters especially intestinal-associated pathogens such as *Salmonella*, but also commensal *Escherichia* and *Pseudomonas* sp. In addition this gene cluster is encoded in genomes of Gram-positive commensal *Enterococcus* and *Clostridium* sp. (Garsin, 2010). The intestine constantly provides a certain ethanolamine source as a compound of phosphatidylethanolamine, which is part of both mammalian and bacterial cell membranes (Kaval and Garsin, 2018). Physiological concentrations of ethanolamine in the intestine are between 1 and 2 mM (Kaval et al., 2018).

1,2-propanediol utilization is a common degradation pathway in *Salmonella enterica* serovars (Faber et al., 2017). This gene cluster is also encoded in some pathogenic and IBD-related *E. coli* strains. Both ethanolamine and 1,2-propanediol utilization are absent from *Salmonella bongori* the second occurring species of the *Salmonella* genus (Porwollik et al., 2002; Faber et al., 2017). *Salmonella* was reported to evolutionary diverge from *E. coli* and *S. enterica* was separated from *S. bongori* (Porwollik et al., 2002), which does not contain both pathways. Interestingly, intensive horizontal gene transfer (HGT) between *Enterobacteriaceae* species is quite common, especially between *S. enterica* and *E. coli* (Gordienko et al., 2013). Thus, we propose that the pathways became part of *S. enterica* genome through evolution: *S. bongori* might lost ethanolamine and 1,2-propanediol utilization pathways after splitting from *E. coli*, and *S. enterica* gained both gene clusters by HGT. Some *E. coli* strains might have lost the 1,2-propanediol degradation cluster because it is not that frequently occurring in *E. coli* strains, and a pathway which is not a big fitness factor when colonizing a homeostatic intestine. We and others suggest that 1,2-propanediol degradation is more prevalent in IBD isolates (Faber et al., 2017). In order to confirm this observation more IBD isolates, other commensal and also

pathogenic strains need to be analyzed. The fact that 1,2-propanediol utilization clusters are more common in IBD isolates might be due to the characteristic outgrowth of *Enterobacteriaceae* especially *E. coli* during inflammation (Winter and Baumler, 2014) and the resulting assumption that only the most adapted *E. coli* strains have a survival and fitness advantage in such conditions. It might be possible that such strains are not able to be isolated and identified from faeces of healthy donors as *E. coli* are completely underrepresented. Another reason could be that strains lacking pathways conferring advantages during inflammation might have a growth benefit in a homeostatic environment such as *Enterococcus* faecalis; its ability to colonize the gastrointestinal tract was enhanced when ethanolamine utilization was eliminated (Kaval et al., 2018).

Observations of *Salmonella* fitness factors during intestinal inflammation were obtained from analysis of caecal and colonic faeces and therefore restricted to the large intestine (Thiennimitr et al., 2011; Faber et al., 2017) but only IBD patients suffering from ulcerative colitis show inflammatory symptoms restricted to the colon (Forbes et al., 2016). Crohn´s disease (CD) patients on the other hand, suffer from inflammation throughout the whole gastrointestinal tract. All *E. coli* strains we included in our previous *in silico* analysis were isolated from CD patients, where inflammation occurs also in the small intestine (Forbes et al., 2016). Thus, we intended to include analysis of the ileum to our experimental work. In vitro pre-evaluation with mucus-containing medium we could already observe differences between small and large intestinal supplements. In particular, supplementation of ileal scrapings and incubation under anaerobic conditions provided better growth conditions for the *E. coli* wild type strains competing against Δ*eutR* Δ*pduD*. It was reported that ethanolamine and 1,2-propanediol pathways inhibit each other, and only one of these pathways could active at a time (Faber et al., 2017) but from this in

vitro competition data it might be possible that lacking both pathways was an additive nutritional disadvantage. It might be possible that the ability to switch between the pathways could be advantageous. Our results obtained from in vivo observations of wt/ Δ*eutR* Δ*pduD* competition during acute intestinal inflammation in mice verified a significant growth advantage of the wild type strain. However, the competitive advantage of the remaining single knockout strains remains to be analyzed. For the first time we could show that onset of bacterial competition can be observed initially in the terminal ileum, and is further increased in the caecum. It has been reported that also in CD patients with ileal inflammation the abundance of Enterobacteriaceae was observed to be elevated (Naftali et al., 2016). The CIs observed in the colonic faeces might be the result of competition occurred in previous intestinal segments.

References

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. [Online]. Available: http://www.bioinformatics.babraham.ac.uk/projects/fastqc [Accessed May 6th 2015].

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., et al. (2008). The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9, 75. doi: 10.1186/1471-2164-9-75.

Babic, A., Guerout, A.M., and Mazel, D. (2008). Construction of an improved RP4 (RK2)-based conjugative system. Res Microbiol 159(7-8), 545-549. doi: 10.1016/j.resmic.2008.06.004.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19(5), 455-477. doi: 10.1089/cmb.2012.0021.

Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc.Natl.Acad.Sci.U.S.A 97(12), 6640-6645.

Faber, F., Thiennimitr, P., Spiga, L., Byndloss, M.X., Litvak, Y., Lawhon, S., et al. (2017). Respiration of Microbiota-Derived 1,2-propanediol Drives Salmonella Expansion during Colitis. PLoS Pathog 13(1), e1006129. doi: 10.1371/journal.ppat.1006129.

Forbes, J.D., Van Domselaar, G., and Bernstein, C.N. (2016). The Gut Microbiota in Immune-Mediated Inflammatory Diseases. Front Microbiol 7, 1081. doi: 10.3389/fmicb.2016.01081.

Garsin, D.A. (2010). Ethanolamine utilization in bacterial pathogens: roles and regulation. Nat Rev Microbiol 8(4), 290-295. doi: 10.1038/nrmicro2334.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 6(5), 343-345. doi: 10.1038/nmeth.1318.

Gordienko, E.N., Kazanov, M.D., and Gelfand, M.S. (2013). Evolution of pan-genomes of Escherichia coli, Shigella spp., and Salmonella enterica. J Bacteriol 195(12), 2786-2792. doi: 10.1128/JB.02285-12.

Hapfelmeier, S., Ehrbar, K., Stecher, B., Barthel, M., Kremer, M., and Hardt, W.D. (2004). Role of the Salmonella pathogenicity island 1 effector proteins SipA, SopB, SopE, and SopE2 in Salmonella enterica subspecies 1 serovar Typhimurium colitis in streptomycin-pretreated mice. Infect Immun 72(2), 795-809.

Karberg, K.A., Olsen, G.J., and Davis, J.J. (2011). Similarity of genes horizontally acquired by Escherichia coli and Salmonella enterica is evidence of a supraspecies pangenome. Proc Natl Acad Sci U S A 108(50), 20154-20159. doi: 10.1073/pnas.1109451108.

Kaser, A., Zeissig, S., and Blumberg, R.S. (2010). Inflammatory bowel disease. Annual review of immunology 28, 573-621. doi: 10.1146/annurev-immunol-030409-101225.

Kaval, K.G., and Garsin, D.A. (2018). Ethanolamine Utilization in Bacteria. MBio 9(1). doi: 10.1128/mBio.00066-18.

Kaval, K.G., Singh, K.V., Cruz, M.R., DebRoy, S., Winkler, W.C., Murray, B.E., et al. (2018). Loss of Ethanolamine Utilization in Enterococcus faecalis Increases Gastrointestinal Tract Colonization. MBio 9(3). doi: 10.1128/mBio.00790-18.

Lange, A., Beier, S., Steimle, A., Autenrieth, I.B., Huson, D.H., and Frick, J.S. (2016). Extensive Mobilome-Driven Genome Diversification in Mouse Gut-Associated Bacteroides vulgatus mpk. Genome Biol Evol 8(4), 1197-1207. doi: 10.1093/gbe/evw070.

Naftali, T., Reshef, L., Kovacs, A., Porat, R., Amir, I., Konikoff, F.M., et al. (2016). Distinct Microbiotas are Associated with Ileum-Restricted and Colon-Involving Crohn's Disease. Inflamm Bowel Dis 22(2), 293-302. doi: 10.1097/MIB.0000000000000662.

Porwollik, S., Wong, R.M., and McClelland, M. (2002). Evolutionary genomics of Salmonella: gene acquisitions revealed by microarray analysis. Proc Natl Acad Sci U S A 99(13), 8956-8961. doi: 10.1073/pnas.122153699.

Rakitina, D.V., Manolov, A.I., Kanygina, A.V., Garushyants, S.K., Baikova, J.P., Alexeev, D.G., et al. (2017). Genome analysis of E. coli isolated from Crohn's disease patients. BMC Genomics 18(1), 544. doi: 10.1186/s12864-017-3917-x.

Sassone-Corsi, M., Nuccio, S.P., Liu, H., Hernandez, D., Vu, C.T., Takahashi, A.A., et al. (2016). Microcins mediate competition among Enterobacteriaceae in the inflamed gut. Nature 540(7632), 280-283. doi: 10.1038/nature20557.

Spees, A.M., Lopez, C.A., Kingsbury, D.D., Winter, S.E., and Baumler, A.J. (2013a). Colonization resistance: battle of the bugs or Menage a Trois with the host? PLoS Pathog 9(11), e1003730. doi: 10.1371/journal.ppat.1003730.

Spees, A.M., Wangdi, T., Lopez, C.A., Kingsbury, D.D., Xavier, M.N., Winter, S.E., et al. (2013b). Streptomycin-induced inflammation enhances Escherichia coli gut colonization through nitrate respiration. MBio 4(4). doi: 10.1128/mBio.00430-13.

Spor, A., Koren, O., and Ley, R. (2011). Unravelling the effects of the environment and host genotype on the gut microbiome. Nat Rev Microbiol 9(4), 279-290. doi: 10.1038/nrmicro2540.

Staib, L., and Fuchs, T.M. (2015). Regulation of fucose and 1,2-propanediol utilization by Salmonella enterica serovar Typhimurium. Front Microbiol 6, 1116. doi: 10.3389/fmicb.2015.01116.

Thiennimitr, P., Winter, S.E., Winter, M.G., Xavier, M.N., Tolstikov, V., Huseby, D.L., et al. (2011). Intestinal inflammation allows Salmonella to use ethanolamine to compete with the microbiota. Proc Natl Acad Sci U S A 108(42), 17480-17485. doi: 10.1073/pnas.1107857108.

Unden, G., and Bongaerts, J. (1997). Alternative respiratory pathways of Escherichia coli: energetics and transcriptional regulation in response to electron acceptors. Biochim Biophys Acta 1320(3), 217-234.

Vejborg, R.M., Hancock, V., Petersen, A.M., Krogfelt, K.A., and Klemm, P. (2011). Comparative genomics of Escherichia coli isolated from patients with inflammatory bowel disease. BMC Genomics 12, 316. doi: 10.1186/1471-2164-12-316.

Waidmann, M., Bechtold, O., Frick, J.S., Lehr, H.A., Schubert, S., Dobrindt, U., et al. (2003). Bacteroides vulgatus protects against Escherichia coli-induced colitis in gnotobiotic interleukin-2-deficient mice. Gastroenterology 125(1), 162-177.

Weirich, J., Brautigam, C., Muhlenkamp, M., Franz-Wachtel, M., Macek, B., Meuskens, I., et al. (2017). Identifying components required for OMP biogenesis as novel targets for antiinfective drugs. Virulence 8(7), 1170-1188. doi: 10.1080/21505594.2016.1278333.

Winter, S.E., and Baumler, A.J. (2014). Why related bacterial species bloom simultaneously in the gut: principles underlying the 'Like will to like' concept. Cell Microbiol 16(2), 179-184. doi: 10.1111/cmi.12245.

Figures and tables

Table 1: Bacterial strains used in this study

Strains and genotypes	Reference
E. coli mpk2	Mouse fecal isolate
E. coli mpk2 pHP45Ω-cm	This study
E. coli mpk2 pHP45Ω-km	This study
E. coli mpk2 Δ*eutR* pHP45Ω-cm	This study
E. coli mpk2 Δ*pduD* pHP45Ω-cm	This study
E. coli mpk2 Δ*eutR,* Δ*pduD* pHP45Ω-cm	This study
E. coli mpk2 Δ*fucO*	This study
E. coli mpk2 Δ*fucO* Δ*rhaBAD* pHP45Ω-cm	This study
Bacteroides vulgatus mpk	Mouse fecal isolate
Bacteroides thetaiotaomicron VPI-5482	ATCC 29148
E. coli 118 λ pir^+	(Weirich et al 2017)
E. coli β2168 Δ*nic35*	(Babic et al. 2008)

Table 2: List of plasmids used in this study

Plasmid	Characteristics	Reference
pKD3	Carries chloramphenicol resistance inserted in *eutR*	(Daksenko and Wanner 2000)
pKD46	Carries λ Red recombinase under arabinose inducible promoter	(Daksenko and Wanner 2000)
pCP20	FLP-recombination plasmid to delete resistance cassette from *E. coli* mpk2 Δ*eutR*	(Daksenko and Wanner 2000)
pHP45Ω-cm	Low-copy plasmid to mark strains for in vivo experiments with chloramphenicol resistance	ATCC 37654
pHP45Ω-km	Low-copy plasmid to mark strains for in vivo experiments with kanamycin resistance	ATCC 37653
pSB890	Suicide plasmid	(Hapfelmeier et al. 2004)
pSB890-ΔpduD	Suicide plasmid carrying flanking regions of *pduD* knockout	This study
pSB890-ΔfucO	Suicide plasmid carrying flanking regions of *fucO* knockout	This study
		This study

