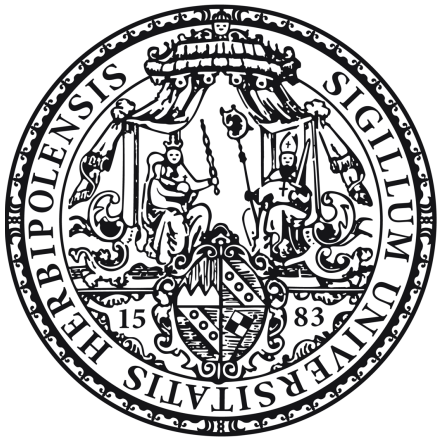


Analysis of gastrointestinal epithelial innate immune barrier using human and murine organoids as a model

Analyse der Gastrointestinalen angeborenen Immunbarriere durch Humane und Murine Organoide als Modell

Doctoral thesis for a doctoral degree

at the Graduate School of Life Sciences,

Julius-Maximilians-Universität Würzburg,

Section Infection and Immunity

submitted by

Özge Kayisoglu-Kaya

from

Altindag, Ankara, Turkey

Würzburg, 2021

Submitted on: ……...……...

Office stamp

Members of the Thesis Committee

Chairperson: Prof. Dr. Markus Sauer

Primary Supervisor: Dr. Sina Bartfeld

Supervisor (Second): Prof. Dr. Thomas Rudel

Supervisor (Third): Dr. Bon-Kyoung Koo

Date of Public Defence: ……..……

Date of Receipt of Certificates:

…….

Table of Contents

Abbreviations

2D	Two-dimensional
3D	Three-dimensional
3R	3R Principles (Replacement, Reduction, Refinement)
AD	Advanced DMEM/F12
ADP	Adenosine diphosphate
ADP-heptose	ADP-D-glycero-β-D-manno-heptose
AIM2	Absent in melanoma 2
ALPK1	Alpha-kinase 1
AMP	Adenosine monophosphate
AP1	Activator protein 1
ASCs	Adult stem cells
ASC	Apoptosis-associated speck-like protein containing a CARD
ATAC-seq	Assay for transposase-accessible chromatin using sequencing
ATP	Adenosine triphosphate
BARX1	BarH-like homeobox 1
BIR	Baculoviral inhibitor of apoptosis repeat
BMP	Bone morphogenic protein
BSA	Bovine serum albumin
CAR1	Carbonic anhydrase 1
CARD	Caspase activation and recruitment domain
CASP1	Caspase-1
CD	Crohn's disease
CD14	Cluster of differentiation 14
CD45	Cluster of differentiation 45
CDX	Caudal related homeobox
cGAMP	Cyclic GMP-AMP
cGAS	Cyclic GMP-AMP synthase
CLR	C-type lectin receptor
CO_2	Carbon dioxide
CpG	5'-C-phosphate-G-3'
ICAM3	Intercellular adhesion molecule 3
DAMP	Damage-associated molecular pattern
DCSIGN	Dendritic cell-specific ICAM3-grabbing nonintegrin
DEG	Differentially expressed genes
DLL	Delta-like protein
DMEM	Dulbecco’s Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid

DTT	Dithiothreitol
E(dd)	embryonic day
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EGFP	Enhanced green fluorescent protein
ENRY	EGF, Noggin, R-spondin1, Rho-kinase inhibitor
ER	EGF, R-spondin1
ERK	Extracellular signal-regulated kinases
ETB	EGF, thiazovivin, BMP-4
FGF10	Fibroblast growth factor 10
FITC	Fluoroscein isothiocyanate
FN1	Fibronectin 1
FS	Forestomach
GALT	Gut-associated lymphoid tissue
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GEO	Gene Expression Omnibus
GFP	Green fluorescent protein
GI	Gastrointestinal
GIF	Gastric intrinsic factor
GLI1	Glioma associated oncogene 1
GMP	Guanosine monophosphate
GO	Gene ontology
GREM1	Gremlin 1
GSDMD	Gasdermin D
GTP	Guanosine triphosphate
H_2O	Water
H3K27ac	Histone H3 lysine 27 acetylated
H3K4me3	Histone H3 lysine 4 trimethylated
H3K36me3	Histone H3 lysine 36 trimethylated
HA	Human influenza hemagglutinin
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HCl	Hydrochloric acid
HH	Hedgehog signaling
HS	Hindstomach
IBD	Inflammatory bowel disease
IFN	Interferon
IGV	Integrative Genome Browser
IκB	Inhibitor of NF-κB
IKK	Inhibitor of NF-κB (IκB) kinase complex
IL	Interleukin

IRAK	IL-1 receptor associated kinase
IRF	Interferon regulatory factor
JNK	c-Jun N-terminal kinase
KH_2PO_4	Monopotassium phosphate
KI67	Marker of proliferation KI67
LBP	Lipopolysaccharide binding protein
LGR5	Leucine-rich repeat-containing G-protein coupled receptors 5
LPS	Lipopolysaccharide
LRR	Leucine-rich repeat
M-cell	Microfold cell
MAMP	Microbe-associated molecular pattern
MAP	Mitogen-activated protein
MAPK	Mitogen-activated protein kinase
MD2	Myeloid differentiation factor 2
MDA5	Melanoma differentiation-associated factor 5
MDP	Muramyl dipeptide
mRNA	Messenger RNA
MUC2	Mucin 2
MUC5AC	Mucin 5, subtypes A and C
MUC6	Mucin 6
MYD88	Myeloid differentiation primary response protein 88
NAIP	NLR family apoptosis inhibitor protein
NBD	Nucleotide-binding domain
NEMO	NF-κB essential modifier
NF-κB	Nuclear factor 'kappa-light-chain-enhancer' of activated B-cells
NLR	NOD-like receptor
NLRA	NOD-like receptor type A
NLRB	NOD-like receptor type B
NLRC	NOD-like receptor type C
NLRP	NOD-like receptor with pyrin domain
NLRX	NOD-like receptor with X effector domain
NOD	Nucleotide-binding oligomerization domain
NOX1	Nicotinamide adenine dinucleotide phosphate (NADPH) oxidase 1
OLFM4	Olfactomedin 4
P(dd)	Postnatal day
PAM3CSK4	Pam3CysSerLys4
PAMP	Pathogen-associated molecular pattern
PAS	Periodic acid-Schiff staining
PBS	Phosphate-buffered saline
PC	Principal component

PCA	Principal component analysis
PCR	Polymerase chain reaction
PSCs	Pluripotent stem cells
PYD	Pyrin domain
RT-qPCR	Quantitative reverse transcription PCR
PFA	Paraformaldehyde
PI	Proximal intestine
PRR	Pattern recognition receptor
PSC	Pluripotent stem cell
PYD	Pyrin domain
RA	Retinoic acid
REG4	Regenerating family member 4
RIG-I	Retinoic acid-inducible gene 1
RIPK2	Receptor-interacting serine-threonine kinase 2
RLR	RIG-I-like receptors
RNA	Ribonucleic acid
dsDNA	Double-stranded DNA
ssRNA	Single-stranded RNA
dsRNA	Double-stranded RNA
RNAPII	RNA polymerase II
ROS	Reactive oxygen species
RT	Reverse transcriptase
SBM	Supplemented basal medium
SD	Standard deviation
SHH	Sonic hedgehog
SI	Small intestine
SOX2	Sex determining region Y (SRY)-box transcription factor 2
STING	Stimulator of interferon genes
STMN1	Stathmin 1
SYK	Spleen tyrosine kinase
TA	Transient-amplifying cells
TAE	Tris-acetate-EDTA
TAK1	Transforming growth factor-β-activated kinase 1
TAM	Transient Abnormal Myelopoiesis
TANK	TRAF family member associated NF-κB activator
TBK1	TANK binding kinase 1
TEER	Transepithelial electrical resistance
TF	Transcription factor
TFF2	Trefoil factor 2
TGF	Transforming growth factor

TICAM	Toll-like receptor adaptor molecule
TIFA	TRAF-interacting protein with forkhead associated domain containing protein A
TIR	Toll-interleukin receptor
TIRAP	TIR domain containing adaptor protein
TLR	Toll-like receptor
TNF	Tumor necrosis factor
TRAF	TNF receptor associated factor
TRAM	Translocation associated membrane protein
TRIF	TIR domain containing adaptor inducing IFN-β
TROY	TNFRSF19 - tumor necrosis factor receptor superfamily, member 19
TSS	Transcription start site
UC	Ulcerative colitis
WENR	WNT, EGF, Noggin, R-spondin1
WT	Wild-type

Abstract

The epithelial layer of the gastrointestinal (GI) tract provides a barrier between the environment and the body. Dysfunction of the epithelium, including changes of the innate immune response facilitated by pattern recognition receptors (PRRs), plays a major role in the development of GI disorders. However, the organization of innate immune sensing, the expression and activity of PRRs and the factors contributing to such possible organization along the GI tract are unclear. In recent years, stem cell-derived organoids gained increasing attention as promising tissue models. Here, a biobank of human and murine organoids comprising three lines from each GI segment; corpus, pylorus, duodenum, jejunum, ileum, colon was generated. RNA sequencing of 42 lines confirmed the preservation of tissue identity and revealed an extensive organization of innate immune signaling components along the cephalocaudal axis, giving each segment a specific innate immune profile. Complementing the region-specific expression analysis, several PRRs in human and murine organoids showed region- and species-specific function. To investigate the factors contributing to the patterning of innate immunity in the GI tract, the impact of microbial components was analyzed using murine embryo-derived, never colonized gastric and proximal intestinal organoids. Transcriptional profiling of embryo-derived organoids showed that while expression of some PRRs may depend on environmental cues as expected, an unexpectedly large part of segment-specific expression of PRR signaling components is independent of prior contact with microbial products. Further, analysis of published RNA-seq data as well as *in vitro* experiments using directed differentiation of organoids into specific cell types showed that expression of innate immune gene also depended on cellular differentiation along the crypt-villus axis. This underlined the importance of cellular differentiation rather than contact to microbial compounds for expression of PRRs. Lastly, analysis of published datasets of RNA-seq and ATAC-seq after knockout of the intestinal transcription factor *Cdx2* demonstrated that *Cdx2* is likely important for the expression of *Nlrp6* and *Naip1* in the murine intestine. Future experiments have to support these preliminary findings. Taken together, the expression of a large part of epithelial innate immunity is developmentally defined and conserved in tissue-resident stem cells. The identification of mechanisms governing expression of genes related to immunity will provide further insights into the mechanisms that play a role in the progress of inflammatory diseases.

Zusammenfassung

Das Epithel des gastrointestinalen (GI) Traktes fungiert als Barriere zwischen der Umwelt und dem Körperinneren. Störungen des Epithels, darunter Veränderungen in der angeborenen Immunantwort, welche über „Pattern Recognition Receptors“(PRRs) ermöglicht wird, spielen eine bedeutende Rolle in der Entstehung gastrointestinaler Krankheiten. Auf welche Weise die angeborene Immunantwort im gastrointestinalen Trakt zwischen symbiotischen und schädlichen Mikroben unterscheidet, wie die Expression und Aktivität von PRRs organisiert sind, und die Faktoren die zu einer möglichen Organization beitragen sind bisher allerdings nicht bekannt. In den letzten Jahren haben aus Stammzellen gewonnene Organoide als vielversprechende Gewebemodelle steigende Aufmerksamkeit erregt. In dieser Arbeit wurde eine „Biobank“ humaner und muriner Organoide, jeweils bestehend aus 3 Linien jedes der gastrointestinalen Segmente Korpus, Pylorus, Duodenum, Jejunum, Ileum und Kolon generiert. Die RNA Sequenzierung von 42 Linien bestätigte den Erhalt der Gewebsidentität und zeigte eine umfangreiche Organization innerhalb der Signalkomponenten des angeborenen Immunsystems entlang der kraniokaudalen Achse, wodurch jedes Segment ein spezifisches Immunprofil erhält. Ergänzend zur regions-spezifischen Expressionsanalyse zeigten einige PRRs sowohl in humanen als auch in murinen Organoiden eine regions- und spezies-spezifische Funktion. Zur Untersuchung der Faktoren, die zur Strukturierung des angeborenen Immunsystems im GI-Trakt beitragen, wurde der Einfluss mikrobieller Komponenten untersucht. Hierfür wurden aus embryonalem Gewebe gewonnene Organoide des Magens und des proximalen Dünndarms verwendet, welche noch nicht mit dem Mikrobiom in Kontakt waren. Transkriptionsprofile embryonaler Organoide zeigten, dass die Expression einiger PRRs wie erwartet wahrscheinlich von Umweltfaktoren abhängt, dass ein unerwartet großer Anteil der segment-spezifischen Expression von Komponenten der PRR-induzierten Signalwege sich allerdings unabhängig vom Kontakt mit mikrobiellen Komponenten entwickelt. Des Weiteren zeigte die Analyse von bereits publizierten RNA Sequenzierungsdaten und *in vitro* Experimenten bei denen durch gezielte Differenzierung von Organoiden spezifische Zelltypen generiert wurden, dass die Expression von Genen des angeborenen Immunsystems auch von der zellulären Differenzierung entlang der Krypten-Zotten-Achse abhängt. Dies unterstreicht die Bedeutung von zellulärer Differenzierung für die Expression von

PRRs, anstelle des Kontaktes zu mikrobiellen Komponenten. Auch konnte durch die Analyse bereits publizierter RNA- und ATAC-Sequenzierungsdaten nach knockout des im Dünndarm exprimierten Transkriptionsfaktors *Cdx2* nachgewiesen werden, dass *Cdx2* mit hoher Wahrscheinlichkeit wichtig für die Expression von *Nlrp6* und *Naip1* im murinen Dünndarm ist. Diese Erkenntnisse müssen in zukünftigen Experimenten validiert werden. Zusammengenommen zeigen die Ergebnisse, dass die Expression eines Großteils der angeborenen Immunität entwicklungsbiologisch festgelegt und in gewebe-spezifischen Stammzellen konserviert ist. Die zukünftige Identifikation von Mechanismen die die Expression von zur Immunität zugehörigen Genen steuern, wird weitere Erkenntnisse über die Mechanismen die eine Rolle in der Entwicklung entzündlicher Erkrankungen spielen bringen.