Table 3: Primers used in this study

Primer	Sequence in 5´-3´	
eutR deletion		
eutR for_EcoRI	tttgaattcGGCTTCGATTTAACCCCTTT	This study
eutR rev_BamHI	aaaggatccGTTCGTCTGCTCAGTCACGA	This study
peutR inv_rev_BglII	ataagatctGAATGGGGGTGAGGGAGTTAC	This study
peutR inv_for_BglII	ataagatctGGCTGTACGGGTCTTTTTCA	This study
cat for_BglII	ataagatctGTGTAGGCTGGAGCTGCTTC	This study
cat rev_BglII	ataagatctATGGGAATTAGCCATGGTCC	This study
eutR_seq_f	GGCTTCGATTTAACCCCTTT	This study
eutR_seq_r	GTTCGTCTGCTCAGTCACGA	This study
pduD deletion		
pduD_a_f	GTTATTCCGCGAAATATAATGACCCTCTTGCTCG CGTCATCCTATGCTTCT	This study
pduD_d_r	CCGCGGTGGGGCTTTTTATTGAGCTTGTGCAGAA TAACGCCAGTCAG	This study
pduD_b_r	CATTTCTGCCAGCACGTCTTC	This study
pduD_bc_f	TATCGAAGACGTGCTGGCAGAAATGCGCGTGAC GTTTTAACAAAGG	This study
pduD_seq_f	ACCGGCGATTATCTCCATAC	This study
pduD_seq_r	TTATTTGCCAGCGGGTAATC	This study
gib_uni_890_r2	CAAGAGGGTCATTATATTTCGCG	(Weirich et al 2017)
gib_uni_890_f2	CAAGCTCAATAAAAAGCCCCAC	(Weirich et al 2017)
p890_seq_f	CGTCACCAAATGATGTTATTCC	(Weirich et al 2017)
p890_seq_r	GTTGAGAAGCGGTGTAAGTG	(Weirich et al 2017)

Figures

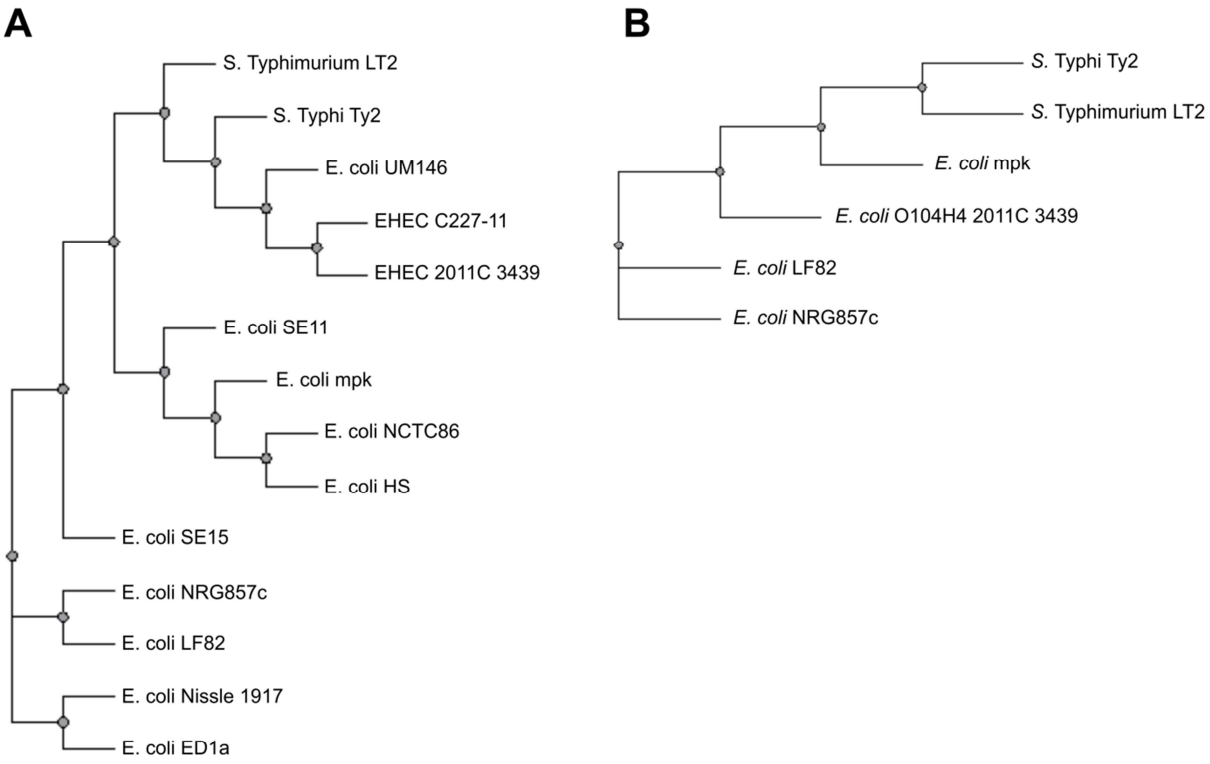


Figure 1: Phylogeny of ethanolamine and 1,2-propaendiol utilization gene clusters. Phylogenetic relations between different ethanolamine utilization (A) and 1,2-propanediol degradation (B) gene clusters of commensal (IBD- and non-IBD isolates) and pathogenic *E. coli* strains, and *Salmonella* Typhimurium and *S.* Typhi, displayed by a cladogram.

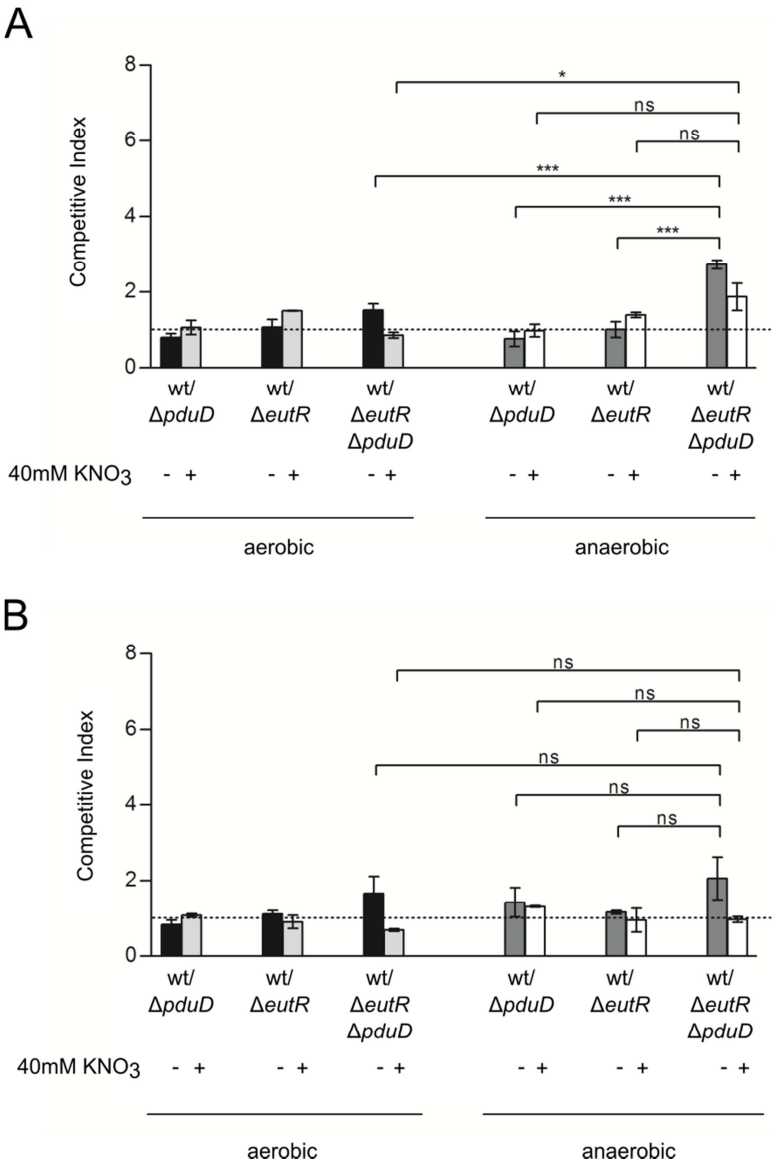


Figure 2: Anaerobic conditions fuel competitiveness of *E. coli* mpk. Anaerobic competition assay of *E. coli* mpk wild type and the *pduD* mutant, *eutR* and *eutR* and *pduD* double mutant in minimal medium supplemented with 15 mg ileal (A) or caecal and colonic mucus (B) and 40 mM KNO_3 when indicated. Bars represent geometric means and SEM of competitive index recovered after 24 h of wild type and mutant co-culture.

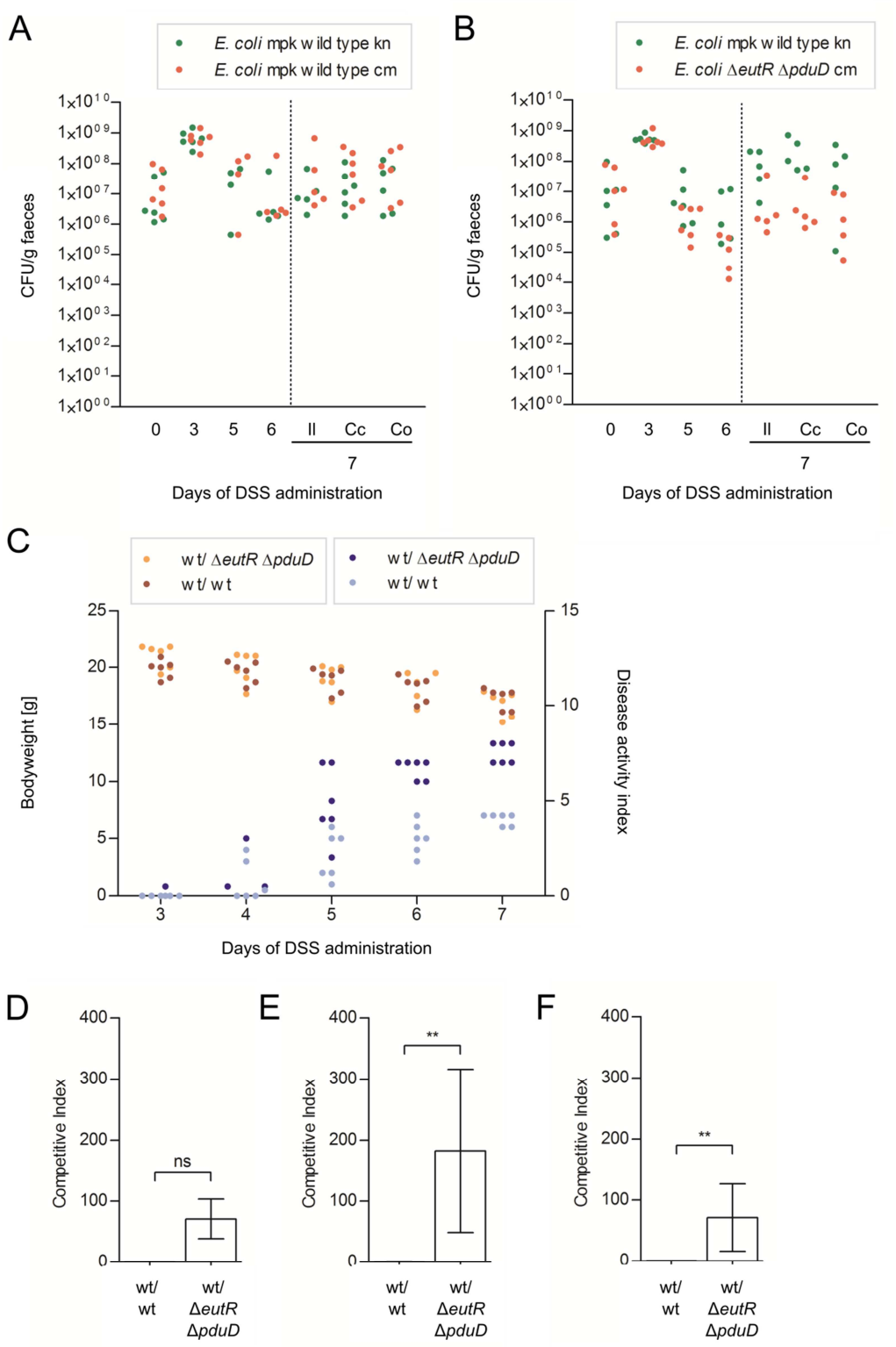
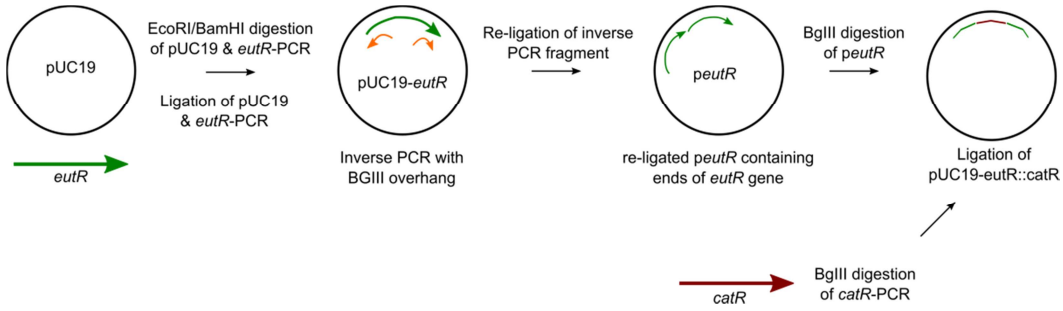


Figure 3: Inflammation boosts bacterial competition during 7 days of acute intestinal inflammation. C57BL/6N wild-type SPF mice were administered with $2x10^8$ CFU of a 1:1 mixture of *E. coli* mpk wild type and Δ*eutR* Δ*pduD* orally in the drinking water containing 3 % DSS. Bacterial colonization was monitored for 7 days under inflammatory conditions (A and B). Weight loss and disease activity index (DAI) was determined between days 3-7 to evaluate disease progression (C). At the end of the experiment the CFU were determined from ileal, caecal and colonic content and the competitive index was calculated for ileal (D), caecal (E), and colonic contents (F).