1 Introduction

As one of the largest organs in the body the gastrointestinal (GI) tract, responsible for all digestive processes, spans the entire human torso. The GI lumen is colonized by a wide variety of commensals, symbionts and occasionally pathogens. The microbial colonization follows a gradient with less than 10^3 microbes/ml in the stomach, to 10^3-10^7 microbes/ml in the small intestine and 10^{11}-10^{12} microbes/ml in the colon [1, 2]. It is now well established that the microbiome is critically involved in a plethora of processes essential for maintaining intestinal homeostasis [3]. On the other hand, bacterial translocation from the gut lumen past the epithelial barrier can lead to a fatal inflammatory response of the whole body. In this context, loss of intestinal epithelial barrier function is considered the driver of multiple organ failure and sepsis [4]. Therefore, the GI tract has the unique and opposing tasks to establish peace at its borders, identify potentially harmful invaders, yet enable adequate exchange of signals and uptake of nutrients [5–8].

A central player in this balancing act is the thin cellular lining that separates the body from the microbiota: the epithelial cells. They provide the physical mucus barrier, secrete the chemical defense proteins and are equipped with innate immune receptors to recognize microorganisms, enabling them to sense invading pathogens and subsequently recruit professional immune cells to help resolve the infection. These innate immune responses must balance the need for providing protection from potentially harmful pathogens with being able to tolerate the exposure to the diverse luminal microbiome. Hence, the innate immune signaling is not only expected to have a significant impact on pathogen sensing but also on tissue homeostasis and inflammatory diseases, including acute gastroenteritis, gastritis and inflammatory bowel diseases (IBDs) [7–10]. Although each of these diseases displays completely different pathogeneses, they share a critical feature: the interface between the environment and the body is disturbed significantly. However, there is an ongoing debate whether changes of the intestinal epithelial layer, in particular those affecting its innate immune function, are a cause or a consequence of the diseases mentioned above.

A major obstacle for gaining a better understanding of the disease-specific contribution of GI epithelial cells has been the lack of appropriate experimental models. However, with the recent development of the organoid technology that enables the generation of primary epithelial cultures from epithelial stem cells, a major step to enable a better understanding of the specific role of epithelium-dependent contributions to GI disorders has been taken.

This thesis analyzes the epithelial innate immunity of the murine and human GI tract using organoids as a model. Here, first, the GI tract of both species will be introduced together with the organoid technology, followed by the developmental patterning of the GI epithelium and finally the innate immune signaling pathways of the GI tract.

1.1 Gastrointestinal tract

The GI tract is required for the digestion of food, and spans from the oral cavity via the esophagus, stomach, small intestine and large intestine to the anus. It is required for the digestion of food by subjecting it to mechanical and chemical treatments to extract and absorb the maximum amount of nutrients and water into the body. Although the GI tract comprises anatomically defined segments with vastly distinct physical functions [11], the basic structural organization is comparable between the segments and the organ walls are made up of four layers (**Figure 1**). The outermost layer, serosa, is followed by muscularis externa which contains two layers of smooth muscle, the submucosa, this in turn is followed by the mucosa which includes a connective tissue layer called the lamina propria (collectively with submucosa, called the mesenchyme) and the epithelial cell layer. While the outer layers are rather similar along the GI tract, the innermost layers show distinct features since they are adapted to the specific functions of each segment and serve as a barrier that separates the lumen of the GI tract from the rest of the organism [12]. The tightly locked epithelial layer lines the GI tract and comprises diverse cell types whose fates are regulated by signals from both the epithelium and the underlying mesenchyme [13].

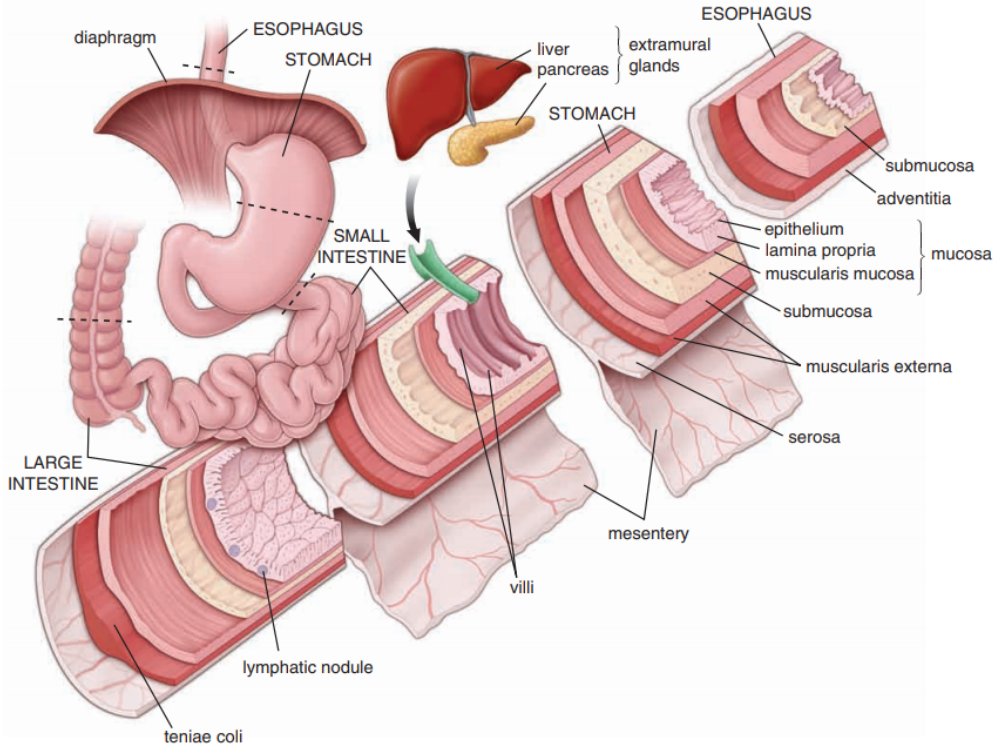


Figure 1: Anatomy of the gastrointestinal tract. Wall structure of esophagus, stomach, small intestine and large intestine (adapted with permission from reference [12]; copyright 2011 Wolters Kluwer Health Inc.).

Of the different segments, the stomach is responsible for storing the food, and it contributes to digestion both chemically and physically by secreting acid and enzymes while contracting to break down the particles. The murine stomach is divided into three anatomical regions, the non-glandular forestomach lined with squamous epithelium, the glandular oxyntic body (corpus) region and the antrum. In contrast to mice, the human stomach does not have a forestomach and is divided into four glandular regions called proximal cardia, fundus, body (corpus) and the antrum (pylorus) [14]. At the lowest point of the stomach, where it transitions into the small intestine, the pyloric sphincter is located and delivers small portions of pre-digested food into the small intestine.

The intestine has the role of digesting the food, absorbing nutrients and water, and maintaining the barrier between the body and the microbiota. It is separated into two main parts, the small intestine and the large intestine. The small intestine comprises the duodenum, jejunum and ileum, which mainly provide the digestion and absorption of the nutrients [15]. The large intestine comprises the cecum, colon and rectum, and its primary role is the reabsorption of water and elimination of the non-digestible food parts.

1.1.1 Anatomy of the GI epithelium

1.1.1.1 Stomach

The gastric mucosa of the stomach is the source of the mucus layer, digestive enzymes and gastric acid. The primary function of the stomach is to start the digestion and direct the food to the small intestine, and here it is not necessary for nutrients to have to reach the epithelium yet. Therefore, the gastric epithelial layer heavily invests in an acidic environment and a thick mucus barrier to protect the epithelial cells from said environment. The gastric mucosa contains numerous tiny invaginated gastric glands with the underlying lamina propria. These glands are located below the luminal epithelium, which deliver the secreted materials of the glands into the lumen. Each gland has distinct regions called gastric pit, isthmus, neck and base (**Figure 2**).

Within the stomach, the glands of the corpus and antrum show differences. The main cell types of the corpus units are the surface mucus (pit) cells, parietal cells, chief cells, isthmus cells, mucus neck cells, endocrine cells and very rare tuft cells, all with different functions. Surface pit cells are located around the gastric pits and line the entire gastric mucosa (**Figure 2**). They are highly differentiated and secrete the mucin 5AC (MUC5AC), which contributes to the formation of the thick mucus layer. Parietal cells, which are heavily concentrated in the neck region, produce and secrete hydrochloric acid (HCl). Chief cells are found at the base of the corpus glands and secrete a peptide-degrading enzyme called pepsinogen. Mucus neck cells produce various mucins, such as mucin 6 (MUC6) and are found among the parietal cells [16]. Endocrine cells of corpus glands produce ghrelin, which is characteristic for the corpus region [17]. Tuft cells are sensory secretory cells rarely found in the stomach. They have a characteristic morphology with a prominent apical extension and form direct synapses with interneurons in the gastric wall [18].

In the antrum, the units are primarily composed of surface pit cells and gland mucus cells. The gland mucus cells produce low levels of zymogenic proteins along with specific mucins such as MUC6 [19] as well as factors for protection of epithelial layer, such as trefoil factor 2 (TFF2) [20]. Although a low number of parietal cells can be observed in human antral units, they are not found in the murine antral units [17]. Endocrine cells are interspersed between the mucus cells which are named according to the type of the hormone they produce, such as G-cells secreting gastrin (**Figure 2**).

Maintenance of the gastric epithelium

In the stomach, adult stem cells reside within the gastric units and constantly proliferate. They are long-lived and self-renewing cells which produce further proliferating undifferentiated progenitor cells. The differentiation of the cell populations occurs along the glands. Cells differentiate at the lower parts of the glands and migrate to the top, where they terminally differentiate into surface pit cells before they are extruded into the gastric lumen [21]. While the surface pit cells are shed into the gastric lumen in a few hours or days, some cells such as parietal cells can survive for weeks or months [22].

The location of the stem cell compartment in corpus and antrum shows differences. The antral glands have a stem cell population in the gland base marked by the stem cell markers leucine-rich repeat-containing G-protein coupled receptors 5 (LGR5) and tumor necrosis factor receptor superfamily member 19 (*Tnfrsf19*) which encodes the protein TROY [23–25]. However, in the corpus, this cell population in the gland base is mostly quiescent and acts as a reserve stem cell population upon injury instead of contributing to the turnover during homeostasis [24, 25]. In the corpus glands, additional to the base region, a constantly proliferating cell population was found in the isthmus region [26]. One study confirmed only recently that two independent stem cell populations are located in the corpus glands. Of those, one population, located at the gland base, consists of slow-cycling cells marked by TROY and LGR5. The other is a rapidly cycling stem cell population found in the isthmus region of the corpus glands, which expresses the proliferation markers stathmin 1 (*Stmn1)* and *Ki67*, and is responsible for the turnover of the pit-isthmus-neck region of the corpus gland [27].

Evolutionarily conserved signaling pathways play a crucial role during gastro-intestinal development and maintenance. A number of extrinsic factors, particularly those provided by the local environment surrounding the stem cells, control the fate of the stem cells. For stem cell maintenance, hedgehog (Hh) signaling has a major role in the gastric stem cell niche by regulating some niche factors, such as bone morphogenic protein (BMP), Notch and fibroblast growth factor (FGF)10 [28] (**Figure 2**). BMP signaling pathway regulates gastric epithelial cell proliferation and differentiation. BMPs are regulatory molecules secreted by stromal cells and they

promote differentiation of the stem cells. It has been demonstrated that inhibition of BMP signaling induces the proliferation of LGR5+ stem cells [29]. Noggin is an antagonist of BMP signaling which acts as an essential factor for maintenance of stem cell niche, by inducing extracellular signal-regulated kinase (ERK) activation [30].

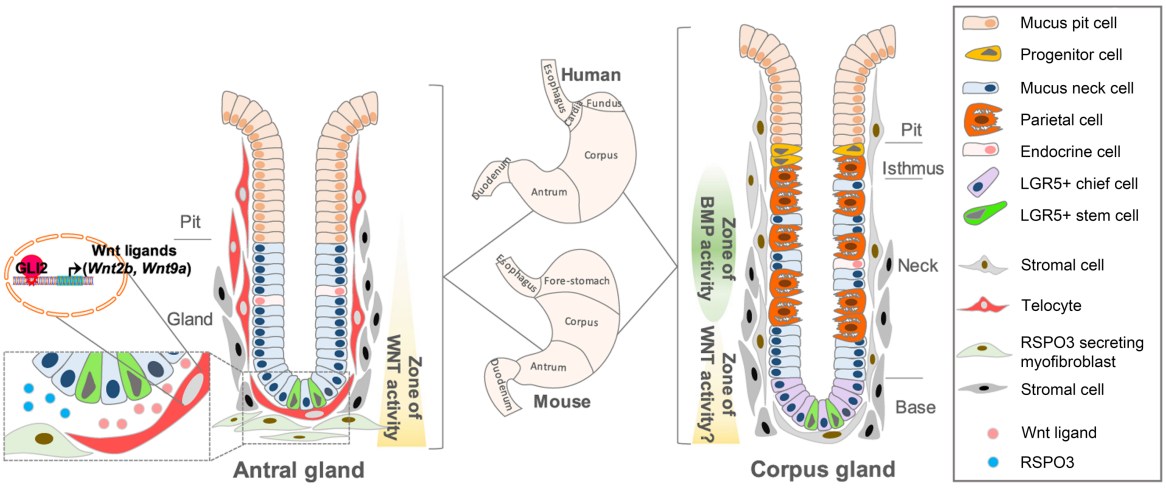


Figure 2: Morphology of gastric epithelial layer and distribution of the niche factors. The cellular architecture of antral (left) and corpus (right) glands, depicting different distribution of epithelial cell types. Corpus glands are comprised parietal cells, chief cells, pit mucus and gland mucus cells, endocrine cells, undifferentiated isthmus cells and tuft cells. Differently than the corpus glands, antral glands do not contain the parietal cells and have LGR5+ stem cells within the gland base. Telocytes and myofibroblasts found underneath them provide the WNT ligands and R-spondin3 to maintain the stem cell activity. In the corpus (right) green circle and yellow tringle indicate the gradients of niche factors. Parietal cells which are found at the neck region of corpus glands secrete Hh and upregulate the mesenchymal BMP expression which leads to the formation of a BMP gradient. The WNT gradient in the corpus glands is not clearly known, however cells at the gland bases are WNT-responsive suggesting a high level of WNT signaling (adapted with permission from reference [13], copyright 2020 Elsevier).