Supplementary information

Cloning strategy for LambdaRed-mediated knockout of *eutR*



The *eutR* gene was amplified from *E. coli* mpk DNA creating EcoRI and BamHI overhangs. pUC19 and the *eutR* PCR fragment were digested using EcoRI and BamHI. After gel purification the digested products were ligated and transformed into chemically competent *E. coli* DH5α cells. The resulting plasmid was purified and an inverse PCR was performed to create a pUC19 plasmid with *eutR* homology regions and BGIII overhangs. The inverse PCR fragment was re-ligated and transformed into *E. coli* DH5α cells. The purified plasmid peutR was BgIII-digested and purified from an agarose gel. The chloramphenicol resistance gene *catR* was amplified from pKD3 creating BGIII overhangs. The PCR product was BgIII-digested and purified from a gel. The BgIII-digested plasmid and the BgIII-digested *catR* fragment were ligated and transformed into *E. coli* DH5α cells. The resulting pUC19-*eutR::catR* plasmid was used as template to generate the *eutR::catR* PCR fragment used for LambdaRed-mediated recombination to create E. coli mpk Δ*eutR::catR*. The *eutR::catR* PCR fragment was transformed into *E. coli* mpk containing pKD46. The resulting clones were checked with *eutR* seq primers. The *catR* gene was flipped out using the pCP20 plasmid as reported elsewhere (Datsenko and Wanner, 2000).

Supplementary Figure

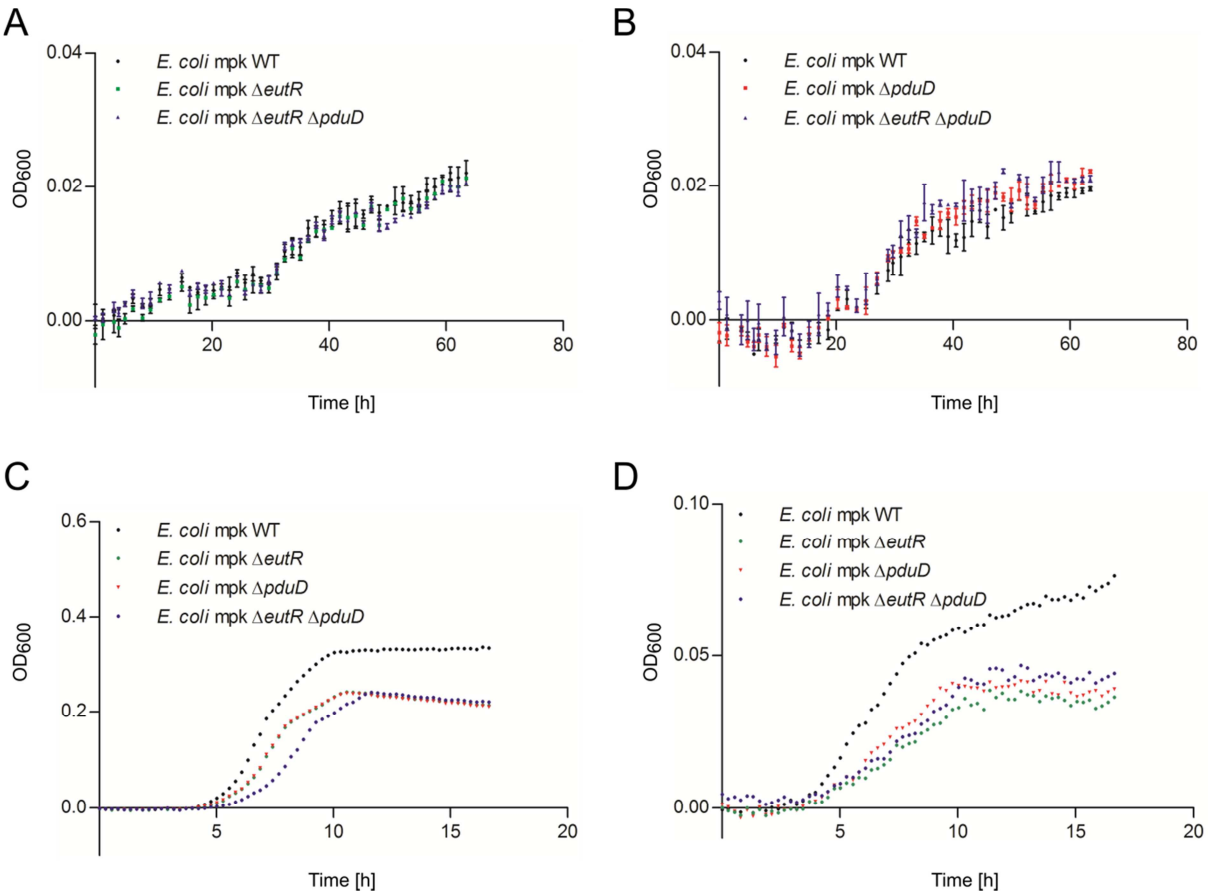


Figure S1: Aerobic growth curves in minimal medium containing different substrates. Aerobic growth assay of E. coli mpk wild type and the pduD knockout, eutR knockout and E. coli ΔeutR and ΔpduD mutant in minimal medium supplemented with 0.5% ethanolamine (A) or 0.5% 1,2-propanediol (B), 15 mg ileal (C) or caecal and colonic mucus (D). Points represent geometric means and SEM of OD600.

Accepted publication d

Lange A., Schäfer A., Bender A., Steimle A., Parusel R, Frick J.S.: *Galleria mellonella*: a novel invertebrate model to distinguish intestinal symbionts from pathobionts. Frontiers in Immunology 9 (2114) (2018)

frontiers in Immunology

ORIGINAL RESEARCH
published: 19 September 2018
doi: 10.3389/fimmu.2018.02114

Galleria mellonella: A Novel Invertebrate Model to Distinguish Intestinal Symbionts From Pathobionts

*Anna Lange[1], Andrea Schäfer[1], Annika Bender[1], Alexander Steimle[1], Sina Beier[2], Raphael Parusel[1] and Julia-Stefanie Frick[1]**

[1] Department for Medical Microbiology and Hygiene, Interfacultary Institute for Microbiology and Infection Medicine, University of Tübingen, Tübingen, Germany, [2] Algorithms in Bioinformatics, ZBIT Center for Bioinformatics, University of Tübingen, Tübingen, Germany

OPEN ACCESS

Edited by:
Lluis Tort, Autonomous University of Barcelona, Spain

Reviewed by:
Sylvia Brugman, Wageningen University & Research, Netherlands
Annalisa Pinsino, Istituto di Biomedicina e di Immunologia Molecolare Alberto Monroy (IBIM), Italy

***Correspondence:**
Julia-Stefanie Frick
julia-stefanie.frick@med.uni-tuebingen.de

Specialty section:
This article was submitted to Comparative Immunology, a section of the journal Frontiers in Immunology

Received: *12 June 2018*
Accepted: *28 August 2018*
Published: *19 September 2018*

Citation:
Lange A, Schäfer A, Bender A, Steimle A, Beier S, Parusel R and Frick J-S (2018) Galleria mellonella: A Novel Invertebrate Model to Distinguish Intestinal Symbionts From Pathobionts. Front. Immunol. 9:2114. doi: 10.3389/fimmu.2018.02114

Insects and mammals share evolutionary conserved innate immune responses to maintain intestinal homeostasis. We investigated whether the larvae of the greater wax moth *Galleria mellonella* may be used as an experimental organism to distinguish between symbiotic *Bacteroides vulgatus* and pathobiotic *Escherichia coli*, which are mammalian intestinal commensals. Oral application of the symbiont or pathobiont to *G. mellonella* resulted in clearly distinguishable innate immune responses that could be verified by analyzing similar innate immune components in mice *in vivo* and *in vitro*. The differential innate immune responses were initiated by the recognition of bacterial components *via* pattern recognition receptors. The pathobiont detection resulted in increased expression of reactive oxygen and nitrogen species related genes as well as antimicrobial peptide gene expression. In contrast, the treatment/application with symbiotic bacteria led to weakened immune responses in both mammalian and insect models. As symbionts and pathobionts play a crucial role in development of inflammatory bowel diseases, we hence suggest *G. mellonella* as a future replacement organism in inflammatory bowel disease research.

Keywords: intestinal commensals, insect, innate immunity, symbiont, pathobiont, reactive oxygen species, antimicrobial peptides

INTRODUCTION

The intestinal microbiota plays a pivotal role in the development of intestinal inflammation and maintenance of homeostasis in humans and the respective mouse models. In this context bacteria of the intestinal microbiota beneficial for the host named symbionts, while commensals promoting intestinal inflammation in a genetically predisposed organism are named pathobionts (1). This is the case in inflammatory bowel disease (IBD), where certain commensals like *Escherichia coli* or *Klebsiella* genus, can accumulate and lead to disturbed immune responses (2, 3). In insects resident bacteria are also not always symbionts. It was demonstrated that *Gluconobacter morbifer* is pathobiotic in *Drosophila* and becomes "pathogenic" when the commensal microbiota composition is disturbed leading to deregulated intestinal immunity (4). Hence, insects might be suitable replacement organisms for discriminating intestinal mammalian symbionts from pathobionts.

Intestinal epithelial cells (IECs) mediate intimate contact between host and microbiota. IECs recognize bacterial components, so called microbe-associated molecular patterns (MAMPs), by pattern recognition receptors (PRRs) e.g., Toll-like receptors (TLRs) resulting in innate immune responses leading to the maintenance of homeostatic balance (5). Consequently, IECs secrete different molecules with antibacterial activity e.g., reactive oxygen species (ROS) (6), reactive nitrogen species (RNS) (7), antimicrobial peptides, (AMPs) or CCL20, an anti-inflammatory cytokine (5).

Recognition and effector components of the gut immune responses are evolutionary conserved (5).

The first line of intestinal innate immune defense is built up by the secretion of antimicrobial ROS and RNS, and is considered to be the most primordial innate immune response mechanism (8). Intestinal ROS can be produced primarily by NADPH oxidases (NOX), and under certain circumstances also by nitrate oxide synthase 2 (NOS2) (6); NOS2 mainly produces RNS (7). ROS and RNS are generated by phagocytes and respiratory and intestinal epithelial cells during infection to eliminate invading microbes (4, 7). However, an *in vivo* role of ROS relevant for host antimicrobial response in mammals needs to be elaborated, as most data involving ROS activity was obtained from *in vitro* experiments (4). ROS/RNS signaling requires detoxification, otherwise the substrates can be harmful for host cells and may lead to DNA, RNA or protein damage, or apoptosis (9). Antioxidant molecules like catalase, superoxide dismutase or glutathione S-transferase (GST) support detoxification (6). In *Drosophila* it has been shown that some bacteria can resist ROS and activate antimicrobial peptide production. Such AMP responses might complement ROS signaling under certain circumstances (4). Thus, gut innate immunity is further enhanced by AMPs, which have a critical role in reinforcing the intestinal barrier function and to prevent the uptake of bacteria (5). AMP expression is tightly regulated by the recognition of present microbes and AMPs are produced following PRR activation by IECs. However, the intestinal epithelium seems to mostly tolerate the intrinsic microbiota, and maintain a basal level of AMPs (5). In mammalian intestines AMPs are produced broadly in epithelial cells and in specialized cells in a constitutive or inducible manner. Among other AMP groups, defensins are central intestinal AMPs (10). Insects express AMPs as well and are often reported to be a very potent component of innate defense responses since insects lack adaptive immune responses (10, 11). Insects produce many AMPs simultaneously to potentiate their antimicrobial effect (11). The above mentioned strategies do not only contribute to fight invading pathogens but also help to balance the commensal microbiota composition and support maintenance of intestinal homeostasis (12).

Here we show that the larva of the greater wax moth *Galleria mellonella* is a suitable organism to distinguish between commensal symbiotic and pathobiotic intestinal microbiota members. In our study we chose an oral force-feeding model to mimic the natural route of colonizing a hosts' intestine with commensals. The *G. mellonella* larvae were administered with symbiotic *Bacteroides vulgatus* mpk, mediating an immune-balancing effect in mouse models of IBD, hence preventing from intestinal inflammation or the pathobiotic *Escherichia coli* mpk, which ignites inflammation (13, 14). We observed different hallmarks of *G. mellonella* innate immune responses like bacterial recognition, induction of ROS and AMP production. Furthermore, we analyzed similar components in mammalian innate immune responses *in vivo* and *in vitro* which are involved in responses toward pathogen infection or inflammatory responses. We demonstrated that *G. mellonella* can differentiate between symbiont and pathobiont homologous to mammalian systems based on the above mentioned innate responses (15).

MATERIALS AND METHODS

Galleria mellonella Rearing

Galleria mellonella breeding was performed by transferring eggs laid by adult moths into wax moth substrate (22% corn grits, 22% wheatmeal, 17.5% beeswax, 11% skimmed milk powder, 11% honey, 11% glycerol, 5.5% dried yeast). Adult moths and larvae were kept in the dark at 30°C. Last instar larvae were used for experiments. Only pale and fast moving larvae were selected for experiments, to exclude interference with any background stress and immunity reactions during assay performance. Further, the larvae were weighed to ascertain than all larvae are in the similar last instar growth stage (180–200 mg).

Bacterial Strains and Conditions

Bacteroides vulgatus mpk (13, 16) was grown at 37°C under anaerobic conditions. *B. vulgatus* was cultivated in liver broth for 2 days and subcultivated overnight in Brain Heart Infusion (BHI) medium. *Escherichia coli* mpk was grown in Luria-Bertani (LB) medium at 37°C under aerobic conditions. *E. coli* was grown overnight in LB broth and subcultivated for 2 h at 37°C on the next day prior to either force-feeding or stimulation of cell culture.

Homology Analysis

For homology assessment the BLASTX tool was used. *G. mellonella* sequences **(Supplementary Table S1)** were obtained from a published transcriptome dataset (17) and aligned to *Mus musculus* and *Homo sapiens* genomes.

As some of the relevant transcripts were only partial sequences, we did not align the fragments to distinct mouse or human protein sequences.

The decision whether proteins were homologous was based on commonly discussed criteria (18, 19). Proteins were considered homologous with 30% identity when completely covered and an *E*-value of 10^{-6}. Further we considered proteins homologous if at least 70% of the sequence was covered with an *E*-value of above 10^{-10} and with at least 25% identity.

Protein Alignments

Protein sequences were aligned using MAFFT (20) and visualized using Unipro UGENE. Protein sequences were obtained from NCBI database: different *Lepidoptera* insect gallerimycin sequences [*Galleria mellonella* (AAM46728.1), *Spodoptera litura* (AEE37278.1), *Spodoptera exigua* (ADJ95798.1), *Spodoptera frugiperda* (AAP69838.1), *Spodoptera exigua* (AKJ54495.1),

Helicoverpa armigera (ADR51151.1, partial), *Trichoplusia ni* (ABV68855.1), *Lonomia obliqua* (AAV91471.1), *Samia ricini* (BAG12297.1)], and mouse β-defensin 2 (BD-2) (P82020.1), human BD-2 (NP_004933.1), mouse (NP_058656.1), and human CCL20 (NP_004582.1) were used for the alignment. The consensus sequence was generated with the strict consensus type with a threshold of 90%. The agreements in hydrophobicity were highlighted.

The phylogenetic tree was created using Unipro UGENE and the PHYLIP Neighbor Joining tree building method with bootstrapping and majority rule (extended) consensus type.

Cell Culture

Transimmortalized mouse intestinal epithelial cells (IECs) (H2b)–mICcl2 cells–were originally derived from the small intestine of a transgenic mouse (21). Cells were grown in advanced Dulbecco's Modified Eagle Medium/Nutrient Mixture (DMEM)/F12 medium (Thermo Fisher) containing insulin-transferrin-selenium (ITS-G) (Thermo Fisher), 10% fetal calf serum (FCS) (Thermo Fisher), 5% Glutamine (Thermo Fisher), 10 nM epidermal growth factor (EGF) (Sigma), 1 nM triiodothyronine (Sigma), 50 nM dexamethasone (Sigma) at 37°C in 5% CO_2/air atmosphere. The medium was used within 2 weeks of preparation to maintain the activity of growth factors.

Human SW480 cells were maintained in L-15 (Leibovitz) medium (Sigma) containing 2 mM Glutamine (Thermo Fisher) and 10% FCS (Thermo Fisher) at 37°C in normal air atmosphere.

Bone marrow-derived dendritic cells (BMDCs) were isolated from wild type C57BL/6J mice seeded and cultivated in petri dishes as described previously (22). Seven days after isolation cells were used for further *in vitro* stimulation experiments.

Stimulation of mICcl2 Cells, SW480 and BMDCs

mICcl2 cells and SW480 cells were seeded in Nunclon cell culture dishes for 24 h (Thermo Fisher). The next day mICcl2 and SW480 cells were stimulated with the bacteria MOI (multiplicity of infection) 10 for 0, 4, and 24 h whereas BMDCs were stimulated for 0.5, 2, and 4 h. PBS was used as mock stimulation control. Cells were harvested by scraping and centrifugation. Supernatants were used for ELISA and the remaining cell pellets were used for RNA extraction and analysis.

Bacterial Association of *Rag1*$^{-/-}$ Mice

Six–Eight weeks old germ-free *Rag1*$^{-/-}$ mice were orally administered with bacteria (*B. vulgatus* or *E. coli*) *via* drinking water. The bacteria were cultivated as described above and adjusted to 10^8 bacteria per ml of sterile drinking water. The drinking water was replaced by sterile drinking water after 24 h. After 2 weeks of stable colonization, mice were transplanted intraperitoneally with 5×10^5 naïve T cells isolated from wild type SPF C57BL/6J mice. After 4 weeks, mice were sacrificed for analysis. Cell scrapings were obtained from the distal ileum and the colon to extract RNA.

Force-Feeding of Larvae With Bacterial Suspensions

The larvae were orally administered with 10^7 bacteria. In order to force-feed the larvae the bacterial suspension was applied to their mouth using a syringe by inserting a blunt-ended needle between their mandibles. The syringe was placed into a microsyringe pump (World Precision Instruments) to ensure the accuracy of the suspension volume applied to each larva. The viable force-fed larvae were incubated in the dark at 37°C between 1 and 6 h. The larvae were snap-frozen after force-feeding. PBS-administered larvae were used as mock control to exclude the influence of stress responses potentially induced by the handling of the larvae.

RNA Extraction

G. mellonella larvae were incubated after force-feeding for 0–6 h. The individual larva was snap-frozen in liquid nitrogen and homogenized. Frozen homogenates were incubated in TRIzol (Sigma-Aldrich) for 1 h. After centrifugation the supernatant was mixed with 1-Bromo-3-Chloropropane (Sigma-Aldrich), incubated for 5 min at room temperature and incubated 10 min on ice. After centrifugation the upper layer was transferred into a new tube and RNA was precipitated with isopropanol. RNA pellets were washed with 70% ethanol. Dried RNA was resuspended with nuclease-free water containing RNasin Ribonuclease Inhibitor (Promega). The RNA contained the larval RNA as well as the bacterial RNA of the respective strain used for oral administration.

Mouse (mICcl2 cells, BMDCs, ileal, and colonic tissue) RNA was isolated from the cell pellets after centrifugation of the stimulated cell lines. For the isolation the RNeasy kit (QIAGEN) was used according to manufacturer's instructions.

RT-PCR

5 μg of the isolated RNA was DNase-digested using DNA-free DNA Removal Kit (Thermo Fisher Scientific). Quantitative RT-PCR was performed using the QuantiFast SYBR Green RT-PCR Kit (Qiagen). Primers for quantitative RT-PCR were designed using NCBI Primer-BLAST design tool **(Supplementary Table S2)**. Primers were checked for gene-specificity, i.e., fragment length, using PCR and agarose gel electrophoresis. In addition the primer efficiency was assessed; suitable primers had an efficiency of $E = 2$. In order to determine RNA gene expression, the ΔΔct method was used for calculations. Ubiquitin was used as housekeeping gene for *G. mellonella* RNA measurements, whereas actin was used as housekeeping gene in mouse RNA. Bacteria stimulated samples were normalized to mock controls.

Quantification of Bacterial Load in Force-Fed *G. mellonella* Larvae

Plasmid standards were generated by blunt-end cloning using pJET (Thermo Fisher Scientific) and the respective specific 16s PCR fragments of *E. coli* [Primer forward: GTTAATACCTTTGCTCATTGA, reverse: ACCAGGGTATCTAATCCTGTT (23)] or *B. vulgatus* (Primer forward: AACCTGCCGTCTACTCTT, reverse: CAACTGACTTAAACATCCAT (24)]. The concentration

of the isolated plasmids was determined and the standard concentrations were prepared in 10-fold serial dilutions in a range of 100,000–10 copies. Groups of force-fed larvae were incubated for 1, 6, and 24 h. At the end of each timepoint the bacterial RNA was isolated as described above. cDNA was synthesized and concentration was measured using Qubit dsDNA High Sensitivity Assay (Thermo Fisher Scientific). For the qPCR measurement cDNA concentrations were adjusted to 5 ng per reaction and PCR was performed using QuantiFast SYBR Green PCR Kit (Qiagen). Bacterial copy numbers were determined by a standard curve for which $\log_{10}$ of standard copy numbers was plotted against ct-values.

ELISA

The supernatants were collected by centrifugation after stimulation of the cultured cells. Human and mouse CCL20 (R&D Systems), mouse CD14 (R&D Systems), and human BD-2 (Arigo Biolaboratories) ELISAs were performed according to manufacturer's instructions. Absorbance was measured in absorbance reader (Tecan).

Statistical Analysis

Statistical analysis of the data was performed with GraphPad Prism 5 software. The distribution of the RNA and ELISA datasets was tested using the Kolmogorov-Smirnov method. RNA expression data was analyzed using the unpaired *t*-test, when the dataset was normally distributed. When datasets did not follow normal distribution, a nonparametric *t*-test was performed. ELISA datasets were analyzed using One-way ANOVA followed by Tukey's multiple comparison test. Differences were considered to be statistically significant if $P < 0.05$ ($^{*}P \leq 0.05$; $^{**}P \leq 0.01$; $^{***}P \leq 0.001$).

RESULTS

Invertebrates and Vertebrates Share Evolutionary Conserved Components of Different Innate Immune Responses

First we analyzed potential evolutionary relations between *Galleria mellonella*, mouse and human marker genes of oxidative stress responses (nitric oxide synthase (NOS), NADPH oxidase (NOX), and glutathione S-transferases (GST)), LPS recognition (apolipophorin, hemolin), and a defensin-like antimicrobial peptide (gallerimycin) on protein level. *G. mellonella* sequences were obtained from a published transcriptome dataset (17) and compared to *Mus musculus* and *Homo sapiens* genomes using BLASTX algorithm.

One transcript could be identifiie encoding for NOS in *G. mellonella*. Unfortunately, the full-length sequence of *G. mellonella* NOS protein was unavailable, but still the query coverage was approximatly 90% and identity greater than 40%. The transcript reaches *E*-values of 1×10^{-6} when aligned with BLASTX to the mouse genome, and *E*-values between 1×10^{-7} when aligned to the human genome **(Table 1)**. We conclude that the sequences are homologous. Further we detected one transcript encoding for *G. mellonella* NOX. BLASTX alignment to the mouse and human genome revealed similarities and resulted in E-values below 10^{-6} **(Table 1)**. However, homology cannot be confirmed resolutely as the query coverage was only 64% with 34% identity for mammalian NOX proteins and 62% with 25% identity for mouse and human DUOX proteins.

The transcriptome analysis revealed 19 different available GST variants (17). We found matches with mammalian GST proteins with *E*-values below 10^{-6} **(Supplementary Table S3)** after BLASTX alignments, but query coverage sometimes even lower than 30%. Only five GST transcripts had query coverage of above 70%. Those proteins were found to have reasonable *E*-values above 10^{-10} and identity of more than 25%.

In the available transcriptome dataset we could further identify relevant LPS-binding molecules. Specifically, five different transcripts encoding for the LPS recognition molecule apolipophorin were detected as well as two transcripts of hemolin, another LPS-binding opsonin. However, none of those transcripts were shown to have similarity with relevant mouse or human proteins after BLASTX analysis.

For *G. mellonella* defensin-like antimicrobial peptide (AMP) gallerimycin, a full-length protein sequence was available in the NCBI database and observed possible similarities with mouse and human β-defensin-2 (BD-2) in more detail. We also integrated other insect gallerimycins available from NCBI in our analysis. Different sequences were chosen from the other *Lepidoptera* order in which *G. mellonella* also belongs. A protein alignment revealed similar regions of hydrophobic (red) and hydrophilic (blue) areas in all sequences independent from the organism **(Supplementary Figure S1).** This suggests that the similar regions account for similar functions such as capabilities to bind other interacting molecules. Further we built a phylogenetic tree with the insect gallerimycin protein sequences and the mouse and human BD-2 sequences **(Figure 1)**. The mouse and human BD-2 peptides showed the closest relation to each other. Interestingly, gallerimycin of the giant silkworm moth *Lonomia obliqua* and the ailanthus silkmoth *Samia ricini* were closer related to mBD-2 and hBD-2 compared to gallerimycin of the other *Lepidoptera* insects.

Additionally, we included the antimicrobial cytokine CCL20 to our analysis, because it shares structural homology with BD-2 (5). We aligned *G. mellonella* gallerimycin, mBD-2, hBD-2, mouse and human CCL20. Interestingly, as shown in the previous alignment, similar hydrophobic (red) and hydrophilic (blue) distribution of amino acids could be observed **(Supplementary Figure S2)**.

G. mellonella and Mice Show Increased Production of LPS-Binding Recognition Molecules in Response to Pathobiont

In previous work we showed that the pathobiont *E. coli* mpk induces colitis in $Il2^{-/-}$ mice, whereas the symbiont *B. vulgatus* mpk maintains homeostasis (13). Intestinal commensals can either be symbionts or pathobionts. This distinction becomes crucial in inflammatory bowel disease (IBD) research as IBD comes along with increased levels of pathobionts (2). *Bacteroides* species on the other hand are reported to exhibit different immunomodulatory effects (14, 25).

TABLE 1 | BLASTX analysis of nitric oxide synthase (*Nos*) and NADPH oxidase (*Nox*) genes.

G. mellonella	*Mus musculus*				*Homo sapiens*			
Transcript	**BLAST Hit**	**Identity**	**Query coverage**	***E* value**	**BLAST Hit**	**Identity**	**Query coverage**	***E* value**
Nitric oxide synthase	Nitric oxide synthase	42%	94%	1×10^{-06}	Nitric oxide synthase	40%	90%	1×10^{-07}
NADPH oxidase	NADPH oxidase 4	34%	64%	4×10^{-79}	NADPH oxidase 4	34%	64%	4×10^{-79}
	Duox2 protein	25%	62%	3×10^{-30}	Dual oxidase 1	24%	62%	7×10^{-26}

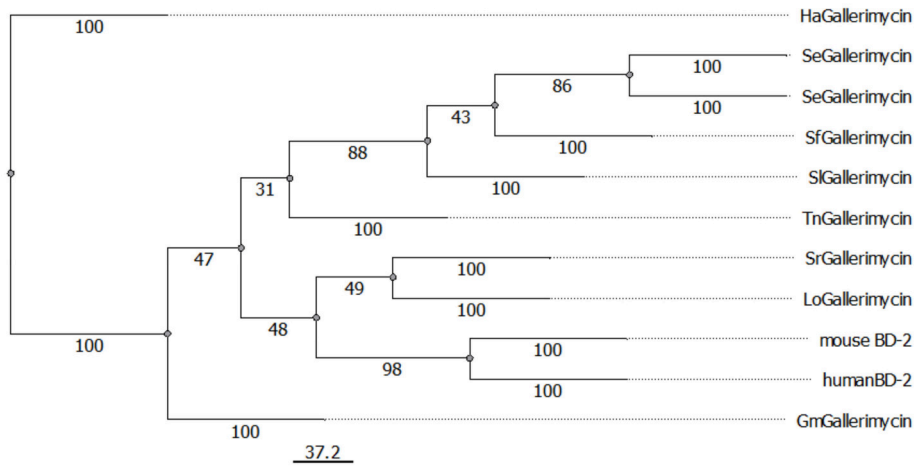


FIGURE 1 | Display of evolutionary relation between lepidopteran gallerimycins and mouse and human β-defensin 2 proteins**.** Phylogenetic relations among gallerimycin sequences of various lepidopteran insects (Gm: *Galleria mellonella*, Ha: *Helicoverpa armigera*, Lo: *Lonomia obliqua*, Sr: *Samia ricini*, Se: *Spodoptera exigua*, Sf: *Spodoptera frugiperda*, Sl: *Spodoptera litura*, Tn: *Trichoplusia ni*) and the mouse and human BD-2 sequences were displayed using PHYLIP Neighbor Joining method. The scale bar for branch length is shown below the tree. NCBI reference IDs for each sequence is available in the material and methods section.