Notch signaling is another pathway important for stem cell maintenance and proliferation. It is especially active in the isthmus region of corpus glands [31] and base of antral glands, below the antral stem cells [32]. When it is inhibited, proliferation in antral and corpus glands diminishes and the number of differentiated cells increases [31]. FGF10 is another important growth factor derived from the mesenchyme. When the pathway is impaired, this results in reduced proliferation and defective gland formation [33], while its overexpression leads to glandular hyperplasia [34]. WNT signaling pathway is required for the expression of the stem cell markers LGR5 in the antral glands [23]. R-spondin3 from stromal cells located beneath the stem cell niche induces the activation of WNT signaling, maintenance and proliferation of stem cells [35]. Thereby, activation of WNT signaling pathway leads to an increased number of progenitor cells in the antrum. In corpus glands

TROY+ chief cells at the gland base express WNT target genes implying the presence of a WNT source around the gland bottom for maintaining the undifferentiated state of the progenitor cells [25]. Finally, epidermal growth factor (EGF) plays an important role in normal proliferation and growth of the mucosa [36].

1.1.1.2 Small intestine and colon

In contrast to the gastric mucosa which is optimized to protect the epithelial layer from the acidic environment, the mucosa in the small intestine is optimized for surface maximization to enable efficient uptake of nutrients. In humans, the small intestine contains permanent circular folds called the plicae. These structures are most prominent in the distal duodenum and proximal jejunum, where most of the digestion and absorption take place and they decrease in number with progression through the ileum [37] (**Figure 3, left**). Protruding from the plicae are densely packed intestinal villi covering the entire small intestine to further increase the surface area. Each villus is a long projection covered by epithelium where the vascularized lamina propria forms the core. In murine mucosa, the villus heights are double those of human, providing increased surface area to murine small intestine, to make up for the lack of plicae [38]. Correlative to the rate of digestion and absorption, the villus length and number progressively decreases from duodenum to the ileum. The villi in the duodenum and jejunum are larger and denser, while those in the ileum are smaller and fewer [39]. Finally, the bases of the villi are surrounded by crypts of Lieberkühn, where primarily secretory, stem and proliferative cells are located. The architecture of the epithelium differs between small intestine and colon, related to their distinct functions. The colonic epithelium is specialized for the absorption of water and ions, and production of mucus barrier. It lacks both plicae and villi which are characteristic for small intestine.

The epithelial cells in the small intestine can be divided into two broad categories: absorptive cells and secretory cells. The majority of the epithelial cells covering the small intestinal villi are well differentiated and polarized absorptive enterocytes which also produce digestive enzymes. The apical side of these columnar epithelial cells is densely packed with finger-like projections called microvilli, forming the brush border to further increase the absorptive surface area of the small intestine [40]. Goblet cells, enteroendocrine cells, tuft cells and cup cells are interspersed among the

enterocytes [41] in both villi and crypts. Goblet cells produce and secrete glycosylated mucins crucial for formation of a mucus barrier on the epithelial layer [42]. With the highest ratio of goblet cells, the ileum has the thickest mucus layer in the small intestine and the lowest rate of digestion and absorption [11, 43]. Enteroendocrine cells are found in both crypts and villi and they secrete various hormones and peptides to regulate the gut physiology. The most distinct feature of enteroendocrine cells is that the secretory granules are located mainly at the basal region of the cell for the produced hormones to be discharged on the basal side of the cell from where they diffuse and pass into the capillary bed underneath. Chemosensory tuft cells respond to various stimuli such as hypoxia and infections. For example, the number of the tuft and goblet cells rapidly increases upon helminth infection, depending on the secretion of cytokines [44].

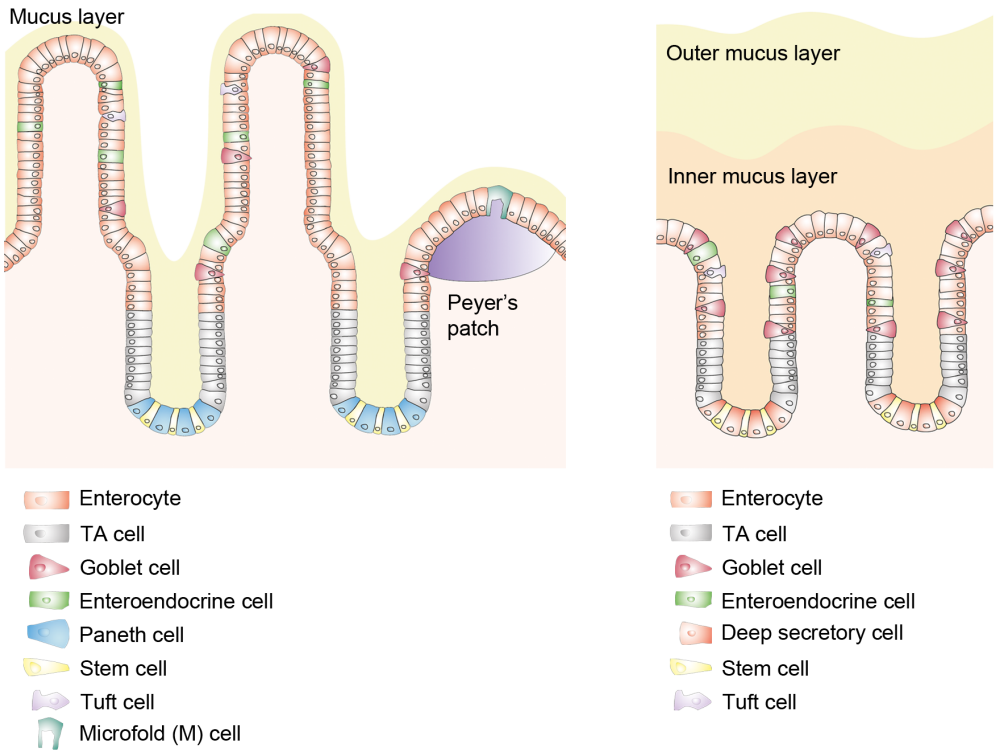


Figure 3: Cellular architecture of small intestinal and colonic epithelial layers. Small intestinal epithelial layer (left) consists of crypts and villi. Stem cells at the crypt base produce more stem cells and also transit-amplifying (TA) cells which migrate upwards toward the villus tips. Paneth cells are also found at crypt bases and they support the stem cell niche by providing necessary signals. Upwards, differentiated cells of absorptive enterocytes and secretory goblet, enteroendocrine and tuft cells reside. M cells are located above the Peyer's patches where they contribute to the crosstalk between outer environment and immune cells. Colonic epithelial layer (right) is similar to the small intestinal one except that it does not have villi and Paneth cells compared to small intestine. Deep secretory cells found at the crypt base secrete antimicrobial peptides, resembling the Paneth cells.

The colonic crypts are similar to those of the small intestine with absorptive enterocytes, goblet cells, enteroendocrine cells, tuft cells and proliferative stem cells at the crypt bases [45] (**Figure 3, right**). A large number of goblet cells are found over the whole length of the crypt in the proximal colon of mice, but at the base of the crypts in the distal colon and the rectum, the number decreases. The goblet cells are predominant in human colon as well, however with the ratio of goblet cells to enterocytes increasing from the proximal toward the distal regions [38]. Mucin 2 (MUC2) is the main mucin of the colonic mucus system which has two major mucus layers. The outer mucus layer is non-attached and less dense, whereas the inner mucus layer is attached and denser, forming netlike structures. While the outer layer is highly colonized by the commensals, the denser inner layer is impenetrable for those, thereby keeping the epithelium sterile [46]. Enteroendocrine cells can also be scattered in the epithelial layer and are identified by their small basally located granules.

The small intestinal crypt bases are populated exclusively with Paneth and stem cells. Paneth cells are pyramidal cells with granules containing antimicrobial molecules to prevent bacterial infiltration, thereby guarding the epithelium and keeping the crypts sterile [47]. Unlike enteroendocrine cells, these granules face towards the lumen of the crypt. The Paneth cells are intermingled with the intestinal stem cells at the base of the crypt, maintaining the stem cell compartment by providing signals [48]. The columnar base cells (CBCs) are the adult stem cells which also produce further proliferating, undifferentiated progenitor cells, called transit-amplifying (TA) cells. These cells move from the crypt bases and proliferate within a region called transit-amplifying zone, located in the lower half of the crypt. They form the amplifying compartment of the crypts which are still capable of cell division and usually committed to fully differentiate as they continue to migrate luminally. The colonic crypts are deeper and lack the Paneth cells [49]. There deep crypt secretory cells replace Paneth cells in supporting the stem cell niche [50] and by producing antimicrobial regenerating family member 4 (REG4) [51].

Beneath the epithelium, the mucosa contains dense networks of blood and lymphatic capillaries to transport the absorbed nutrients and produced hormones. Moreover, the lamina propria houses the peripheral immune cells responsible for guarding

against any ingested pathogens which might have survived the hostile environment in the stomach. These cells constitute the gut-associated lymphatic tissues (GALTs). Peyer's patches, which are large aggregates of lymphoid tissue beneath the small intestinal epithelium, are components of GALTs (**Figure 3**). They are diffusely distributed and are most prominent in the human ileum and murine jejunum [38]. The overlaying cell layer on a Peyer's patch is called follicle-associated epithelium where the villous architecture is less developed. Microfold (M) cells are located in the follicle-associated epithelium and have a unique morphology with reduced microvilli structures. Their role is to transport antigens in the gut lumen across the epithelium to the underlying lymphoid tissue for regulation of immune responses [52]. In the colon, GALT is more extensively developed with large lymphoid follicles. This extensive development of the immune system in the colon reflects the large number and variety of the microbiota and the end products of the metabolism.

Maintenance of small intestinal and colonic epithelia

The intestinal epithelial layer is maintained by rapidly proliferating stem cells at the base of the crypts which generate TA cells. TA cells divide multiple times before maturing into enterocytes, while secretory cells mature directly from stem cells [53]. Cells above the amplification region move conveyor belt-like towards the surface, where they are shed into the lumen after a lifetime of only 3-5 days [12]. Conversely, cells below have a relatively long lifetime. For example, Paneth cells residing at the base of the crypt have a lifespan of about 1-2 months [54].

Intestinal stem cell activity is maintained by the continuous communication with the intestinal mesenchyme and multiple epithelial cells. Canonical WNT/R-spondin, Notch and EGF signaling pathways promote the proliferation, while BMPs favor the differentiation [55–57] (**Figure 4**). In the small intestine, Paneth cells located at the crypt bases support the stemness by producing stem cell factors such as WNT3 and EGF [58, 59]. They also present Notch ligands such as delta-like protein (DLL) 1 and DLL4 to the stem cells which trigger the activation of NOTCH1 and NOTCH2 on the stem cells. Notch signaling happens only through cell-cell contact and leads to the signal transmission between Paneth and stem cells. Notch inhibition leads to terminal differentiation of stem cells into secretory cells. In the colon, although deep secretory cells fulfill the functions of Paneth cells in secreting EGF and Notch ligands, they do

not produce WNT ligands to support the stem cells. Therefore, in the colon WNT ligand production to support the maintenance of the stem cells is done solely by the subepithelial mesenchymal niche populations [60, 61]. Likewise, these populations secrete R-spondins, BMP ligands and BMP inhibitors for the precise control of stem cell maintenance and epithelial differentiation in both small intestine and colon [60, 62]. Peri-cryptal stromal cells create a WNT/R-spondin enriched and BMP-poor microenvironment by secreting WNT/R-spondin ligands and BMP inhibitors, such as noggin and gremlin, while the villus lamina propria creates the reverse by producing BMPs [55, 62, 63]. Hence, because of the local production of the mediators and limited diffusion capacity, opposing WNT and BMP gradients are established along the crypt-villus axis which control the balance between self-renewal and lineage differentiation [64, 65]. While the cells migrate upwards from the crypt bases, they are exposed to decreasing levels of WNT and increasing levels of BMP. This gradient provides the niche for the stem cells to maintain their stem cell identity while allowing the differentiation of cells along the crypt-villus axis.

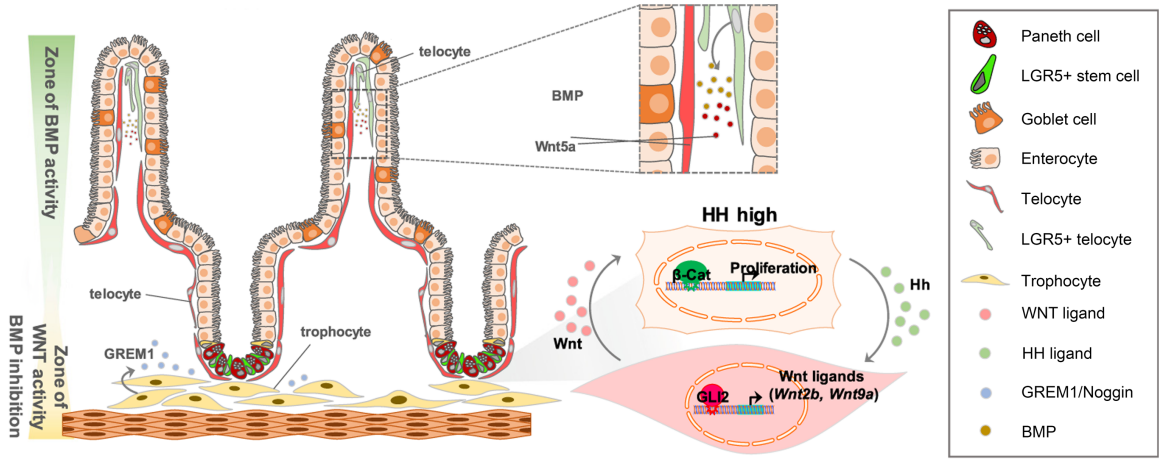


Figure 4: Distribution of niche factors in small intestinal epithelial layer. Peri-cryptal mesenchymal cells such as trophocytes and telocytes provide the niche signals to epithelial cells. Trophocytes are located at crypt bases, close to the stem cells and express the BMP inhibitor gremlin 1 (GREM1) to support the stem cell niche. LGR5+ telocytes found at the villus tips secrete BMP to drive differentiation of the enterocytes (adapted with permission from reference [13], copyright 2020 Elsevier).

1.1.2 Regional identity and patterning in the gut

The general GI organization is established during development. During the late gastrula stage, the definitive endoderm develops into a primitive gut tube, surrounded by mesoderm. This mesodermal tissue, later mesenchyme, controls the anterior-posterior and proximal-distal patterning by regulating the FGF, WNT, BMP,

retinoic acid (RA), Hedgehog and Notch pathways in a time-dependent manner. These morphogens form concentration gradients to instruct the gut tube to regionalize into foregut, midgut and hindgut. While the foregut develops into esophagus and stomach of the GI tract, the midgut develops into the small intestine and the hindgut into the large intestine [66, 67] (**Figure 5**).

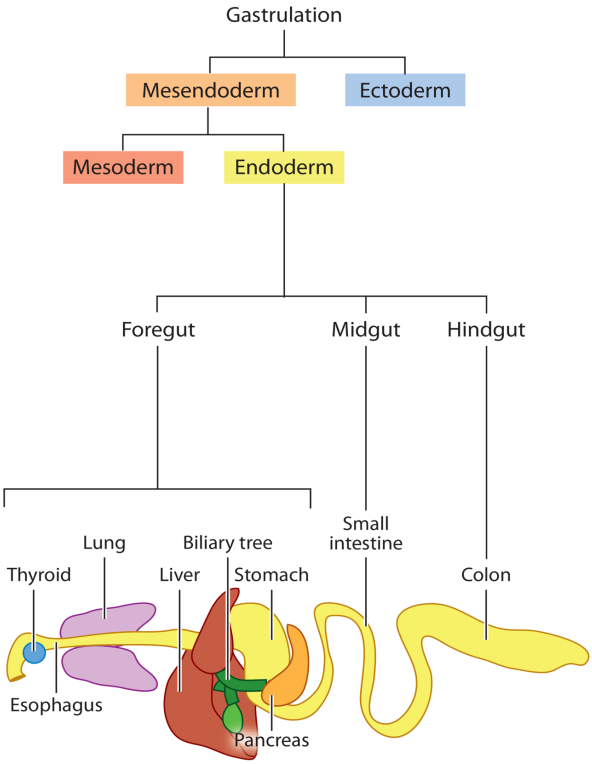


Figure 5: Overview of gastrointestinal development. After the formation of three main germ layers, endoderm arranges into foregut, midgut and hindgut. From these sections, esophagus and stomach develop from foregut, while proximal small intestine develops from midgut and large intestine from hindgut (adapted with permission from reference [67]; copyright 2009 Annual Reviews).

The boundaries between the organs are established during developmental transitions of the primitive gut tube. Signals from surrounding mesodermal tissues determine the regional identity and drive the anterior-posterior patterning by expression of region-specific transcription factors. Active WNT signaling, FGF10 and BMP originating from the mesoderm drive intestinal patterning of the mid- and hindgut characterized by the expression of the transcription factor caudal type homeobox 2 (CDX2) [68, 69] (**Figure 6**). CDX2 expressed in these tissues functions as a master regulator of intestinal epithelial development and regionalization of the GI epithelium [70–72]. During the early stages of development, CDX2, when ectopically expressed, can posteriorize the foregut and induce an intestinal-like identity. Loss of

CDX2 in the early endoderm converts the posterior intestinal epithelium into esophageal squamous type [70]. Conversely, its loss later in development only leads to partial anteriorization, with expression of gastric epithelial genes in the intestine [71, 73].

For proper regionalization and development of the stomach, WNT signaling needs to be minimized in the foregut, since it would lead to expression of CDX2 at this stage of development [68]. BarH-like homeobox 1 (BARX1) expressed by the mesenchyme of the developing stomach induces the secretion of WNT antagonists, thereby locally inhibiting the WNT pathway in the posterior foregut [74]. Additionally, sex determining region Y (SRY)-box 2 (SOX2), which is a transcription factor directing the development of stratified epithelium, is expressed in the early foregut [69, 75]. At early stage of the development, the BMP needed for the hindgut would lead to SOX2 inhibition in the foregut. Therefore, to keep the regional identity in the foregut, epithelial cells secrete BMP antagonists to inhibit the pathway [76]. Thus, the stomach-intestine boundary is determined by antagonism between the intestine-promoting WNT/CDX2 and the stomach-promoting BARX1/SOX2. This antagonism persists throughout the lifetime of the organism.

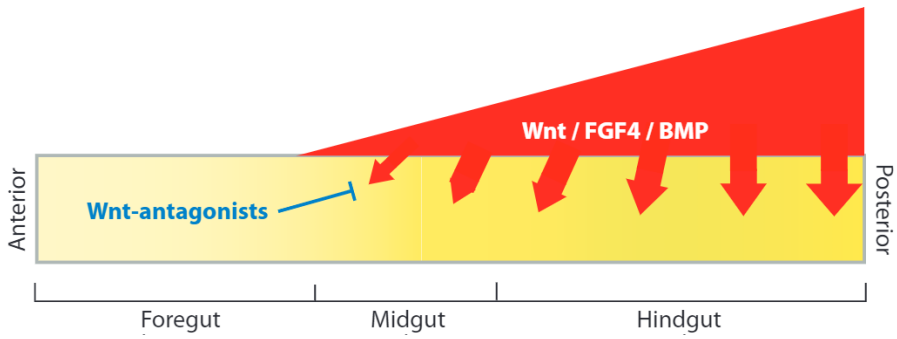


Figure 6: Signaling pathways responsible of regionalization in GI tract. The patterning of anterior-posterior regions. Differential WNT, FGF and BMP signaling results into anterior-posterior axis patterns which leads to formation of foregut, midgut and hindgut. The inhibition of WNT signaling is crucial to maintain foregut identity and further development (adapted with permission from reference [67]; copyright 2009 Annual Reviews).

Maintenance of regional identity is also observed in stem cells taken from adult tissue. Therefore, once defined, the regional identity remains encoded in the adult stem cells, which constantly regenerate the tissue [77]. For example, intestinal stem cells already express low levels of intestinal markers such as villin, mucin 2 and lysozyme [78], while the gastric stem cells express the gastric intrinsic factor (GIF) [23]. Taken together, regional differences in the GI tract are established during

development and are maintained in adult stem cells as part of the cell identity during adulthood.

1.1.3 Organoids to model the GI epithelial layer

Organoids are defined as non-transformed stem cell-derived, three-dimensional (3D) cell cultures that have self-organizing capacity and retain some of the function of the original organ e.g., secretion, filtration, absorption or contraction (**Figure 7**). Organoids can be grown either from tissue-resident adult stem cells (ASCs) or from pluripotent stem cells (PSCs). Together, the two types of organoids cover an extensive repertoire of organs that can be mimicked [64, 79, 80]. Both technologies have their individual advantages. For example, cultures of ASC-derived organoids have a tremendous expansion potential and are relatively homogeneous, whereas PSC-derived organoids are more complex in the sense that they combine cells of very different developmental origin (e.g., epithelial and mesenchymal cells). PSC-derived organoids allow the analysis of the developmental steps but may not reach the full level of differentiation into an epithelial layer as found *in vivo* [81].

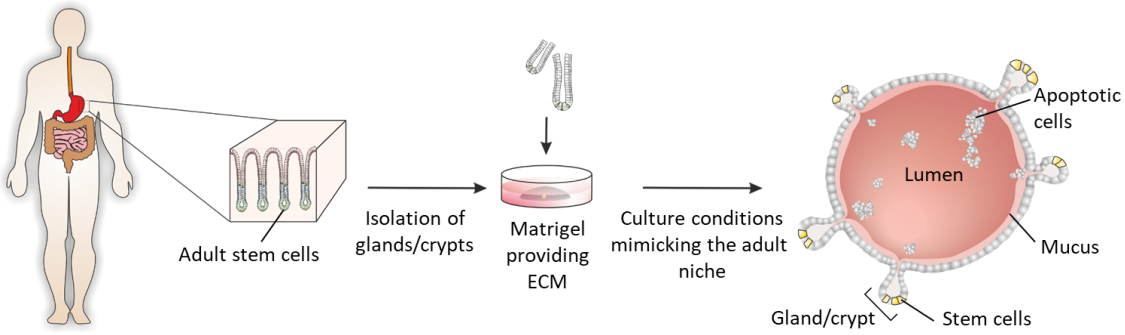


Figure 7: Generation of organoids. Adult stem cells can be isolated from tissue sections of the GI tract (exemplarily shown for human tissue here) and cultured *in vitro*. When the respective culture conditions are applied and the cells are grown in an extracellular matrix such as Matrigel, they form 3D structures called organoids. These organoids contain the majority of the cell types which are found in the original organ and additionally mimic the organization of the inner structure of the organ.

Murine intestinal organoids are the prototype of ASC-derived organoids [82]. They contain stem cells as well as differentiated cells. Moreover, the cells self-organize into different domains that harbor the stem cells - representing the crypts of the tissue, or the differentiated cells, such as enterocytes and goblet cells representing the villi of the tissue [82]. The GI organoids derived from ASCs retain this tissue identity and harbor cell types of the intestine or stomach [23, 82, 83]. The intrinsically programmed regional identity of the ASCs is not dependent on the niche factors provided by the

environment. For example, organoids grown from the different segments of the small intestine, express location-specific genes and show functional properties of the gut segment the organoids have been generated from despite being grown in the same culture conditions [81, 84]. The expression profile of intestinal organoids is also maintained when they are grown in gastric medium, which contains several common niche factors (R-spondin, Noggin, EGF), but also gastric-specific factors such as FGF10 [73]. The strongest evidence for the intrinsically programmed location specificity is provided by a transplantation experiment: Murine small intestinal organoids transplanted into the colon grow into patches of new small intestinal tissue at the transplantation site, marked by high CDX2 expression and even the presence of Paneth cells, a cell type that is restricted to the small intestine [85].

The GI epithelial cells are highly polarized with apical side facing the lumen of the gut with its microbiota, and the basolateral side facing the tissue free of the luminal content. Organoids now allow direct functional testing of side-specific immune responses, since the cellular polarization is retained in organoids. Under standard conditions when organoids are grown in an extracellular matrix, the apical side faces the lumen of the organoid and the basal side faces the extracellular matrix [23, 61, 82, 86]. When cells from organoids are seeded onto standard cell culture surfaces, such as culture dishes or transwells, the apical side faces the lumen of the well [87–90].

Regional identity is already encoded in the tissue-resident stem cells of the embryo to a certain extent, which is visible in the deoxyribonucleic acid (DNA) methylation profile [81]. However, while DNA methylation profiles of adult-derived organoids are stable over many weeks of culture, the DNA methylation profiles of fetal organoids show dynamic changes, indicating *in vitro* maturation [81] which is also found on the transcriptome level [81, 91]. Therefore, ASC-derived organoids generated from fetal tissues may also allow the study of maturation of fetal epithelia, since they age in culture [81, 92].

Taken together, organoids allow for the study of regional differences, their establishment during embryogenesis and their maintenance in adulthood. In addition, they also open up the possibility to study the extent of the regional identity.

1.2 Innate immune barrier

The healthy GI tract maintains its homeostasis by establishing the necessary contact with the lumen for digestion and nutrition, concurrent with the safe co-existence with microorganisms. This is apparent in the tissue architecture and cellular dynamics of the epithelial layer: surface cells are constantly exposed to the luminal content for nutrient uptake, therefore also open to possible infection by pathogens, but this is mitigated by stem cell proliferation which ensures constant renewal of these surface cells. The host defenses include further barriers, such as the physical barrier maintained by tightly interlocked epithelial cells and the mucus barrier. Furthermore, it is protected by chemical defenses, such as the antibacterial proteins, and by immunological defenses, provided by the innate and the adaptive immune systems. A key function of the innate immune system is to recognize microbial stimuli to mount quick defensive responses. This interaction is mediated by evolutionarily preserved receptors called pattern recognition receptors (PRRs).

1.2.1 Pattern recognition receptors

PRRs recognize microbe-associated molecular patterns (MAMPs) expressed by microorganisms but not by the host cells, and damage-associated molecular patterns (DAMPs) which are endogenous danger signals released from damaged tissues. Currently, based on their molecular structures and functions, the major classes of PRRs are toll-like receptors (TLRs), nucleotide-binding oligomerization domain (NOD)-like receptors (NLRs), C-type lectin receptors (CLRs), retinoic acid-inducible gene-I (RIG-I)-like receptors (RLRs), as well as several putative sensors such as absent in melanoma 2 (AIM2) and alpha-kinase 1 (ALPK1), cyclic guanosine monophosphate (GMP)-adenosine monophosphate (AMP) synthase (cGAS) and stimulator of interferon genes (STING) (**Figure 8**) [93, 94].

PRRs trigger innate immune responses in response to MAMP or DAMP detection. While recognizing different patterns, many of them share a common central signaling pathway, the nuclear factor-κB (NF-κB) pathway and interferon regulatory factors (IRFs) leading to production of proinflammatory cytokines, or formation of signaling complexes called inflammasomes. Through the coordination of several pathways, the type, magnitude and duration of the resulting host response are regulated to control the invading microorganisms and recruit professional immune cells [5, 7, 8, 95–97].