We monitored if the symbiont *B. vulgatus* mpk or the pathobiont *E. coli* mpk were differentially recognized. Both bacteria are Gram-negative and contain Lipopolysaccharides (LPS). Thus, regulation of the LPS-binding molecules apolipophorin (*ApoIII*) and hemolin in *G. mellonella* and an analogous component of mammalian cells, CD14, which is an important LPS-binding molecule involved in microbe recognition (26), was observed.

E. coli administration of *G. mellonella* larvae lead to stronger gene expression of both *ApoIII* and hemolin after 2 h and remained elevated over time, compared to administeration with the symbiont *B. vulgatus* **(Figures 2A,B)**.

Likewise small intestinal mouse mICcl2 cells stimulated with *E. coli* provided higher *Cd14* gene expression after 4 h **(Figure 2C)**. The corresponding cell culture supernatants of the mICcl2 cells were assayed for secretion of CD14 protein by ELISA. CD14 levels were significantly enhanced in supernatants of *E. coli*-stimulated cells after 24 h compared to *B. vulgatus* stimulation **(Figure 2D)**. In compliance with this, we reported about increased intestinal *Cd14* expression levels in *E. coli*-associated gnotobiotic $Il2^{-/-}$ mice compared to *B. vulgatus*-associated gnotobiotic mice (13).

In addition, we stimulated BMDCs of C57BL/6 mice with either *E. coli* or *B. vulgatus* for 0.5, 2, and 4 h. Gene expression levels of *Cd14* were significantly higher in *E. coli*-stimulated cells compared to *B. vulgatus* stimulation similarly demonstrated in small intestinal cells **(Figure 2E)**.

Increased ROS-Related Gene Expression by Intestinal Commensals in Insect and Mammalian Hosts

Rapid environmental changes may force an organism to quickly adapt by triggering a cytoprotective survival program. This program can lead to increased production of reactive oxygen and nitrogen species (ROS/RNS) and can cause oxidative stress (27). In previous work, we already indicated a higher accumulation of ROS in *E. coli*-stimulated bone marrow-derived dendritic cells (BMDCs) compared to *B. vulgatus* stimulation (14).

In order to test whether different commensals cause different ROS-related gene expression in *G. mellonella* and gnotobiotic mice, *E. coli* or *B. vulgatus* were orally-administered to both organisms. Gene expression of *Nos* and *Nox* was determined from RNA isolated from either whole larval individuals or from ileal and colonic sections of bacteria-administered mono-associated $Rag1^{-/-}$ mice. Oral administration of *E. coli* to *G. mellonella* induced significantly higher *Nos* gene expression compared to *B. vulgatus* administration after 2 h and *Nox-4* expression after 3 h **(Figures 3A,B)**. After 5 h of bacterial

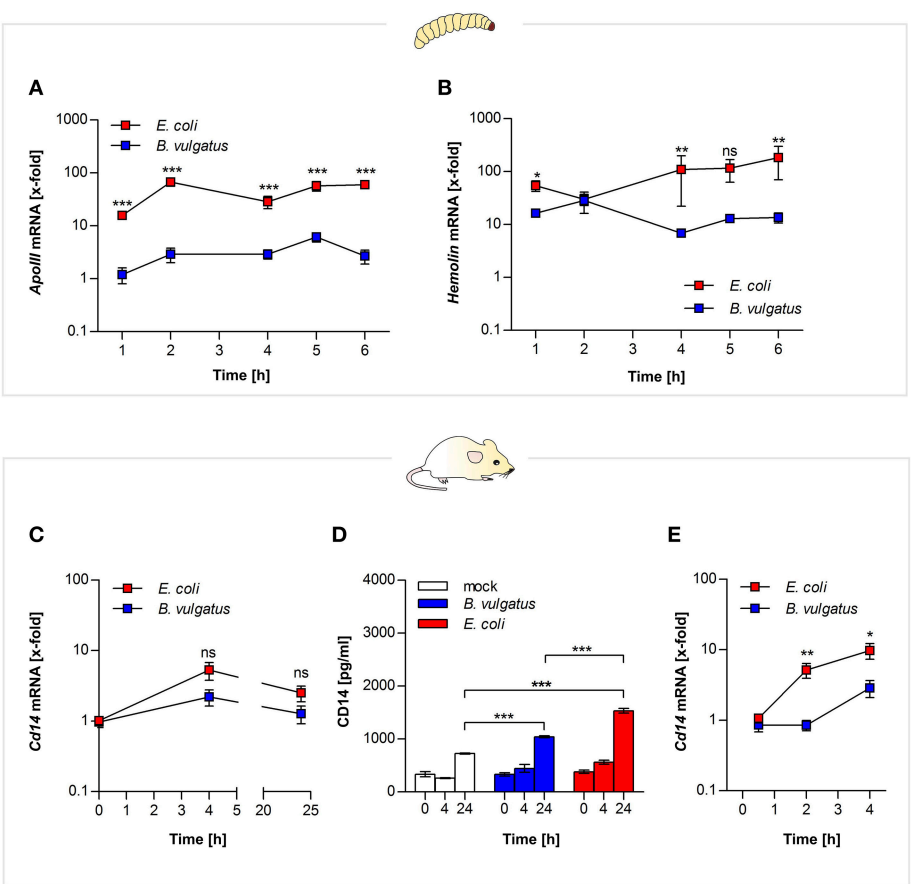


FIGURE 2 | MAMP recognition by both vertebrate and invertebrate pattern recognition receptors. Larvae were force-fed with bacteria and immune responses were observed over time and RNA was isolated from individuals. Gene expression of LPS recognition molecule apolipophorin **(A)** and opsonin hemolin **(B)** was measured from *G. mellonella* RNA ($n = 8$). Mouse mICcl2 cells were stimulated with *B. vulgatus* or *E. coli* for 0, 4, and 24 h. Gene expression of *Cd14* an LPS receptor molecule was determined in the isolated RNA (n=5) **(C)**. Secretion of CD14 in mICcl2 cell culture supernatants was analyzed by ELISA ($n = 3$) **(D)**. *Cd14* gene expression was further detected in supernatants of BMCDs of C57BL/6 mice stimulated with *B. vulgatus* or *E. coli* for 0.5, 2, and 4 h. ($n = 5$) **(E)**. Data represent the geometric means ± SEM.

administration, when initial *Nos* and *Nox-4* gene expression levels had decreased, the larvae induced *Gst* gene expression **(Figure 3C)**.

To investigate further if mouse intestinal cells display similar expression of the homologous ROS-responsive *Nos2* gene, we used small intestinal mICcl2 cells. The cells were stimulated with *E. coli* and *B. vulgatus* for 0, 4, and 24 h and *Nos2* expression was determined by quantitative PCR. *Nos2* Expression of *E. coli*-stimulated cells was significantly increased after 4 and 24 h **(Figure 3D)**. *Gstp1* gene expression on the other hand was not regulated at any time point **(Figure 3E)**.

The assessment of *in vivo* relevance of oxidative stress in response to symbiont or pathobiont association was performed in Rag1$^{-/-}$ gnotobiotic mice. After 4 weeks of mono-association with either *E. coli* or *B. vulgatus Nos2* gene expression was determined in ileal and colonic sections. *Nos2* gene expression in response to *E. coli*-association was more pronounced compared to *B. vulgatus* association **(Figures 3F,G)**, whereas *Gstp1* expression was not induced **(Figures 3H,I)**.

Pathobionts Activate Stronger Antimicrobial Peptide Responses

Recognition of microbes by sensing their microbe-associated molecular patterns (MAMPs) like LPS promotes a bacteria- or pathogen-dependent response program. Thus to fortify the intestinal barrier and to maintain intestinal homeostasis, epithelial cells express antimicrobial peptides (AMPs) as innate effectors molecules (5). Like vertebrates, insects are also able to secrete AMPs in the gut lumen and as insects lack an adaptive response it is a major component of host defense mechanism (28).

We analyzed AMP production in *G. mellonella* and *Rag1*$^{-/-}$ mice in response to oral administration of bacteria. In *G. mellonella* gene expression of defensin-like AMP gallerimycin was increased after 4 and 5 h in response to *E. coli* administration, whereas gene expression in *B. vulgatus*-administered larvae was significantly lower **(Figure 4A)**. Interestingly, we observed a substantially decreasing bacterial load in *G. mellonella* larvae within 6 h after administration **(Supplementary Figure S3)**.

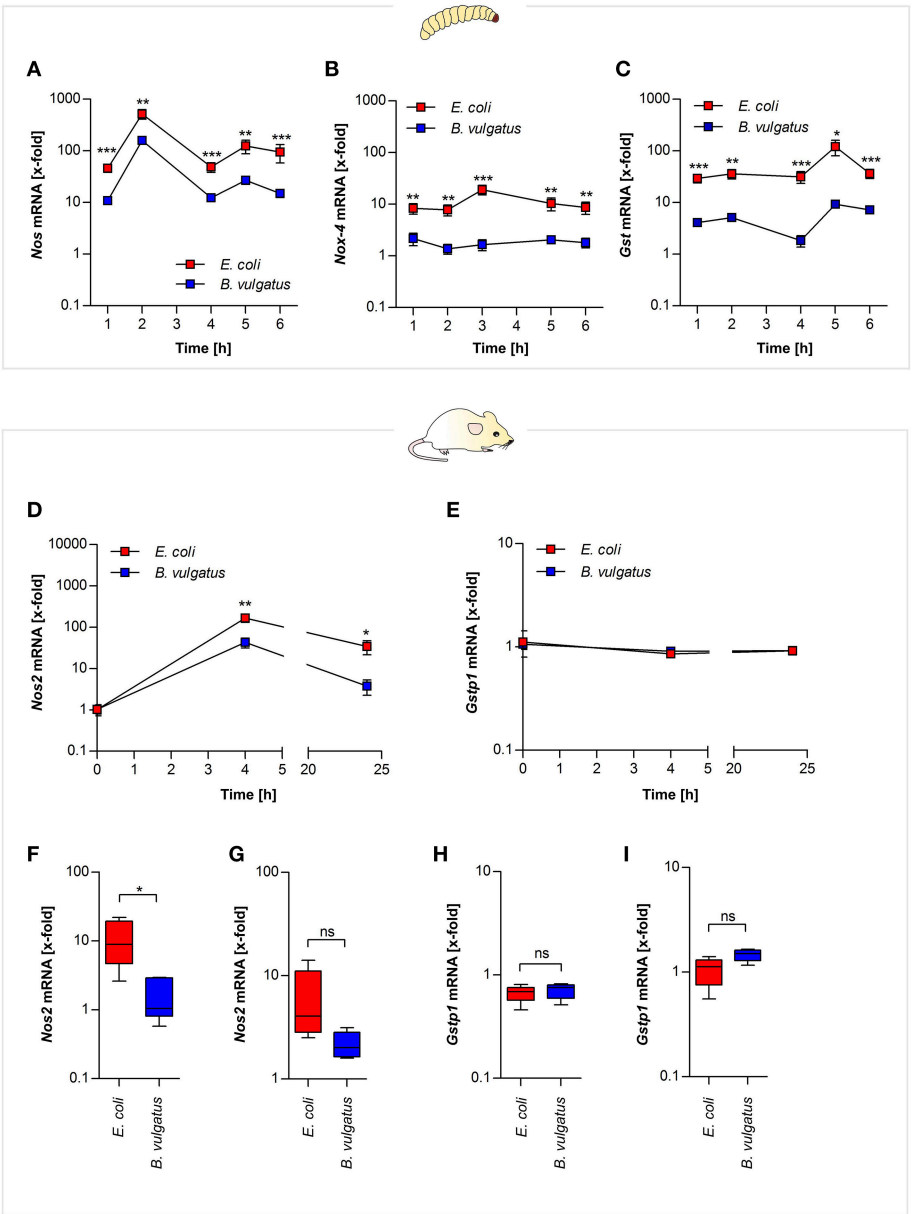


FIGURE 3 | ROS- and RNS-related gene expression after symbiont or pathobiont challenge**.** *E. coli* and *B. vulgatus* orally administered larvae were monitored over time. Gene expression of ROS defense response markers was investigated. *Nos* **(A)**, *Nox-4* **(B)**, and *Gst* **(C)** gene expression was measured in RNA isolated from whole individuals ($n = 8$). RNA was isolated from mouse mICcl2 cells stimulated with *B. vulgatus* or *E. coli* for 0, 4, and 24 h. *Nos2* **(D)** and *Gstp1* **(E)** gene expression was determined ($n = 5$). Gnotobiotic *Rag1*$^{-/-}$ mice were associated with *B. vulgatus* or *E. coli* for 4 weeks. *Nos2* gene expression was determined from RNA isolated from ileal **(F)** and colonic **(G)** sections. In addition, ileal **(H)** and colonic **(I)** *Gstp1* gene expression was measured ($n = 5$). Data represent the geometric means ± SEM.

Additionally, we could also detect increased gene expression of other AMPs in *G. mellonella*, such as gloverin, cecropin, moricin, and lysozyme in response to *E. coli* administration **(Supplementary Figures S4A–D)**. In summary, commensal-administered *G. mellonella* larvae produced a wide range of different antimicrobial peptides.

Similar responses could be observed in the mammalian host: in supernatants of *E. coli*-stimulated mouse small intestinal mICcl2 cells, higher secretion of CCL20 protein was detected compared to *B. vulgatus* stimulated cells **(Figure 4B)**. As mentioned above, gallerimycin, BD-2, and CCL20 share structural and functional homology.