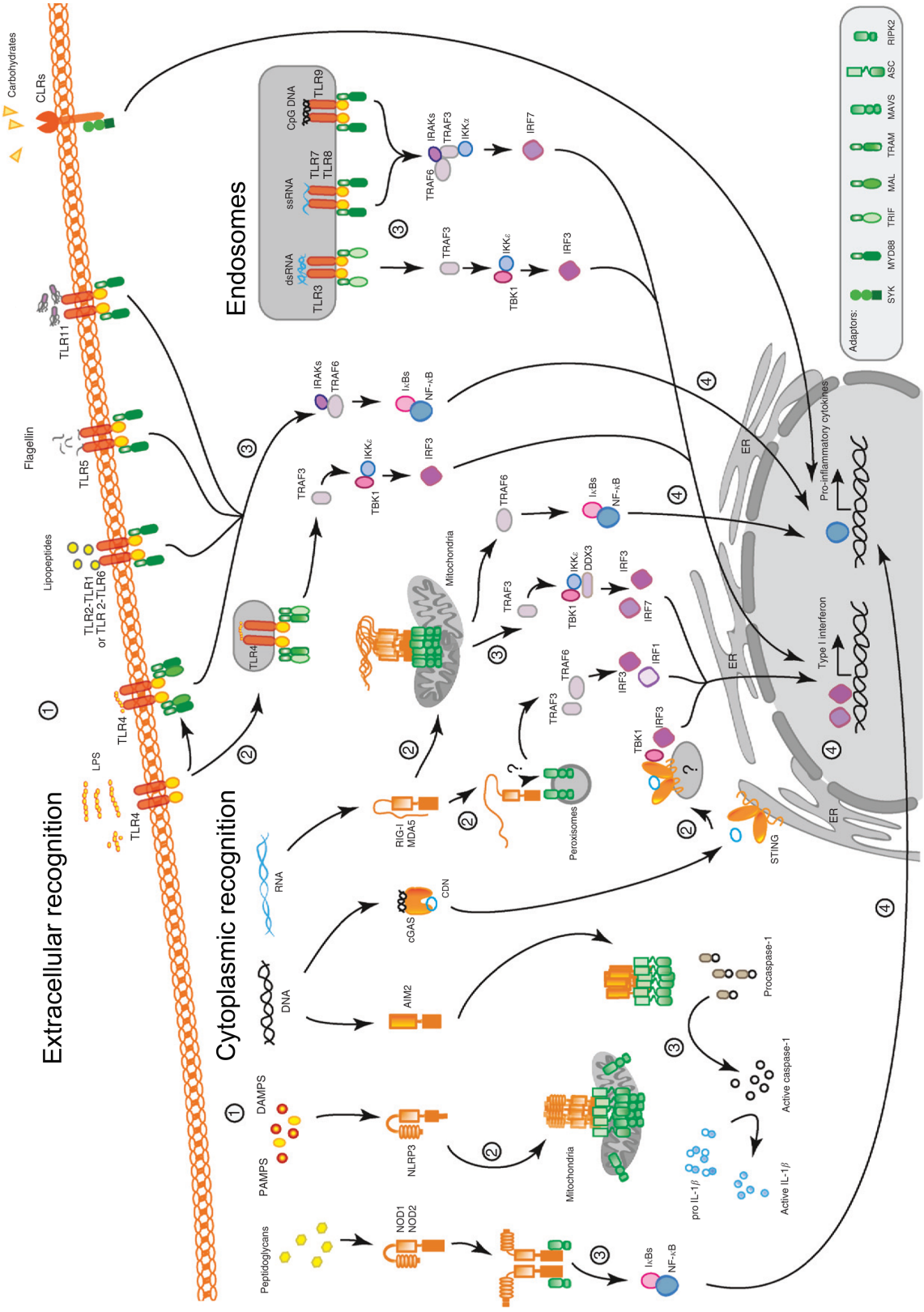


Figure 8: Overview of the major innate PRR signaling pathways triggered by extracellular and intracellular PRRs. MAMPs and DAMPs are recognized by the PRRs (1) and with the help of adaptor proteins (2) they activate the downstream mediators (3) and transcription factors to activate the immune response (4) (adapted with permission from reference [98]; copyright 2016 Elsevier.)

1.2.1.1 Toll-like receptors

TLRs are transmembrane proteins that contain three structural domains: a ligand-binding domain with leucine-rich repeats (LRRs) motif, a transmembrane segment and a cytosolic toll/IL-1 receptor (TIR) domain for signal transmission [99]. Currently, the mammalian TLR family includes 13 members. 10 functional TLRs have been characterized in humans (TLR1-TLR10) and 12 receptors were identified in mice (TLR1-TLR9, TLR11-TLR13). Each TLR recognizes a distinct MAMP derived from microbes (**Figure 8**).

TLR1 which is functionally associated with TLR2 by MAMP-mediated dimerization [100] recognizes triacyl lipopeptides derived from bacteria [101]. TLR2 also dimerizes with TLR6, for the recognition of bacterial diacyl lipopeptides [102]. TLR3 is involved in detection of double-stranded (ds)ribonucleic acid (RNA) produced by viruses during their replication [103]. TLR4 detects the lipid A contained in lipopoly-saccharide (LPS) localized in the outer membrane of Gram-negative bacteria. TLR5 recognizes bacterial flagellin [104, 105]. TLR7 and 8 recognize single-stranded (ss)RNA [106, 107]. TLR9 recognizes unmethylated 5'-C-phosphate-G-3' (CpG) containing ssDNA [108]. A ligand for TLR10, which is a pseudogene in mice, has not been clearly identified yet. Murine TLR11 has been demonstrated to detect flagellin from *Salmonella,* while TLR11 and 12 both recognize profilin derived from *Toxoplasma gondii* [109]. Lastly, TLR13 is associated with the recognition of bacterial 23S ribosomal RNA [110]. These receptors can also be grouped into those that are localized on the cell surface (TLR1, 2, 4, 5, 6, 10) and those that are found inside the cellular compartments called endosomes (TLR3, 7, 8, 9, 11, 12, 13). However, TLR4 can also be found intracellularly [111] and TLR9 is additionally expressed on the cell surface of intestinal epithelial cells [112]. Activation of TLRs leads to production of proinflammatory cytokines and chemokines, such as tumor necrosis factor alpha (TNF-α), IL-6 and IL-8 which recruit phagocytes to the site of infection. In addition to the proinflammatory cytokines, stimulation of intracellular TLRs, TLR3, 7, 8 and 9 also induces production of type 1 interferons, which are crucial for innate antiviral immunity.

When the ligand interacts with the receptor and induces the hetero- or homodimerization, the intracellular TIR domains of TLRs (except TLR3) associate

with the cytosolic TIR-containing adaptor, the myeloid differentiation primary response protein 88 (MYD88) (**Figure 9**). Recruitment of MYD88 leads to interaction of IL-1R-associated kinases (IRAKs) and the TNF receptor associated factor (TRAF) 6. TRAF6 recruits transforming growth factor (TGF)-β–activated protein kinase 1 (TAK1) and TAK1 activates the inhibitor of NF-κB (IκB) kinase complex (IKK), which consists of NF-κB essential modifier (NEMO), IKKα, and IKKβ, leading to phosphorylation of IκB proteins which in turn leads to their degradation. Once NF-κB is released from inhibition, the NF-κB subunits such as p50 and p65 translocate to the nucleus where they mediate the transcriptional activation of proinflammatory genes. Another transcription factor, IRF5, which is required to produce proinflammatory cytokines is also activated as a result of the activated TRAF6. In addition, TAK1 activation also leads to activation of p38 mitogen-activated protein (MAP) kinase and c-Jun N-terminal kinase (JNK), which then activates the activator protein 1 (AP1) transcriptional complex [113]. In contrast to the other TLRs, TLR3 recruits the adaptor protein TIR domain–containing adaptor- inducing interferon (IFN)-β (TRIF) upon ligand recognition, leading to recruitment of TRAF3 and activation of TRAF family member associated NF-κB activator (TANK)-binding kinase 1 (TBK1), IKKε, and IRF3, and activation of NF-κB. Activation of IRF3 then induces IFN-β production. TRIF also recruits TRAF6, which leads to the activation of NF-κB.

TLR4 is unique in the sense that it is the only TLR which requires an accessory molecule, myeloid differentiation factor 2 (MD2). The TLR4/MD2 complex recognizes LPS [114]. LPS is an amphipathic molecule consisting of a hydrophilic core oligosaccharide, an O-antigen and a lipid A region with hydrophobic multi-acyl chains which causes micelle formation in aqueous solutions [114]. To be presented to TLR4/MD2 complex, these hydrophobic chains need to be extracted from the bacterial membrane. Thus, two accessory proteins called cluster of differentiation protein 14 (CD14) and lipopolysaccharide binding protein (LBP) are required to mediate this transfer and increase the sensitivity of LPS detection [115]. LBP is a glycoprotein with a hydrophobic region which forms a high-affinity complex with LPS micelles to deliver them to CD14 [115]. CD14 is found anchored to the plasma membrane on myeloid cells or in a soluble form in endothelial and epithelial cells [116]. By transiently binding, it transfers LPS to TLR4/MD2 and can mediate LPS-

induced endocytosis of TLR4. Upon binding of LPS, TLR4, MD2 and LPS together with adaptor molecules, such as MYD88, TIR domain containing adaptor protein (TIRAP), TRIF and translocation associated membrane protein (TRAM), form a complex [117] for activation of TLR4 signaling cascade [116].

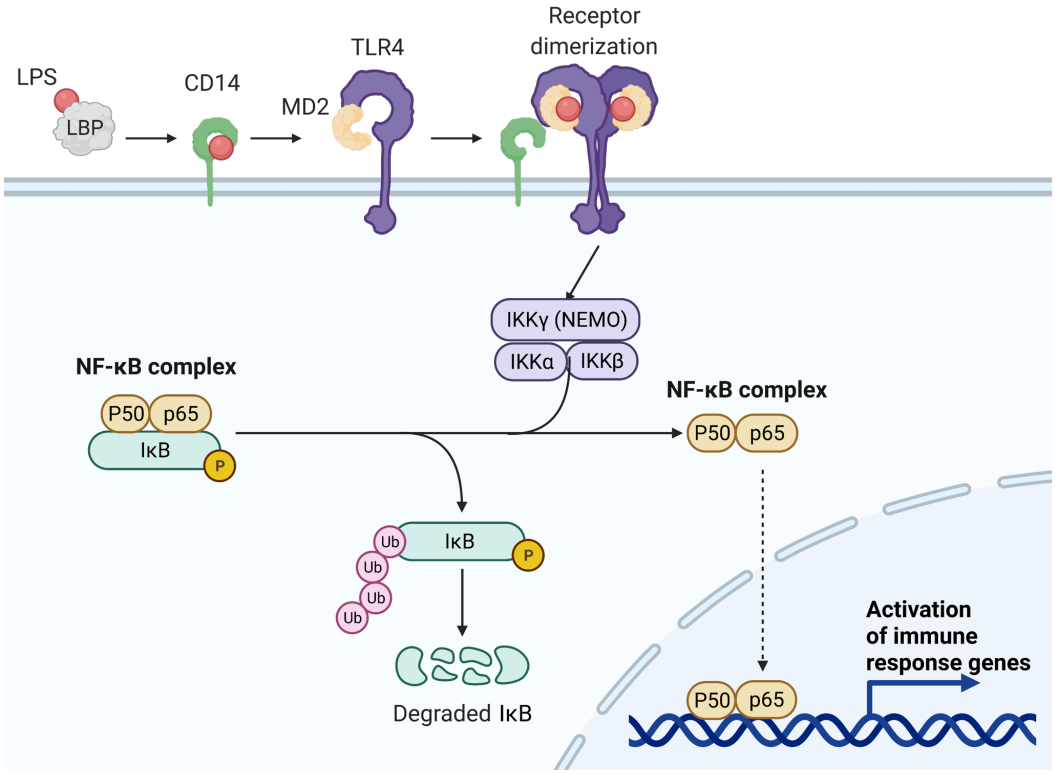


Figure 9: Overview of the TLR4 signaling pathway. LPS, the ligand of TLR4, binds to LBP to be transferred first to CD14 and then to TLR4/MD2 complex to initiate the dimerization of TLR4. The TLR4/MD2/LPS complex can activate the NF-κB pathway by triggering the signal transduction via recruitment of different adaptor molecules which lead to the activation of IKK complex formed by IKKα, IKKβ and IKKγ (NEMO). This complex phosphorylates the inhibitor of NF-κB which leads to its degradation. The released subunits of NF-κB can translocate to nucleus and initiate the transcription of immune response genes.

1.2.1.2 NOD-like receptors

NLRs are intracellular PRRs ensuring cytosolic protection. They are structurally characterized by their modular organization of three domains with distinct functions. A variable C-terminal LRR domain binds to the ligands, the evolutionarily conserved and centrally located NOD contributes to the oligomerization of the receptor, and a variable N-terminal protein-protein interaction domain mediates the downstream signaling cascades [93]. Depending on this N-terminal domain, the NLRs are classified into four types: NLRAs contain an acidic transactivation domain, NLRBs, which are also called NLR family apoptosis inhibitor proteins (NAIPs) contain a baculoviral inhibitor of apoptosis repeat (BIR) domain, NLRCs have at least one

caspase activation and recruitment domain (CARD), and NLRPs possess a pyrin domain (PYD) instead. In contrast to the other subclasses, the effector domain of NLRX1 does not share homology with the N-terminal region of any of the other subclasses and is instead classified as a CARD-related X effector domain [118]. Both CARD and PYD are involved in a range of cellular processes including production of proinflammatory cytokines and apoptosis.

MAMPs and DAMPs recognized by the C-terminal LRR domain of the NLRs triggers a conformational change and allow for the oligomerization of the proteins through the central NOD domains. This in turn facilitates the recruitment of downstream signaling adaptors and effector proteins culminating in formation of an oligomeric complex. Within the NLR family, the activating stimuli, the responsiveness and the activation mechanism differ strongly and the physiological functions and signaling pathways of only some of these receptors have been well defined. Among the intensively studied NLRs, NLRP1, 3, 6, 12, NLRC4 and NAIP form a cytoplasmic signaling complex called inflammasome that activates caspase-1, which cleaves and activates the proinflammatory cytokines pro-IL-1 and pro-IL-18. These cytokines in turn recruit professional immune cells of the innate and adaptive immune system and can also trigger an inflammatory type of cell death called pyroptosis [119, 120] (**Figure 10**). On the other hand, NLRP10, NOD1, NOD2, NLRC3, NLRC5 and NLRX1 form signaling platforms independently of the inflammasome and mediate innate immunity through the activation of NF-κB [121, 122] and mitogen-activated protein kinase (MAPK) pathways [123].

NOD1 and NOD2 (also known as NLRC1 and NLRC2) are the first NLRs to have been described. These receptors both detect bacterial cell wall moieties derived from bacterial peptidoglycan, although the exact molecular pattern each one recognizes differs: NOD1 binds to gamma-D-glutamyl-meso-diaminopimelic acid and NOD2 is activated by muramyl dipeptide (MDP) from bacteria [124]. Upon sensing their ligands, the CARD is exposed and the receptor-interacting serine-threonine kinase 2 (RIPK2) adaptor protein is recruited which leads to polyubiquitination of the inhibitor of NF-κB kinase subunit gamma (IKKγ). This activates the IKK complex, which then degrades and activates NF-κB [125, 126].

On the other hand, in the case of the other inflammasome forming NLRs, such as NLRP3, conformational changes in the molecule following ligand sensing expose the PYD resulting in interaction with a similar domain on the adaptor protein apoptosis associated speck-like protein containing a CARD (ASC), which associates with pro-caspase-1, resulting in its cleavage to active caspase-1 [127]. The latter then cleaves the inactive forms of IL-1β and IL-18 to generate active cytokines. Subsequently, once the inflammasome is assembled, this leads to induction of pyroptosis.

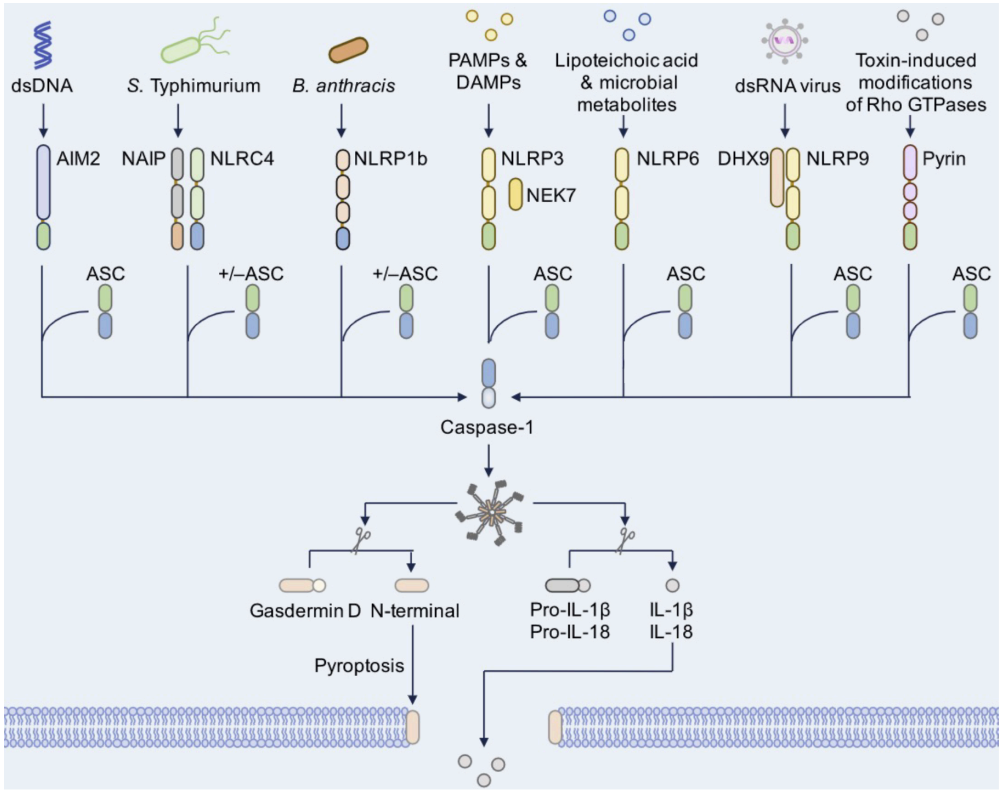


Figure 10: Inflammasome formation pathways. NLRs are cytosolic receptors. Inflammasomes occur when the MAMPs or DAMPs are recognized. These receptors interact with their respective ligands which leads to the activation of inflammasomes. Once the caspase-1 is activated, it cleaves GSDMD which leads to pyroptosis. Activated caspase-1 also cleaves the inactive pro-IL-1b and pro-IL-18 resulting in their activation which then are released through the pores formed as a result of pyroptosis.

Pyroptosis differs from apoptosis and necroptosis mechanistically, since it leads to pore formation in the cellular membrane for the release of the proinflammatory cytokines [128]. Following the activation of the inflammasome, caspase-1 cleaves a gasdermin protein, gasdermin D (GSDMD), leading to self-oligomerization of its N-terminal region. This creates a protein complex driving the pore formation in the cellular membrane for the cytokines to be released from the cell and this results in GSDMD-mediated pyroptotic cell death [129].

1.2.1.3 C-type lectin receptors

CLRs are proteins that are either transmembrane or soluble in nature and have a wide range of function. These receptors are mainly engaged in recognition of an extensive range of fungal carbohydrate structures; however, they may also detect a number of structurally varied microbe- or host-associated ligands [130]. CLRs capable of detecting carbohydrates are distinguished by the presence of a C-type lectin domain (**Figure 8**). Over a thousand members of the CLR family are known to be involved in endocytosis, phagocytosis, cell adhesion, complement activation, tissue remodeling as well as antimicrobial, proinflammatory, and anti-inflammatory responses. Dectin-1, dectin-2 and dendritic cell-specific intercellular adhesion molecule (ICAM)3-grabbing nonintegrin (DCSIGN) are some of the well-known examples of the CLR family. Upon interacting with its ligand, the tyrosine residues of CLRs are phosphorylated and spleen tyrosine kinase (SYK) is recruited, leading to the activation of NF-κB pathway [98].

1.2.1.4 RIG-I-like receptors

The RIG-I-like receptors (RLRs) are intracellular sensors for RNA. The RLR family involving RIG-I and melanoma differentiation factor 5 (MDA5) exhibit unique domains: A C-terminal domain, a central DExD/H box helicase domain for RNA binding, and two N-terminal CARD domains that are involved in downstream signaling (**Figure 8**). Both receptors were discovered to have a function in the detection of RNA viruses. RIG-I binds to short (< 300 bp) dsRNAs with blunt ends and a 5′ triphosphate moiety, thus, recognizes mainly negative-sense ssRNA viruses or positive-sense ssRNA/dsDNA viruses. MDA5 on the other hand, detects long dsRNA strands (~41,000 bp) with no end specificity and primarily detects the intermediate dsRNA generated during replication of positive ssRNA viruses. Despite the fact that these receptors detect RNA viruses in different ways, they activate similar downstream cascades in which the ligand-bound receptor interacts with mitochondrial or peroxisomal adaptor proteins to activate the cytosolic protein kinases IKK and TBK1 leading to transcription of NF-κB and IRF3, facilitating the production of type I interferon and other proinflammatory cytokines [98].

1.2.1.5 Other intracellular sensors

In addition to the classical PRRs, further sensors of microbial activities have recently been discovered. For example, a recent study confirmed ALPK1 as a cytosolic innate immune receptor recognizing the LPS metabolite ADP-β-D-manno-heptose (ADP-heptose) [131]. ALPK1 stimulation, results in phosphorylation of TRAF-interacting protein with forkhead associated domain containing protein A (TIFA), mediating the formation of TIFAsomes as a response to gram-negative bacteria [132, 133].

Cytoplasmic DNA derived from replicating intracellular pathogens or from dying host cells has a potential to serve as a danger signal. A variety of proteins including AIM2 and cGAS were reported to monitor the cytoplasm for host or pathogen derived DNA (**Figure 8**). Upon sensing dsDNA, AIM2 forms an inflammasome with ASC and caspase-1, leading to enzymatic cleavage of the proinflammatory cytokines IL-1β and IL-18 [134]. The cGAS-STING pathway consists of the DNA-binding protein cGAS, the messenger product cyclic GMP-AMP (cGAMP) and STING. Together these constitute a DNA sensing machinery in cytoplasm capable of sensing microbial or self-DNA and activating the type I interferon pathway [135]. Firstly, DNA of various origin in the cytoplasm binds and activates cGAS, leading it to convert adenosine triphosphate (ATP) and guanosine triphosphate (GTP) into cGAMP. In turn, cGAMP binds to STING which is an adaptor protein found in the endoplasmic reticulum membrane, leading to its oligomerization and activation [136]. Activated STING is translocated to other regions such as Golgi apparatus and triggers the production of type I interferons [137, 138]. Another important function for clearance of cytoplasmic DNA is autophagy occurring downstream of the cGAS pathway is autophagy which is again induced by activated STING [139].

1.3 GI innate immune signaling in homeostasis and disease

After birth, the GI tract is rapidly colonized by a diverse community of microorganisms commonly referred to as the microbiota. They live symbiotically with the host by participating in the digestion of food and production of vitamins. In addition, signals from the microbiota under basal conditions play a vital role in maintaining homeostasis and proper barrier function of the epithelium [140]. For example, germ-free mice have lower cell turnover rates and an altered brush border compared to wild-type mice [141, 142]. Moreover, commensals (non-pathogenic

microbiota) protect the host from colonization by pathogens through establishing a colonization barrier [143]. However, the mutualistic host-microbe relationship may be damaged by various factors, one of which is the local or systemic inflammatory responses by the host immune system. Although PRR signaling is a crucial part of the innate immune system, it might cause excessive inflammation and tissue damage when dysregulated at mucosal surfaces. On the other hand, emerging research on epithelial barriers indicates that the epithelial PRR signaling significantly contributes to GI homeostasis [5, 7, 144]. In line with these statements, two concepts about epithelial PRR sensing in GI tract have emerged:

(1) PRR activation is important for homeostatic regeneration and barrier integrity. The expression and function of PRR signaling is expected to not only have a major impact on pathogen sensing, but also on tissue homeostasis and inflammatory diseases [7, 145–147]. For example, when NF-κB is inhibited, epithelial cells become sensitized to environmental stimuli from the microorganisms or immune cells and this might lead to dysregulation of immune homeostasis by triggered inflammation [147]. Mice deficient for either of the TLR2, 4, 5 or 9, the TLR-pathway mediator MYD88 or the NLRP6 are highly susceptible to experimentally induced colitis [148–151]. NLRP6 deficiency also leads to impairment of goblet cell function and reduced mucus secretion [152]. While these studies clearly point to the importance of innate immune signaling for epithelial homeostasis, it is unclear, whether the observed impact is due to the innate immune signaling in epithelial cells or in professional immune cells.

To disentangle the interplay of professional immune cells and epithelial cells, several studies have used epithelium-specific knockouts or organoids. For example, the stimulation of murine small intestinal organoids with the NOD2 agonist MDP increased the number of organoids growing out of isolated stem cells, indicating that the innate immune signaling supported survival of the stem cells [153, 154]. Intestinal epithelial cell-specific ablation of IKKα and IKKβ or single knockout of IKKγ, also known as NEMO, led to spontaneous intestinal inflammation and an IBD-like phenotype. This was dependent on the presence of the TNF-α receptor on epithelial cells, indicating that the anti-apoptotic function of the NF-κB pathway is necessary to protect the epithelial cells from TNF-α-induced apoptosis [146]. Interestingly, in

humans polymorphisms in innate immune genes such as *NOD2* and *TLR4* are associated with an increased risk to develop IBD [155] and inhibition of TNF-α is currently the most efficacious treatment for IBD in a subgroup of patients [156]. Thus, a picture emerges in which a low level of innate immune stimulation is important for mucus secretion, barrier integrity and epithelial cell survival. Its impairment may allow translocation of intestinal bacteria from the lumen into the subepithelial tissue, thereby leading to inflammation.

(2) The epithelium needs to minimize the inflammation by the commensals. It should be kept in mind that pathogens and commensals are in theory sensed by similar PRRs. Therefore, to preserve gut homeostasis and prevent disproportionate reactions against the gut microbiota, tight PRR-regulating mechanisms are necessary. For example, while the GI immune system can mount a strong inflammation against invading pathogens, it can also provide many layers of inhibitory and tolerogenic mechanisms not to mount excessive immune response against the commensals [9, 10].

Considering these concepts, the studies on the epithelial barrier partially highlighted regulation of microbial receptors on various levels, such as by limiting the access of the ligands, by polarized expression of the receptors or inactivating downstream signaling cascades [5, 95]. However, the points summarized below need to be addressed and studied further to reveal the mechanisms contributing to the organization of GI epithelial innate immunity:

(1) The knowledge on the expression and function of the PRRs is controversial. Most of the knowledge on epithelial innate immunity has been collected from animal models which might not be enough to model the distinctive characteristics of human cells, and from human cancer cell lines which may have altered signaling pathways. It remains a challenge to disentangle the immune cell-epithelial cell interplay *in vivo* or *ex vivo*, where even minor contaminations with immune cells can confound the results. As a consequence of these technical challenges, the results of expression and functionality of PRRs in GI tract have been highly contradictive and the knowledge on whether a given PRR is expressed in the GI epithelium and would activate downstream signaling upon recognition of its ligand is limited [157].

(2) The presumable relation between the PRR expression and microbial load varying between the segments of the GI tract is unclear. The spatial compartmentalization along the GI tract which is important for regional functions, such as nutrient uptake and digestion, was speculated to also determine the expression of the PRRs by following a pattern according to the microbial load and eventually be silenced towards the gut lumen as a tolerogenic approach [95]. To date, only a few studies have directly assessed the variation in epithelial TLR expression along the length of the GI tract or only the intestines [158, 159]. Northern blots for messenger RNA (mRNA) of *Tlr2* and *Tlr4* in *ex vivo* isolated epithelium indicated that expression levels of these two *Tlr* molecules were segment-specific: *Tlr2* was expressed mainly in the colon, while *Tlr4* was mainly expressed in stomach and colon. The authors already termed this "strategic compartmentalization" of these TLRs [158]. For another study, five strains of reporter mice were generated, enabling the expression analysis of TLR2, 4, 5, 7 and 9, respectively. TLR2 and 5 were expressed in the small intestine and the proximal colon, TLR4 in colon, TLR7 and 9 were not expressed. What exactly shapes this organization is not yet understood though.

Moreover, numerous GI disorders are confined to a specific section of the GI tract. IBD includes Crohn's disease (CD) and ulcerative colitis (UC), which show a differential and disease-specific pattern of inflammation: while UC begins in the rectum and is found in the colon, CD can affect all parts of the GI tract from the oral cavity to the anus. Furthermore, CD is characterized by segmental, discontinuous inflammation within the GI tract while UC is usually described as a continuous inflammation of the colon [160, 161]. In the esophagus, stomach and colon, cancer incidence is high, and infection and inflammation can promote development and progression of these cancers [162, 163]. In contrast, malignant transformation in the small intestine is very rare [164]. The apparent segment-specificity of these diseases in the GI tract remains enigmatic, but it is tempting to speculate that they may be due to regional disturbances of the tightly balanced system of epithelial barrier function, innate immunity and mucosal regeneration. Therefore, it would be insightful to highlight differences between segments within the GI tract.