Therefore we additionally investigated if human colonic epithelial SW480 cells produce AMPs in response to symbiont or pathobiont stimulation after 0, 4, and 24 h. We could demonstrate that in contrast to B. vulgatus, stimulation with

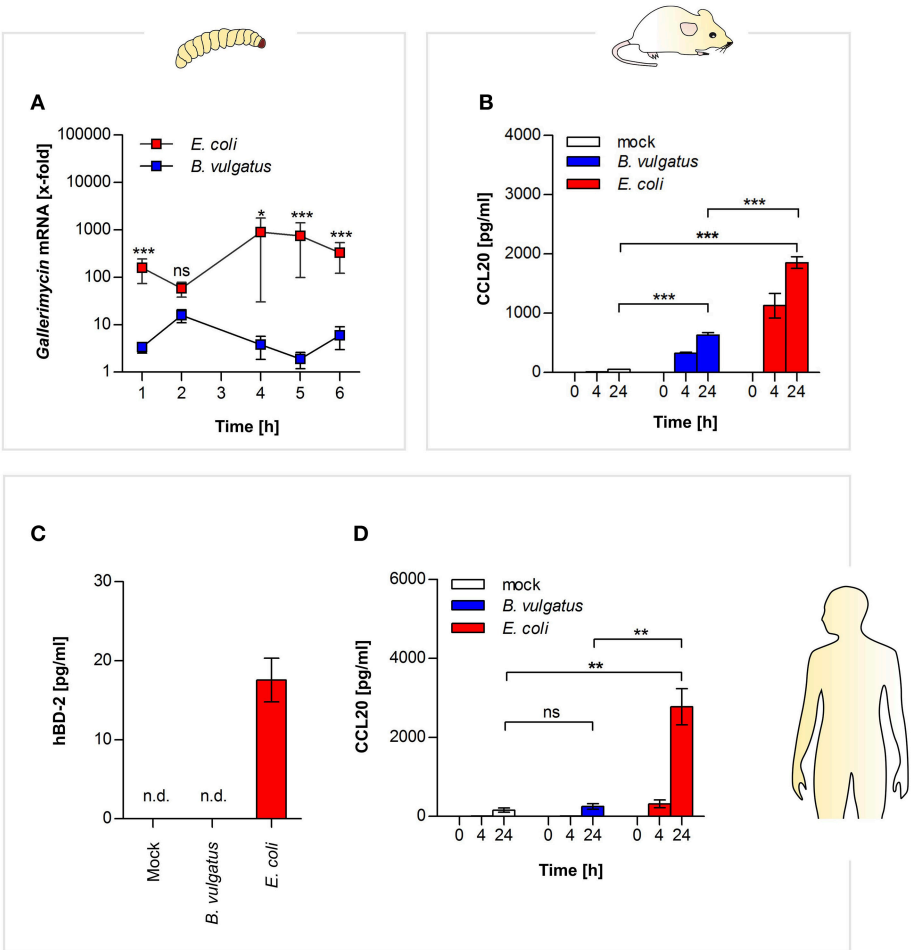


FIGURE 4 | Commensal-induced antimicrobial responses in insect larvae, and mammalian epithelial cells. *Gallerimycin* gene expression was measured in RNA from *G. mellonella* larvae after *E. coli* or *B. vulgatus* oral administration ($n = 8$) **(A)**. Secretion of CCL20 was determined in mouse mICl2 cell culture supernatants after commensal stimulation ($n = 3$) **(B)**. Human SW480 cells were stimulated with *E. coli* or *B. vulgatus* for 0, 4, and 24 h. hBD-2 **(C)** and human CCL20 **(D)** was measured by ELISA in culture supernatants ($n = 3$). Data represent the geometric means ± SEM. (n. d.: not detected).

E. coli leads to secretion of human BD-2 after 24 h in cell culture supernatants **(Figure 4C)**. Secretion of human CCL20 protein was measured by ELISA, and significant higher CCL20 was produced after 24 h in supernatants of *E. coli*-treated cells as compared to *B. vulgatus*-stimulated cells **(Figure 4D)**.

In summary we provide evidence that symbionts or pathobionts trigger similar innate immune responses in both mammals and insects and that hence *G. mellonella* may be a suitable replace organism to distinguish mammalian intestinal symbionts form pathobionts.

DISCUSSION

In our study we present a comprehensive characterization of different innate immune mechanisms toward symbiotic and pathobiotic intestinal commensals in an invertebrate *G. mellonella* and vertebrate mouse model.

Previous *in silico* analysis of different components of innate immune responses, such as enzymes relevant in ROS signaling and AMP secretion, verified how closely related innate mechanisms among the models are and revealed evolutionary conservation of those genes involved in creating the responses. After *in silico* data generation, we created the hypothesis that in response to intestinal exposition of either a bacterial symbiont or pathobiont both vertebrates and invertebrates might activate comparable immune mechanisms. This hypothesis was successfully confirmed by our experimental data.

Both *G. mellonella,* as an insect representative, and mice, belonging to the mammalian class, are able to differentiate self from non-self and are capable to further discriminate between commensals. The ability to recognize foreign non-self elements is ancient and conserved among many primitive invertebrates (ameba), advanced invertebrates (sea stars) and of course among higher vertebrates like mice or rats (29). In order to react toward a non-self signal of potentially harmful invaders e.g., pathogens,

hosts require PRRs that are able to sense microbial components–MAMPs–like LPS, peptidoglycan, flagellin, or lipoteichoic acid (LTA). The recognition of these MAMPs is a prerequisite for the initiation of appropriate follow-up mechanisms to eliminate potential intruders (30, 31). Innate immune responses are major contributors to acute inflammation induced by bacterial infection (32).

In mammals LPS released from Gram-negative bacteria can bind to the extracellular LPS-binding protein followed by binding to the CD14 molecule, a surface-expressed opsonic co-receptor. This reaction leads to the transfer of LPS to the MD2 accessory molecule and association with the TLR4 receptor, resulting in further immune signaling (31, 33). The upregulation of TLR4 receptor during intestinal inflammation was observed (10). Intestinal dendritic cells (DCs) have a key role in innate microbe recognition and influence adaptive immune responses (34). Intestinal DCs are decisive to differentiate between commensals and potential pathogens, and decide on maintenance of homeostasis or active immune responses (35). In previous work, we conveyed that *E. coli* and *B. vulgatus* mpk are differentially recognized by TLR2- and TLR4-overexpressing human embryonic kidney (HEK) cells. *E. coli* leads to much stronger stimulation of both TLR2 and 4 than *B. vulgatus* (36).

In *G. mellonella* two LPS-binding proteins were identified, which are considered to be PRRs (30). A highly abundant PRR in lepidopteran insects is apolipophorin which binds to LPS and triggers antimicrobial responses (30, 37). It is reported that *G. mellonella* manages to induce immune responses with differing intensities toward various microbes, which assumes also the differential recognition (38). However, all available information of *G. mellonella* MAMP recognition was obtained from studies involving systemic hemolymph infection. Until now, no evidence has been provided for differential intestinal recognition of Gram-negative microbes in *G. mellonella*. Interestingly, we also found upregulated hemolin gene expression, which is abundant in insect hemolymph and interacts with LPS and LTA (39). Hemolin a member of the immunoglobulin superfamily is also proposed to function as a recognition molecule (39). Both apolipophorin and hemolin work like opsonins (15). These results might suggest that certain bacterial compounds might cross the intestinal barrier and involve the hemolymph responses.

Insects and mammals share similar effectors that are secreted upon microbial challenge, such as production of ROS/RNS or AMPs (4, 29, 40), but only little is reported concerning insects and their immune responses toward intestinal microbial challenge. The majority of information available about insect intestinal immune responses toward microbes was obtained from *Drosophila*. The involvement of ROS responses to maintain intestinal homeostasis was recently reported in *Drosophila* (40). Xiao and colleagues showed that DUOX-related signaling was involved in balancing the load of the intestinal microbiota (40). DUOX belongs to the family of NADPH oxidases, which we could also identify in *G. mellonella* (41). These mentioned results are in agreement with our data generated in the present study. We could show that *G. mellonella* induces ROS- and RNS-related responses by *Nox-4* and *Nos* expression, which contributes to a decrease of bacterial load. Besides, we could provide evidence that *G. mellonella* induced differential ROS/RNS responses upon *E. coli* and *B. vulgatus* oral administration. The oxidative stress responses generated to clear the excessive bacteria was antagonized by induced *Gst* gene expression, which acts as an antioxidative molecule (6, 42). We assumed that the produced ROS and RNS were not intended to substantially harm the host, as the stress response was antagonized to retain cellular homeostasis. *G. mellonella* larvae have never been reported before to maintain gut homeostasis and preserve microbiota balance *via* intestinal epithelial cells. The *G. mellonella* model is widely used as a systemic infection model, in which ROS responses were determined in hemocytes after pathogen infection of the hemolymph (43).

Intestinal infections and inflammation in mammals are characterized by production of reactive oxygen and nitrogen species (44). ROS and RNS both similarly contribute to suppression of microbial growth and intestinal homeostasis in mammals (6). We were also able to detect elevated *Nos2* gene expression upon microbial challenge in mouse host cells using *in vitro* cell culture and *in vivo* in $Rag1^{-/-}$ T cell-transplanted mice mono-associated with bacteria. Higher *Nos2* expression was found in response to *E. coli* stimulation. *Nos2* is highly expressed during pathogen infection and accumulates nitric oxide (NO) (45).NOS2 is mainly responsible for NO production, but it is further able to produce ROS, when concentrations of tetrahydrobiopterin or the co-substrate L-arginine is available in low quantities (6). $Rag1^{-/-}$ mice are a well-established model for the influence of commensal bacteria on intestinal immune balance. The commensal induced innate immune response differs depending on the activation potential of the commensal and results in proliferation and polarization of T cells, which in turn might limit the commensal induced innate activation (46). Although here we used T cell transplanted $Rag1^{-/-}$ mice we can observe innate immune responses comparable to that of *G. mellonella*, suggesting that there is no downregulation of the innate response by the transplanted T cells.

Some bacteria can overcome the oxidative stress responses, but can be eliminated by AMP responses, which serves as complementary second line antibacterial defense response (4, 41). We could show in both our invertebrate and in the mouse model that *E. coli* caused a higher production of AMPs compared to *B. vulgatus*. In *G. mellonella* all measured AMPs showed higher expression in *E. coli* administered larvae, which indeed points toward a distinct recognition and defined immune response according to different Gram-negative commensals. This is in line with the previous differential expression of PRRs in responses to those bacteria. We analyzed five out of 18 known AMPs from different AMP classes such as defensin-like peptides, cecropins, or moricins which are frequently investigated. When the *G. mellonella* immune system is challenged by a certain stimulus there is always a set of structurally and functionally distinct AMPs which is produced to enable a strong and effective response. A recent report showed that such AMP co-production along with the synergy of AMPs was an effective defense response due to potentiated antimicrobial effects (11).

Mice and humans secrete AMPs as well to maintain intestinal homeostasis. Different classes of antimicrobial peptides are secreted in the intestine such as C-type lectins, cathelicidins, or defensins. The latter two represent the most abundant types of AMPs (47). We were interested in the mammalian response of defensins, as similar molecules are also present in *G. mellonella*. The intestinal defensins are divided into α- and β-defensins and are at least in part dependent on innate PRRs to induce their transcription or secretion. α-defensins (HD5 and HD6) are produced by specialized Paneth cells whereas β-defensins are produced within different epithelial cells including enterocytes which contributes to a broader significance of β-defensins (10). BD-1 is constitutively expressed whereas BD-2, BD-3, and BD-4 show increased expression in response to pathogen attack (5). BD-2 and the antimicrobial chemokine CCL20 share structural homology and are able to bind to the CCR6 receptor. CCR6-CCL20 and CCR6-BD-2 respectively, play an important role in maintenance of intestinal homeostasis and restrict overgrowth of commensal bacteria (5, 48).

Using *in vitro* mouse small intestinal cells we could determine increased amounts of CCL20 after *E. coli* stimulation. We could confirm these results in human epithelial cells and even detect low amounts of the structurally similar AMP hBD-2 after stimulation with *E. coli*. AMP production is TLR-dependent and hBD-2 accumulates after TLR2, TLR3, and TLR4 recognition (5). It was shown that both hBD-2 and CCL20 were induced upon Caco-2 cell infection with EPEC, and are highly abundant in IBD patients (5). We propose that the high accumulation of both antimicrobial molecules might also be the consequence of the increased amounts of *E. coli* during the infection and inflammation, and its potential to stimulate TLR4 than only the inflammatory conditions.

Interpretation of our data suggests that the monitored immune reactions occur in different sequential timeframes and can be divided into faster and later responses: ROS- and RNS-related gene expression was upregulated quite early between 2 and 3 h, and antagonized by GST after 5 h, whereas the secretion of AMPs was induced only after oxidative stress responses have been completed. Thus, bacterial recognition must have occurred previously, since expression of apolipophorin, which recognizes and interacts with LPS structures, was induced already after 2 h.

For the first time we provide evidence that *G. mellonella* recognizes commensals or their MAMPs by intestinal PRRs, and triggers effector molecules to cause oxidative stress responses, as well as the production of AMPs. Further we could uncover hints for differential intestinal commensal recognition that allow *G. mellonella* for discrimination of Gram-negative symbionts from pathobionts. Administration of the symbiont *B. vulgatus* leads to less pronounced immune responses (**Figure 5**, right) whereas administration of *E. coli* initiates stronger immune responses (**Figure 5**, left). Nevertheless, *E. coli* is not necessarily a pathobiont in *G. mellonella*, but still *E. coli* and *B. vulgatus* caused immune responses with different intensities and show a different immunogenic potential. The findings obtained from insects were therefore comparable to data obtained from vertebrates. We conclude that intestinal innate immune responses among *G. mellonella* and mammals are evolutionary conserved in response to challenge with intestinal commensals. Our results indicate that invertebrates can initiate proper immune responses to excessive loads of microbes and clear bacterial dysbiosis.