(3) It is unclear whether PRRs are located apically and in contact with microbial stimuli constantly. Gastrointestinal epithelial cells are highly polarized, with a specialized apical side facing the lumen of the gut with its microbiota, and a basolateral side facing the tissue. Under homeostasis, MAMPs only reach the apical side. However, when the epithelial barrier is breached, microorganisms can also challenge the basolateral side. It has thus been hypothesized that epithelial cells may only selectively mount a proinflammatory response when stimulated from the basolateral side, in order to match the threat posed by the signal. For example, TLR5 only induced the NF-κB response gene IL-8 when stimulated from the basal side [104] and TLR9 was demonstrated to induce distinct signaling pathways when stimulated from the apical or basolateral side in cancer cell lines [165]. However, while earlier studies using antibody labeling against TLRs have reported specific expression on one side only [95, 96], analysis of TLR reporter mice using staining of an hemagglutinin (HA)-tag did not confirm this but instead demonstrated TLR2, 4, and 5 receptors on both apical and basal sides as well as some intracellular TLR4 [159]. These apparent differences are likely due to the different technical approaches and it is unclear whether functional PRRs are also apically expressed.

(4) The mechanisms contributing to the organization of GI epithelial innate immunity are not fully understood. The main concept in this regard has been the induction of tolerance after colonization of the sterile gut during birth, the so-called "window of opportunity" [166–168]. This concept postulates a priming period of the innate and adaptive immune system after birth, which sets the stage for immune homeostasis and subsequent host-microbial interactions. Previous studies had reported regulation of PRR signaling in response to stimulation with MAMPs. For example, TLR4 responsiveness decreases after birth, presumably because of the exposure to LPS during delivery and subsequent colonization of the gut [169]. Also, stimulation of TLR9 with its ligand CpG-DNA leads to a decrease of *Tlr4* expression and inhibits TLR4 signaling [170]. Thus, it was proposed that contact with the microorganisms, their molecules and metabolites in the GI tract would lead to silencing of PRR expression towards the gut lumen [95]. However, germ-free vs specific pathogen-free mice did not show differences in TLR expression in either small intestine or colon, indicating that neither upregulation of TLR3 nor downregulation of TLR5 in this early period depend on the microbiota [159, 171]. Thus, further research is needed to

better determine whether the expression of the PRRs is regulated by the postnatal colonization of the GI tract or endogenously by the developmental processes.

Another long-standing concept in epithelial innate immunity is that PRRs are expressed in a lineage-dependent manner, where the cellular heterogeneity and localization of the cells in crypt-villus axis or along the gastric gland can strategically help to activate and regulate proliferative and pro- or anti-inflammatory pathways. For example, NOD2, which was detected first in the crypt region of the murine small intestine [172], appears to be restricted to the stem cells [154]. Stimulation with NOD2 ligand regulates the secretion of several alpha-defensins by Paneth cells, which in turn leads to activation of adaptive immunity and also increases the survival of stem cells, indicating that stimulation of PRRs may also regulate gut epithelial regeneration directly. *Tlr4* was also reported to be expressed mainly in the intestinal crypts [173, 174], and was not found on the murine small intestinal villi or Paneth cells [175]. Similarly, *Tlr1*, *2* and *4* were observed only in epithelial cells of murine colonic crypts, which suggests that their expression might be lost as the cells mature and migrate upwards [158]. However, a broad analysis on how the cellular heterogeneity in epithelial layer affects the localization of PRRs and related signaling elements have not been performed.

2 Aims of this study

The GI epithelium is an important barrier to separate the body from the gut lumen. Its capacity to sense pathogens and tolerate the microbiota is thought to play a central role in infection control and healthy homeostasis. However, the development, organization and function of GI epithelial innate immune sensing remains enigmatic. In this thesis, I aim to systematically profile gene expression in GI epithelial cells and characterize the innate immune signaling in these cells.

For this, I pursue the following 3 objectives:

1. Establishment and characterization of an organoid biobank from 6 different segments of human and murine GI tract, each represented by lines from 3 donors. The organoids consist purely of primary epithelial cells and will be used as GI epithelial model throughout the thesis. I will use RNA sequencing to gain a global overview of the gene expression.

2. Systematic profiling of the innate immune barrier in GI epithelial cells using the organoids of the biobank. I will use RNA sequencing data to focus on the differential expression of innate immune signaling components among the segments and analyze the function of some specific differentially expressed PRRs.

3. To address the question, which factors contribute to the regulation of epithelial innate immunity, I will analyze the impact of (i) exposure to microbial components, (ii) cellular heterogeneity within the epithelial layer of one GI segment, and (iii) the tissue identity exemplarily shown by expression of the transcription factor *Cdx2*.

3 Results

The results are presented in 3 parts: 1. Characterization of biobanks of murine and human GI organoids, 2. Analysis of innate immunity in murine and human GI tract, 3. Analysis of the factors regulating epithelial innate immunity

3.1 Characterization of biobanks of murine and human GI organoids

3.1.1 Tissue-resident stem cells from different segments of GI tract formed into organoids

To better understand the different cellular functions in murine and human GI epithelial cells, a biobank of organoids for the murine and human GI tract was generated. For this, glands of the gastric corpus and pylorus, as well as crypts of the small intestinal duodenum, jejunum, ileum and colon were isolated from three individuals each by incubating the tissue in ethylenediaminetetraacetic acid (EDTA). The amount and duration of EDTA treatment for extraction of glands and crypts were optimized depending on the organism and the segment of GI tract (Methods-7.1.1, **Table 16**, page 95). The isolated glands and crypts were either used for RNA analysis directly or were used to generate organoids (**Figure 11**). Murine corpus and pylorus organoids were grown in murine gastric medium, duodenal and jejunal organoids in murine small intestinal medium, and ileal and colonic organoids were grown in murine colonic medium. In contrast to the rest of the small intestinal segments, ileal organoids did not grow in small intestinal medium which did not contain WNT. For human organoids, only two types of media were used, human gastric medium and intestinal medium (for a complete overview over the different media and components refer to Methods-7.1.2**, Table 17**, page 98).

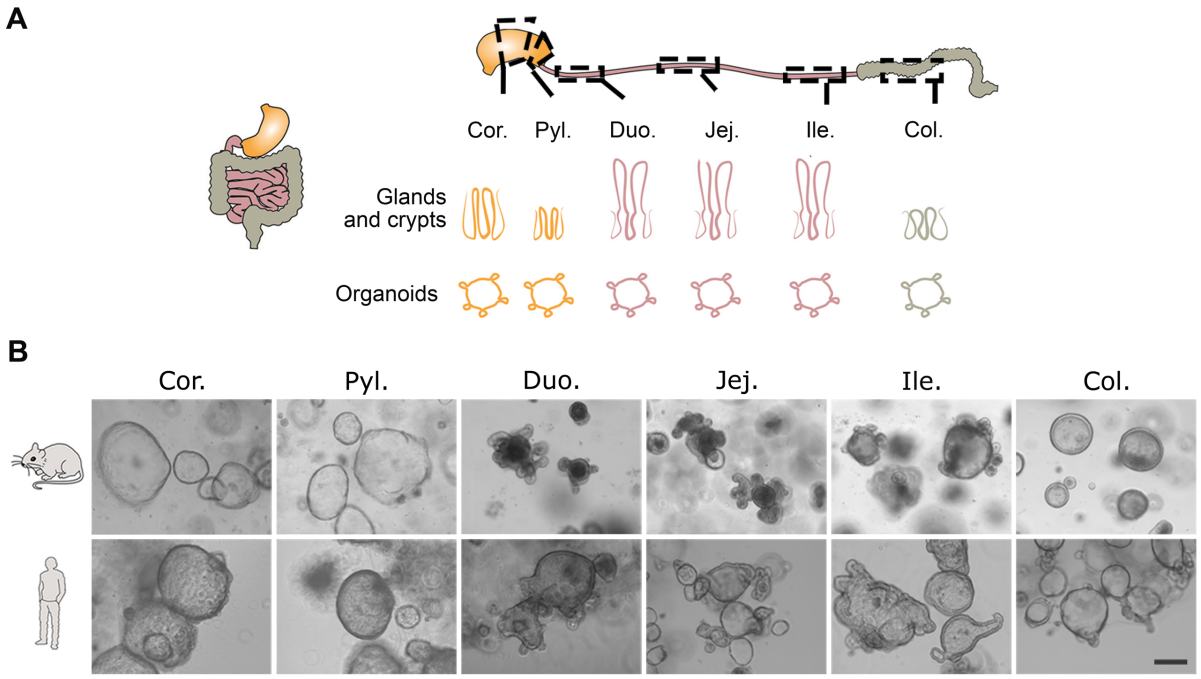


Figure 11: A living biobank was generated from human and murine GI organoids. (A) Organoids were derived from indicated segments of the human and murine GI tract, from three individuals each. Illustration is taken from Kayisoglu *et al.* [176]. **(B)** Representative light microscopic images of human and murine GI organoids. Scale bar: 200 µm.

To analyze the purity of the organoids, they were expanded for a maximum of five passages to obtain enough material, then lysed and used for analyzing the RNA expression. PCR analysis of cell type-specific markers showed that organoid cultures were devoid of immune cells (*Cd45*) and mesenchymal cells (glioma associated oncogene 1 (*Gli1*) for murine samples, fibronectin 1 (*FN1*) for human samples) (**Figure 12A and B, top rows**). The *ex vivo* isolated crypts and glands on the other hand still could contain a small amount of these cell types (**Figure 12**, **left**). From this it can be inferred that purely epithelial murine and human organoids were obtained after passaging them for a couple of times.

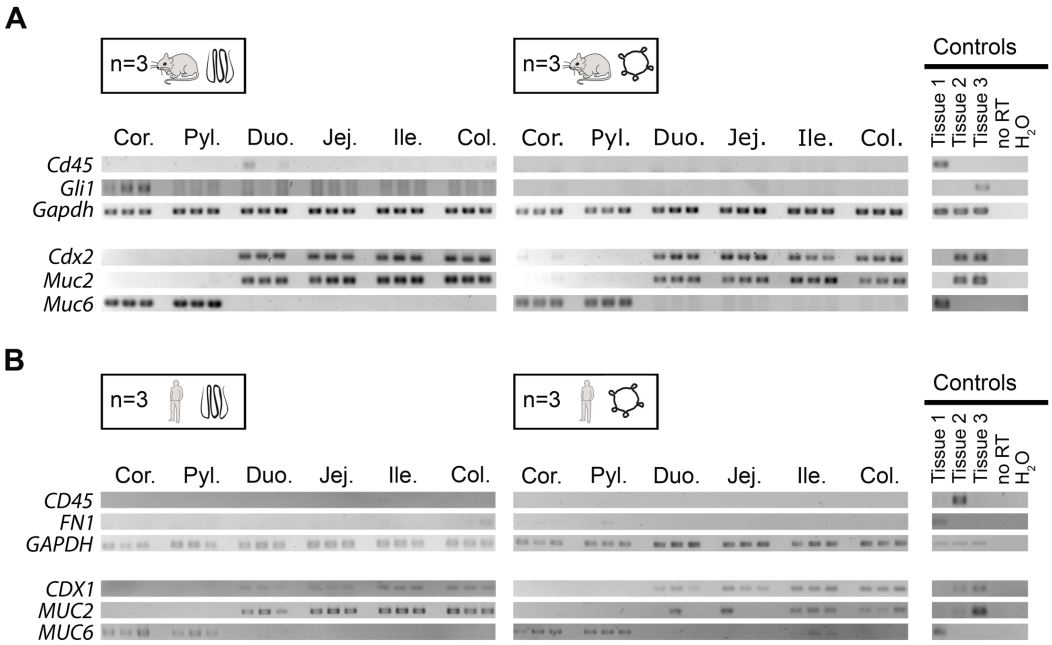


Figure 12: Purity of organoid cultures compared to freshly isolated epithelium. Crypts and glands were isolated from **(A)** 3 mice or **(B)** 3 patients per tissue from the segments of the GI tract. Organoids were grown from the tissue and expanded for a maximum of 5 passages. RNA was isolated, and conventional PCR was performed. Control tissues in **(A)** are murine spleen (positive control of *Cd45*), murine organoid pool (negative control), and mesenchyme tissue (positive control of *Gli1*). Control tissues in **(B)** are human mesenchyme tissue (positive control of *FN1*), PBMCs isolated from blood (positive control of *CD45*) and human duodenal organoids (negative control).

3.1.2 Organoids and respective monolayers consist of polarized epithelial cells

GI epithelial cells are polarized: The apical side faces the lumen of the GI tract and the basal side faces the lamina propria. To characterize the apical-basal orientation of the cells, human and murine gastric organoids were seeded as 3D organoids or as two-dimensional (2D) monolayers on transwells (schematically depicted in **Figure 13A and D**). To better visualize the polarization, WNT in the medium was withdrawn, which led to the differentiation of cells into mainly MUC5AC-producing surface pit cells. Cells were fixed, processed for histology and mucus was stained using periodic acid-Schiff (PAS) staining. The observation of apically located pink mucus clusters and accumulation of mucus inside of the 3D organoids or on 2D monolayers suggested a polarized epithelial cell population for both murine (**Figure 13B and E**) and human (**Figure 13C and F**) cells.