Symbionts and pathobionts and their recognition play crucial roles in intestinal homeostasis (2) and with our invertebrate *G. mellonella* model we provide an alternative model to

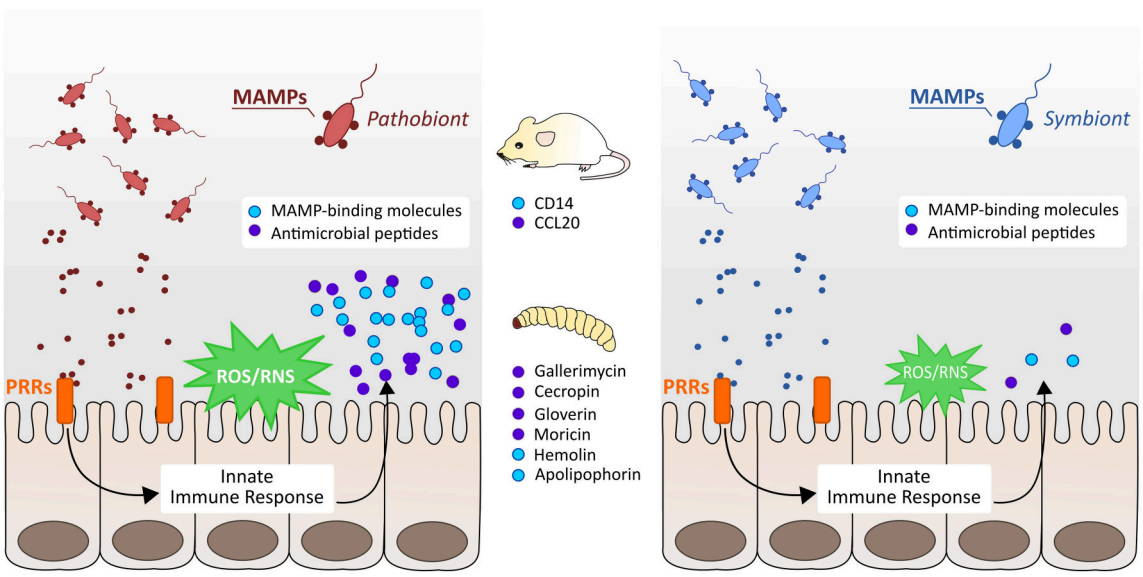


FIGURE 5 | Schematic overview of vertebrate and invertebrate innate immune responses to symbiotic and pathobiotic bacteria. In both mice and *G. mellonella* larvae, pathobionts cause strong immune responses by the production of MAMP-binding recognition proteins, induction of reactive oxygen and nitrogen species (ROS/RNS), and secretion of different antimicrobial peptides (AMPs) (left). Symbionts cause only less pronounced immune responses and induce only low expression of immunity markers (right).

screen commensals for their immunogenic potential. Hence we suggest *G. mellonella* as an invertebrate replacement model to discriminate intestinal commensals according to their immunogenic potential. Our future work will aim on the investigation of further microbes and their potential to activate innate immune responses via the intestinal tract in the insect model *G. mellonella*.

ETHICS STATEMENT

This study was carried out in accordance with the principles of the Basel Declaration. Protocols and experiments involving mice were reviewed and approved by the responsible Institutional Review Committee and the local authorities within H1/15 approval.

The work involving invertebrates does not need ethical permission according to German law.

AUTHOR CONTRIBUTIONS

AL, ASt, and J-SF conceived and designed the experiments. AL, AS, AB, and RP performed the experiments. AL, ASt, and J-SF analyzed the data. SB performed the protein alignments. AL, ASt, and J-SF wrote the manuscript. All authors gave final approval to publish the article.

FUNDING

This work was supported by the DFG (SPP1656), the DFG research training group 1708, the Bundesministerium für Bildung und Forschung (BMBF) and the German Center for Infection Research (DZIF).

ACKNOWLEDGMENTS

We would like to thank Brianna Ball, Jan Maerz, and Brigitte Beifuss for their support with editing and proof-reading of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2018.02114/full#supplementary-material

REFERENCES

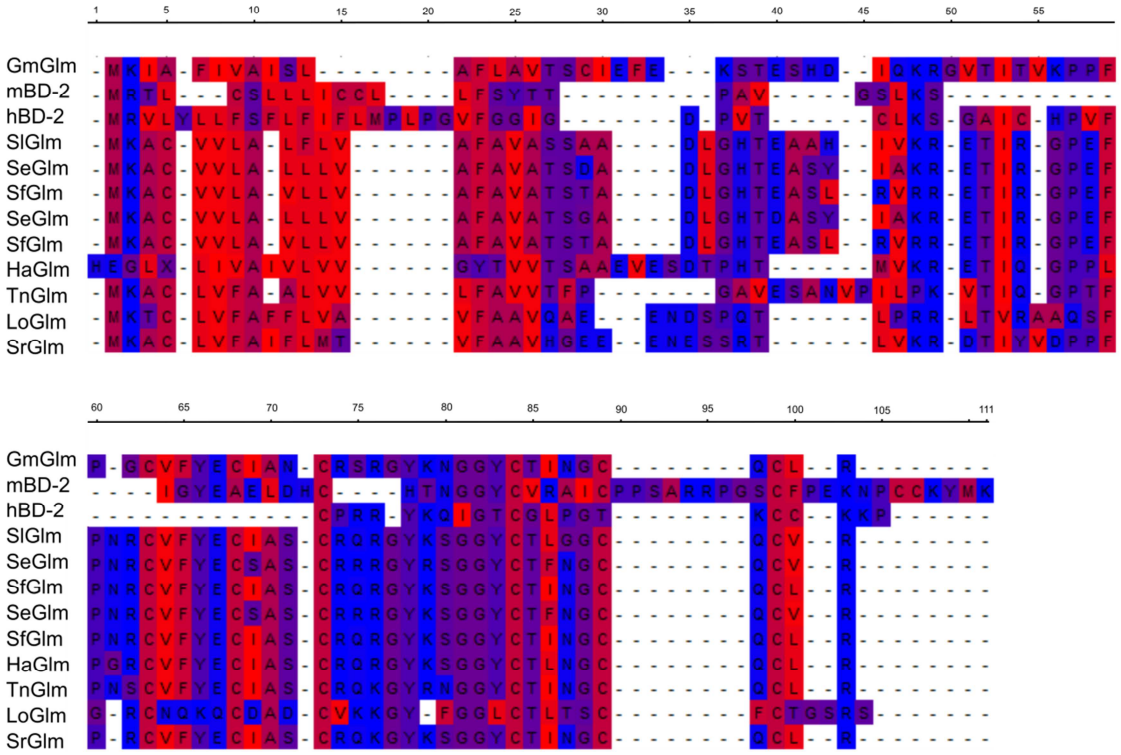
1. Ayres JS. Inflammasome-microbiota interplay in host physiologies. *Cell Host Microbe* (2013) 14:491–7. doi: 10.1016/j.chom.2013.10.013
2. Kamada N, Seo SU, Chen GY, Nunez G. Role of the gut microbiota in immunity and inflammatory disease. *Nat Rev Immunol.* (2013) 13:321–35. doi: 10.1038/nri3430
3. Atarashi K, Suda W, Luo C, Kawaguchi T, Motoo I, Narushima S, et al. Ectopic colonization of oral bacteria in the intestine drives TH1 cell induction and inflammation. *Science* (2017) 358:359–65. doi: 10.1126/science.aan4526
4. Kim SH, Lee WJ. Role of DUOX in gut inflammation: lessons from Drosophila model of gut-microbiota interactions. *Front Cell Infect Microbiol.* (2014) 3:116. doi: 10.3389/fcimb.2013.00116
5. Muniz LR, Knosp C, Yeretssian G. Intestinal antimicrobial peptides during homeostasis, infection, and disease. *Front Immunol.* (2012) 3:310. doi: 10.3389/fimmu.2012.00310
6. Nathan C, Cunningham-Bussel A. Beyond oxidative stress: an immunologist's guide to reactive oxygen species. *Nat Rev Immunol.* (2013) 13:349–61. doi: 10.1038/nri3423
7. Fang FC. Antimicrobial reactive oxygen and nitrogen species: concepts and controversies. *Nat Rev Microbiol.* (2004) 2:820–32. doi: 10.1038/nrmicro1004
8. Grasberger H, Gao J, Nagao-Kitamoto H, Kitamoto S, Zhang M, Kamada N, et al. Increased expression of DUOX2 is an epithelial response to mucosal dysbiosis required for immune homeostasis in mouse intestine. *Gastroenterology* (2015) 149:1849–59. doi: 10.1053/j.gastro.2015.07.062
9. Bae YS, Oh H, Rhee SG, Yoo YD. Regulation of reactive oxygen species generation in cell signaling. *Mol Cells* (2011) 32:491–509. doi: 10.1007/s10059-011-0276-3
10. Ostaff MJ, Stange EF, Wehkamp J. Antimicrobial peptides and gut microbiota in homeostasis and pathology. *EMBO Mol Med.* (2013) 5:1465–83. doi: 10.1002/emmm.201201773
11. Bolouri Moghaddam MR, Tonk M, Schreiber C, Salzig D, Czermak P, Vilcinskas A, et al. The potential of the *Galleria mellonella* innate immune system is maximized by the co-presentation of diverse antimicrobial peptides. *Biol Chem.* (2016) 397:939–45. doi: 10.1515/hsz-2016-0157
12. Mukherjee S, Hooper LV. Antimicrobial defense of the intestine. *Immunity* (2015) 42:28–39. doi: 10.1016/j.immuni.2014.12.028
13. Waidmann M, Bechtold O, Frick JS, Lehr HA, Schubert S, Dobrindt U, et al. Bacteroides vulgatus protects against *Escherichia coli*-induced colitis in gnotobiotic interleukin-2-deficient mice. *Gastroenterology* (2003) 125:162–77. doi: 10.1016/S0016-5085(03)00672-3
14. Steimle A, Gronbach K, Beifuss B, Schafer A, Harmening R, Bender A, et al. Symbiotic gut commensal bacteria act as host cathepsin S activity regulators. *J Autoimmun.* (2016) 75:82–95. doi: 10.1016/j.jaut.2016.07.009
15. Tsai CJ, Loh JM, Proft T. *Galleria mellonella* infection models for the study of bacterial diseases and for antimicrobial drug testing. *Virulence* 7:214–29. doi: 10.1080/21505594.2015.1135289
16. Lange A, Beier S, Steimle A, Autenrieth IB, Huson DH, Frick JS. Extensive mobilome-driven genome diversification in mouse gut-associated *Bacteroides vulgatus* mpk. *Genome Biol Evol.* (2016) 8:1197–207. doi: 10.1093/gbe/evw070
17. Vogel H, Altincicek B, Glockner G, Vilcinskas A. A comprehensive transcriptome and immune-gene repertoire of the lepidopteran model host *Galleria mellonella*. *BMC Genomics* (2011) 12:308. doi: 10.1186/1471-2164-12-308
18. Lewis AC, Jones NS, Porter MA, Deane CM. What evidence is there for the homology of protein-protein interactions? *PLoS Comput Biol.* (2012) 8:e1002645. doi: 10.1371/journal.pcbi.1002645
19. Pearson WR. An introduction to sequence similarity ("homology") searching. *Curr Protoc Bioinformatics* (2013) Chapter 3, Unit3 1. doi: 10.1002/0471250953.bi0301s42
20. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* (2002) 30:3059–66. doi: 10.1093/nar/gkf436
21. Mennechet FJ, Kasper LH, Rachinel N, Li W, Vandewalle A, Buzoni-Gatel D. Lamina propria CD4+ T lymphocytes synergize with murine intestinal epithelial cells to enhance proinflammatory response against an intracellular pathogen. *J Immunol.* (2002) 168:2988–96. doi: 10.4049/jimmunol.168.6.2988
22. Lutz MB, Kukutsch N, Ogilvie AL, Rossner S, Koch F, Romani N, et al. An advanced culture method for generating large quantities of highly pure dendritic cells from mouse bone marrow. *J ImmunolMethods* (1999) 223:77–92. doi: 10.1016/S0022-1759(98)00204-X
23. Hermann-Bank ML, Skovgaard K, Stockmarr A, Larsen N, Molbak L. The gut microbiotassay: a high-throughput qPCR approach combinable with next generation sequencing to study gut microbial diversity. *BMC Genomics* (2013) 14:788. doi: 10.1186/1471-2164-14-788

24. Sato K, Kumita W, Ode T, Ichinose S, Ando A, Fujiyama Y, et al. OmpA variants affecting the adherence of ulcerative colitis-derived Bacteroides vulgatus. *J Med Dent Sci.* (2010) 57:55–64. doi: 10.11480/jmds.570107
25. Mazmanian SK, Round JL, Kasper DL. A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature* (2008) 453:620–5. doi: 10.1038/nature07008
26. Zanoni I, Ostuni R, Marek LR, Barresi S, Barbalat R, Barton GM, et al. CD14 controls the LPS-induced endocytosis of Toll-like receptor 4. *Cell* (2011) 147:868–80. doi: 10.1016/j.cell.2011.09.051
27. Muralidharan S, Mandrekar P. Cellular stress response and innate immune signaling: integrating pathways in host defense and inflammation. *J Leukoc Biol.* (2013) 94:1167–84. doi: 10.1189/jlb.0313153
28. Ganz T. The role of antimicrobial peptides in innate immunity. *Integr Comp Biol.* (2003) 43:300–4. doi: 10.1093/icb/43.2.300
29. Buchmann K. Evolution of innate immunity: clues from invertebrates via fish to mammals. *Front Immunol.* (2014) 5:459. doi: 10.3389/fimmu.2014.00459
30. Brivio M, Mastore M, Pagani M. Parasite-host relationship a lesson from a professional killer. *Invert Surv J.* (2005) 2:41–53.
31. Mogensen TH. Pathogen recognition and inflammatory signaling in innate immune defenses. *Clin Microbiol Rev.* (2009) 22:240–73. doi: 10.1128/CMR.00046-08
32. Takeuchi O, Akira S. Pattern recognition receptors and inflammation. *Cell* (2010) 140:805–20. doi: 10.1016/j.cell.2010.01.022
33. Aderem A, Ulevitch RJ. Toll-like receptors in the induction of the innate immune response. *Nature* (2000) 406:782–7. doi: 10.1038/35021228
34. Hart AL, Al-Hassi HO, Rigby RJ, Bell SJ, Emmanuel AV, Knight SC, et al. Characteristics of intestinal dendritic cells in inflammatory bowel diseases. *Gastroenterology* (2005) 129:50–65. doi: 10.1053/j.gastro.2005.05.013
35. Ng SC, Kamm MA, Stagg AJ, Knight SC. Intestinal dendritic cells: their role in bacterial recognition, lymphocyte homing, and intestinal inflammation. *Inflamm Bowel Dis.* (2010) 16:1787–807. doi: 10.1002/ibd.21247
36. Maerz JK, Steimle A, Lange A, Bender A, Fehrenbacher B, Frick JS. Outer membrane vesicles blebbing contributes to *B. vulgatus* mpk-mediated immune response silencing. Gut *Microbes* (2018) 9:1–12. doi: 10.1080/19490976.2017.1344810
37. Staczek S, Zdybicka-Barabas A, Mak P, Sowa-Jasilek A, Kedracka-Krok S, Jankowska U, et al. Studies on localization and protein ligands of *Galleria mellonella* apolipophorin III during immune response against different pathogens. *J Insect Physiol.* (2018) 105:18–27. doi: 10.1016/j.jinsphys.2017.12.009
38. Mak P, Zdybicka-Barabas A, Cytrynska M. A different repertoire of *Galleria mellonella* antimicrobial peptides in larvae challenged with bacteria and fungi. *Dev Comp Immunol.* (2010) 34:1129–36. doi: 10.1016/j.dci.2010.06.005
39. Yu XQ, Kanost MR. Binding of hemolin to bacterial lipopolysaccharide and lipoteichoic acid. An immunoglobulinsuperfamily member from insects as a pattern-recognition receptor *Eur J Biochem.* (2002) 269:1827–34. doi: 10.1046/j.1432-1033.2002.02830.x
40. Xiao X, Yang L, Pang X, Zhang R, Zhu Y, Wang P, et al. A Mesh-Duox pathway regulates homeostasis in the insect gut. *Nat Microbiol.* (2017) 2:17020. doi: 10.1038/nmicrobiol.2017.20
41. Ha EM, Lee KA, Seo YY, Kim SH, Lim JH, Oh BH, et al. Coordination of multiple dual oxidase-regulatory pathways in responses to commensal and infectious microbes in drosophila gut. *Nat Immunol.* (2009) 10:949–57. doi: 10.1038/ni.1765
42. Mone Y, Monnin D, Kremer N. The oxidative environment: a mediator of interspecies communication that drives symbiosis evolution. *Proc Biol Sci.* (2014) 281:20133112. doi: 10.1098/rspb.2013.3112
43. Bergin D, Reeves EP, Renwick J, Wientjes FB, Kavanagh K. Superoxide production in *Galleria mellonella* hemocytes: identification of proteins homologous to the NADPH oxidase complex of human neutrophils. *Infect Immun.* (2005) 73:4161–70. doi: 10.1128/IAI.73.7.4161-4170.2005
44. Baumler AJ, Sperandio V. Interactions between the microbiota and pathogenic bacteria in the gut. *Nature* (2016) 535:85–93. doi: 10.1038/nature18849
45. Okumura R, Takeda K. Roles of intestinal epithelial cells in the maintenance of gut homeostasis. *Exp Mol Med.* (2017) 49:e338. doi: 10.1038/emm.2017.20
46. Gronbach K, Flade I, Holst O, Lindner B, Ruscheweyh HJ, Wittmann A, et al. Endotoxicity of lipopolysaccharide as a determinant of T-cell-mediated colitis induction in mice. *Gastroenterology* (2014) 146:765–75. doi: 10.1053/j.gastro.2013.11.033
47. Agier J, Efenberger M, Brzezinska-Blaszczyk E. Cathelicidin impact on inflammatory cells. *Cent Eur J Immunol.* (2015) 40:225–35. doi: 10.5114/ceji.2015.51359
48. Lin YL, Ip PP, Liao F. CCR6 deficiency impairs iga production and dysregulates antimicrobial peptide production, altering the intestinal flora. *Front Immunol.* (2017) 8:805. doi: 10.3389/fimmu.2017.00805

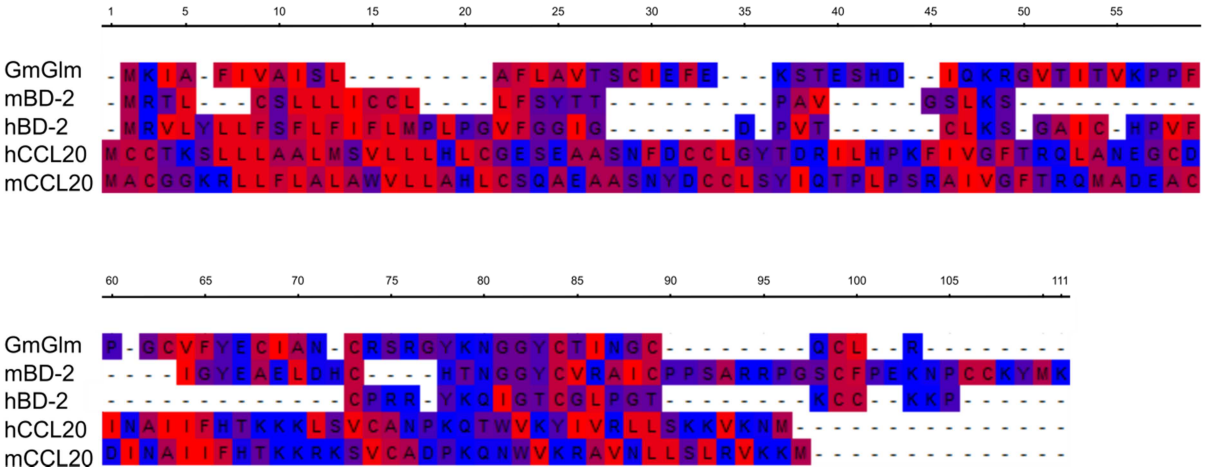
Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary Material

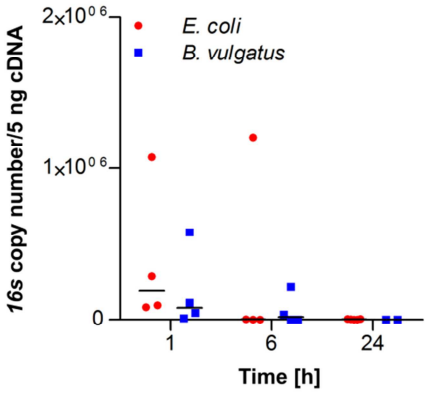
Supplementary Figure S1: Protein alignment of *Galleria mellonella* and other Lepidoptera insect gallerimycins, human and mouse beta-defensins with highlighted hydrophobic (red) and hydrophilic (blue) areas.



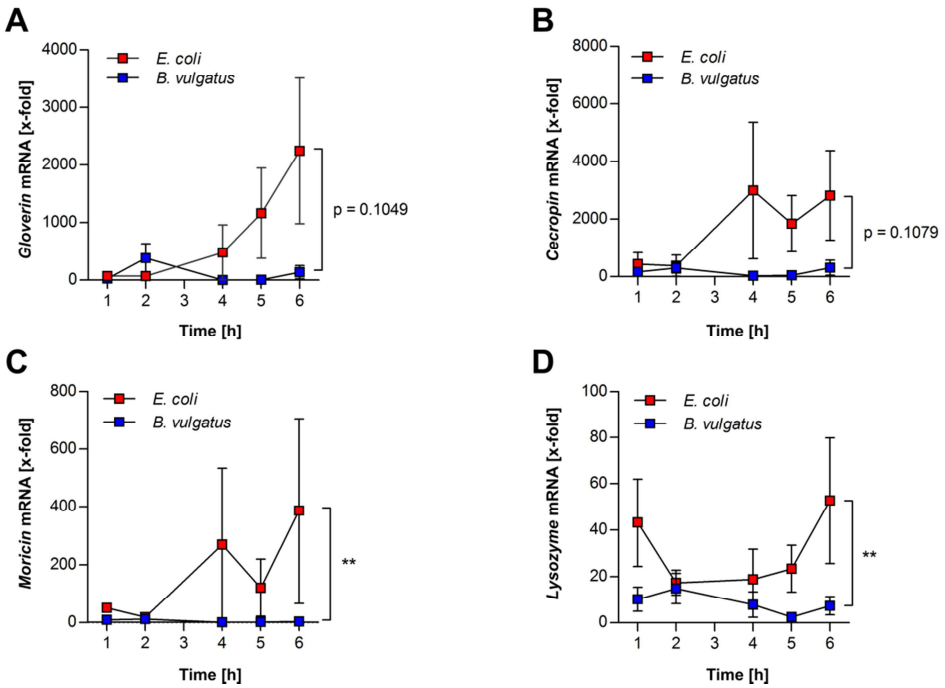
Supplementary Figure S2: Protein alignment of *Galleria mellonella* gallerimycin, human and mouse beta-defensin and CCL20 with highlighted hydrophobic (red) and hydrophilic (blue) areas.



Supplementary Figure S3: Persistence of bacterial load in *Galleria mellonella* larvae after force-feeding. Copy numbers of *B. vulgatus*- and *E. coli*-specific 16s rDNA genes were determined from 5 ng of cDNA at different time points using RT-PCR.



Supplementary Figure S4: Analysis of antimicrobial peptide (AMP) gene expression.
RNA from orally administered larvae was isolated and RT-PCR was used to investigate gene expression of different AMPs: LPS-responsive gloverin (A), cecropin (B) and moricin (C) with activity against Gram-negative bacteria, and lysozyme (D).



Supplementary Table S1: *Galleria mellonella* sequences obtained from Transcriptome dataset

Excel table online:

https://www.frontiersin.org/articles/10.3389/fimmu.2018.02114/full#supplementary-material

Supplementary Table S2: Primers used in this study

Primer name	Sequence 5´-3´	Reference
Gallerimycin f	GAAGTCTACAGAATCACACGA	This study
Gallerimycin r	ATCGAAGACATTGACATCCA	
ubiquitin f	TCAATGCAAGTAGTCCGGTTC	Virulence (2014) 5: (4) 547-554
ubiquitin r	CCAGTCTGCTGCTGATAAACC	
Apo III f	AGACTTGCACGCCATCAAGA	This study
Apo III r	TGCATGCTGTTTGTCACTGC	
Gloverin f	GTGTTGAGCCCGTATGGGAA	This study
Gloverin r	CCGTGCATCTGCTTGCTAAC	
GST1 f	GACAGAAGTCCTCCGGTCAG	This study
GST r	TCCGTCTTCAAGCAAAGGCA	
Lysozyme f	GGACTGGTCCGAGCACTTAG	This study
Lysozyme r	CGCATTTAGAGGCAACCGTG	
Cecropin f	CTGTTCGTGTTCGCTTGTGT	This study
Cecropin r	GTAGCTGCTTCGCCTACCAC	
Moricin f	GCTGTACTCGCTGCACTGAT	This study
Moricin r	TGGCGATCATTGCCCTCTTT	
Hemolin f	CTCCCTCACGGAGGACAAAC	This study
Hemolin r	GCCACGCACATGTATTCACC	
NOX-4 f	TGGCACGGCATCAGTTATCA	This study
NOX-4 r	ACAGCGACTGTCATGTGGAA	
NOS f	ATGAAGGTGCTGAAGTCACAA	This study
NOS r	GCCATTTTACAATCGCCACAA	
NOS2 f	GTTCTCAGCCCAACAATACAAGA	This study
NOS2 r	GTGGACGGGTCGATGTCAC	
GSTP1 f	CGGCAAATATGTCACCCTCAT	This study
GSTP1 r	GCCAGGACTTGGTGGATCAG	
CD14 f	GACCATGGAGCGTGTGCTTG	This study
CD14 r	GGACCAATCTGGCTTCGGAT	
β-Actin f	CCCTGTGCTGCTCACCGA	This study
β-Actin r	ACAGTGTGGGTGACCCCGTC	

Supplementary Table S1: BLASTX analysis of glutathione S-transferase (GST) transcripts

G. mellonella	*Mus musculus*				*Homo sapiens*			
Transcript	**BLAST Hit**	**Identity**	**Query coverage**	**E value**	**BLAST Hit**	**Identity**	**Query coverage**	**E value**
GST1	GST theta	27%	47%	$1x10^{-12}$	GST theta-1	31%	36%	$3x10^{-14}$
GST2	GST theta-3	31%	59%	$2x10^{-18}$	GST theta-4	33%	57%	$9x10^{-20}$
GST3	GST theta-4	34%	39%	$1x10^{-12}$	GST theta-2	34%	41%	$3x10^{-11}$
GST4	GST Mu 4	27%	86%	$2x10^{-13}$	GST M2-3	26%	87%	$1x10^{-13}$
GST5	GST omega-1	38%	68%	$2x10^{-45}$	GST Omega 1	39%	68%	$4x10^{-46}$
GST6	GST omega-1	33%	68%	$4x10^{-35}$	GSTO1	36%	71%	$1x10^{-38}$
GST7	GST theta-3	28%	70%	$9x10^{-25}$	GST T1	33%	69%	$1x10^{-23}$
GST8	GST theta-3	38%	48%	$5x10^{-41}$	GST theta 1	38%	48%	$2x10^{-40}$
GST9	GST theta-2	30%	27%	$1x10^{-19}$	GST theta-1	29%	25%	$1x10^{-17}$
GST10	microsomal GST 1	41%	48%	$3x10^{-27}$	microsomal GST 1	43%	48%	$1x10^{-26}$
GST11	GST theta 3	28%	71%	$2x10^{-21}$	GST theta-2	34%	68%	$1x10^{-22}$
GST12	GST theta-4	39%	23%	$8x10^{-11}$	GST theta-1	34%	25%	$3x10^{-09}$
GST13	GST	29%	57%	$5x10^{-19}$	GST A5	30%	56%	$8x10^{-15}$
GST14	GST theta-4	29%	68%	$3x10^{-20}$	GSTT2	30%	87%	$6x10^{-23}$
GST15	no match	-	-	-	no match	-	-	-
GST16	GST theta	27%	77%	$3x10^{-15}$	GST	31%	67%	$6x10^{-15}$
GST17	GST theta	29%	71%	$2x10^{-24}$	GST T1	30%	71%	$1x10^{-23}$
GST18	GST Zeta 1-1	47%	66%	$2x10^{-59}$	GST zeta 1	45%	65%	$5x10^{-56}$
GST19	GST theta	30%	67%	$3x10^{-18}$	GST T1	33%	67%	$1x10^{-17}$