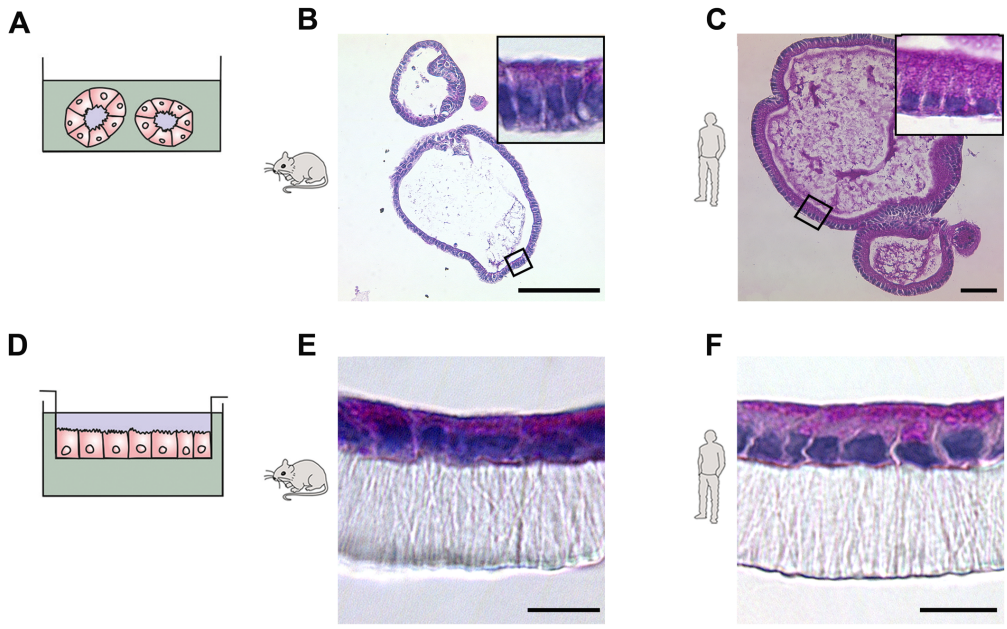


Figure 13: Human and murine epithelial cells forming the 3D organoids and 2D monolayers are polarized. Illustration of **(A)** 3D organoids and **(D)** 2D monolayers on transwells. **(B, C, E, F)** Organoids and monolayers were differentiated by removal of WNT from the culture medium. Differentiated and MUC5AC-producing surface pit cells were fixed, processed for histology and PAS staining was performed, which shows mucus on the apical side (pink). **(B)** Murine gastric organoids. **(C)** Human gastric organoids. **(E)** Murine 2D monolayer grown from organoid cells. **(F)** Human 2D monolayer grown from organoid cells. Boxed areas are shown in higher magnification in the upper right corner. Scale bars: 100 µm **(B, C)** or 10 µm **(E, F)**. *(These experiments were performed, and data were collected by Carolin Niklas.)*

3.1.3 GI segment-specific RNA expression is maintained in human and murine organoids

To profile and compare the gene expression data among the samples of the organoid biobank generated from 3 murine or 3 human donors, the isolated RNA was subjected to bulk RNA sequencing (RNA-seq). The inter- and intragroup variabilities among different segments were visualized using a correlation heatmap depicting the samples from each segment clustering together, suggesting that there were no outliers among the biological replicates. Additionally, gastric segments (corpus and pylorus), proximal intestinal segments (duodenum and jejunum) and distal intestinal segments (ileum and colon) clustered close to each other in both murine and human datasets (**Figure 14A**). Likewise, principal component analysis (PCA) was plotted displaying all 18 samples per species to visualize the first two linearly transformed dimensions reflecting the total variation of the datasets (**Figure 14B**). Principal component 1 (PC1), which describes the highest variation within the data, accounted for 70% of variance in murine samples and 68% in human samples. PC2, the second

highest, accounted for 20% of variance in murine samples and 14% in human samples. Accordingly, the PCA plot demonstrated a clear separation of the organoids from different regions and grouping of organoids from the same region, where PC1 corresponded to the difference between the gastric and intestinal segments and PC2 to the difference among the intestinal segments. Therefore, organoids derived from the two gastric regions clustered together, as did those derived from duodenum and jejunum. Ileal organoids clustered between the other small intestinal segments and colon in both species. These data indicate that the tissue and organ identity is encoded and maintained in the adult stem cells within the organoids.

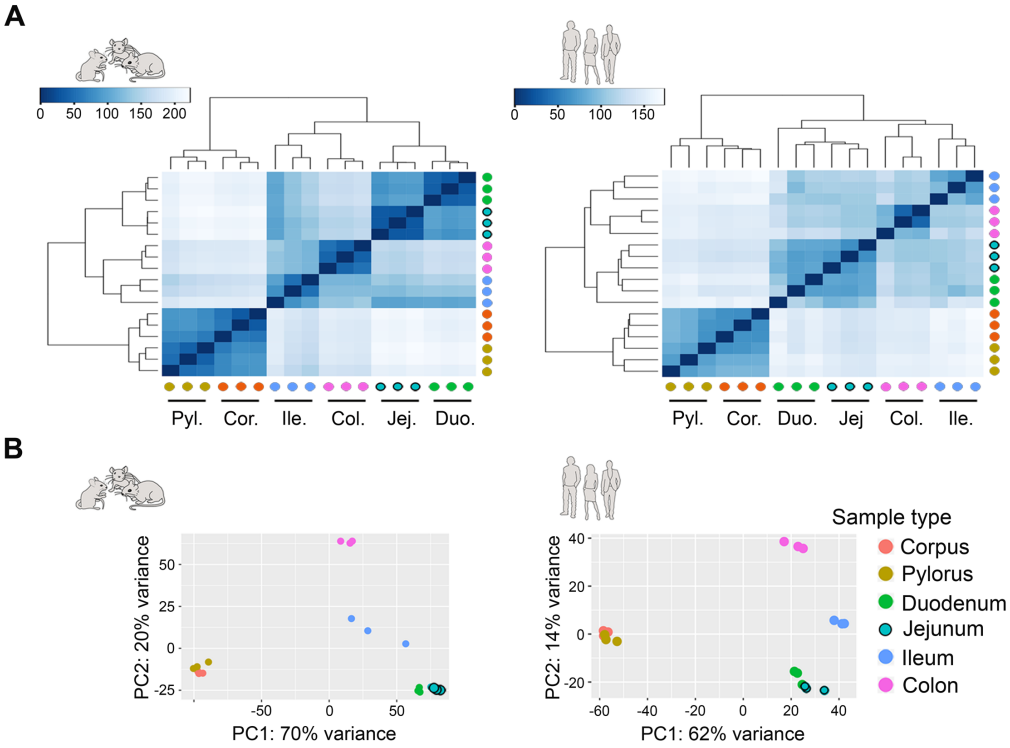


Figure 14: Transcriptome profiling of human and murine GI organoids reveals segment-specific expression. **(A)** Heatmaps showing the correlation between the murine **(left)** and human samples **(right)**. **(B)** PCA plot of the RNA sequencing data of 18 samples from each species along PC1 and PC2, highlighting the inter- and intragroup variability among organoids.

To investigate the cellular identity of the epithelial cells of different GI segments, RNA-seq data were analyzed further. A heatmap of a manually curated list of commonly used markers of GI segments defined the expression profiles along the cephalocaudal axis, where the organoids derived from the GI segments still show the expression profiles of the tissues they were derived from. For both murine and human organoids, the heatmaps of four groups of genes defined expression profiles

along the cephalocaudal axis (**Figure 15**, **supplementary Figure S 1**). Claudin 18 (*Cldn18*), which is a gastric cell marker, was expressed in all gastric samples but not expressed in the intestinal ones. In addition, expression of the mucins *Muc1*, *Muc5ac*, *Muc6*, pepsinogen C (*Pgc*) and *Tff2* in the gastric regions suggested the existence of differentiated cell types specialized for secretion of mucins and pepsinogen. Likewise, the marker genes of the intestinal segments such as *Cdx1* and *Cdx2* were not expressed in the stomach but only in the intestinal segments. The marker genes of the mature enterocytes, alkaline phosphatase (*Alpi*) and fatty acid binding protein 2 (*Fabp2*) were highly expressed in the proximal small intestine, but their expression was low in stomach and colon. Although murine ileal and colonic organoids grown in colonic medium had lower levels of mature enterocyte markers, intestinal identity markers such as *Cdx1* and *Cdx2* were expressed in similar amounts to the other intestinal segments. Expression of another group of genes including *Nox1*, which is abundantly expressed by the colonocytes [177] was mainly observed in the colonic organoids and the ileal organoids to some degree (**Figure 15, left**). In human organoids, tissue identity was similarly conserved and though not as quite as distinctively distributed as in murine organoids, tissue-specific genes such as *CDX1*, *CDX2*, *VIL1*, *MUC5AC* and *NOX1* were expressed in organoids from different regions as expected (**Figure 15, right**).

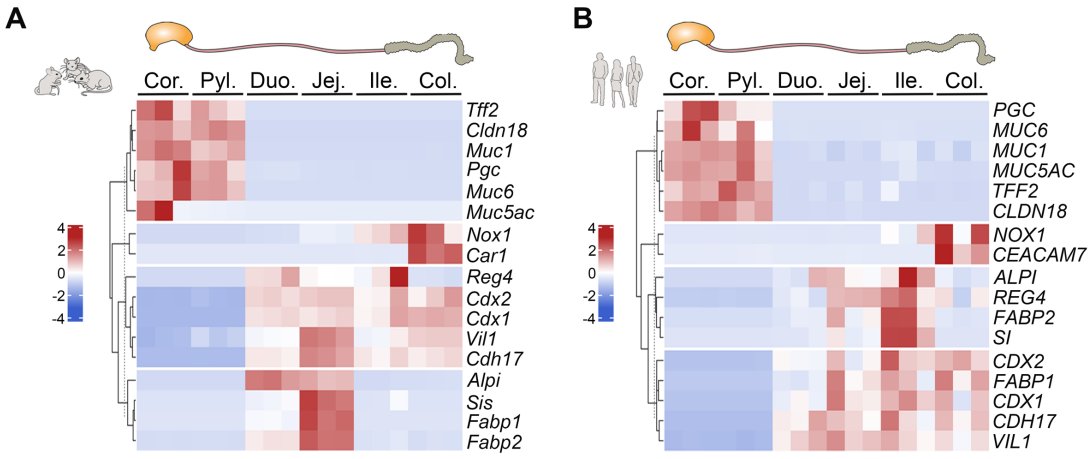


Figure 15: Organoids express tissue identity markers in a segment-specific manner. Heatmaps displaying normalized and scaled gene counts of selected gastric and intestinal cell marker genes along the GI tract for murine **(A)** and human **(B)** organoids. For these genes, **supplementary Figure S 1** depicts the bar graphs of expression values.

3.2 Analysis of innate immunity in murine and human GI tract

3.2.1 Expression of innate immune signaling components are both GI segment- and species-specific

To further explore this intriguing conservation of tissue identity, the genes which were characteristic for the three major regions were further explored by performing differential gene expression analysis (DEG). To simplify the dataset, three major regions of interest were defined: stomach, proximal small intestine and colon. Accordingly, the data of the two gastric regions, corpus and pylorus were combined (= "stomach"), as were the two proximal small intestinal regions, duodenum and jejunum (= "prox. int."), with the sole colon as the third group. DEGs between stomach and proximal intestine; and between proximal intestine and colon were identified. 2916 murine and 3575 human genes were significantly up- or downregulated by at least 2-fold between stomach and proximal intestine, but not between proximal intestine and colon. As expected, these contained well-known genes defining the gastric-intestinal border, such as *Cdx1*, *Cdx2* as the master regulators of intestinal development, and the tissue-specific mucins *Muc5ac*, *Muc6* (**Figure 16, black gene names**). Additionally, genes related to innate immunity were observed in these comparisons, such as *Nlrp6* and *Nod2* for murine, *TLR2* and *NOD2* for human organoids. 1328 murine and 1139 human genes were up- or downregulated by at least 2-fold in both comparisons, stomach versus proximal intestine and proximal intestine versus colonic organoids. Once again, these analyses also identified innate immune genes, such as *Tlr1*, *Tlr2* and *Tlr4* for murine organoids, *TLR6* and *NLRP6* for human organoids (**Figure 16, blue gene names**). Taken together, various innate immune signaling components displayed patterned expression among different segments of the GI tract in both species.

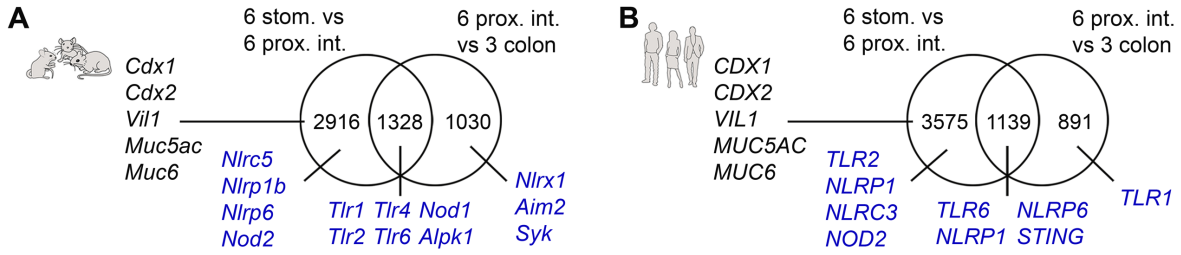


Figure 16: Expression of innate immune signaling components is GI segment- and species-specific. For the murine **(A)** and human **(B)** organoids lists of differentially expressed genes were generated, comparing 6 gastric organoids (3 corpus and 3 pylorus organoids combined) with 6 proximal intestinal organoids (3 duodenal and 3 jejunal organoids combined) or 6 proximal intestinal with 3 colonic organoids. Venn diagrams display number of differentially expressed genes ($p<0.05$ and log2FoldChange > ±1 (2-fold change)). Selected genes chosen for biological interest are listed. Black: Genes known for tissue identity. Blue: Genes known for immune function.

To assign biological relevance to the lists of genes generated as a result of DEG analyses of stomach versus proximal intestine and proximal intestine versus colon, gene lists were subjected to functional enrichment analysis using the website g:Profiler [178]. By comparing the provided gene lists to a background set of genes of each species, significance of enrichment for previously annotated and defined gene ontology (GO)-terms were assessed. Additional to the expected GO-terms for anatomical and developmental processes, the GO-term "Response to external stimulus" (GO:0009605), which also contains genes related to innate immunity, was highly significant and in the top 10 for both murine and human samples (**Figure 17**).

To further investigate whether the expression of immune-related genes follows a gradient along the cephalocaudal axis, possibly mirroring the microbial load, hierarchical clustering of all genes that are included in the GO-term "Response to external stimulus" (1939 murine 2012 human genes) was performed. The heatmap demonstrated a clear segregation among the segments of GI along the cephalocaudal axis, each segment containing a portion of highly expressed genes. This was more pronounced in the genetically identical laboratory mice but still visible in the more diverse patient samples (**Figure 18**).

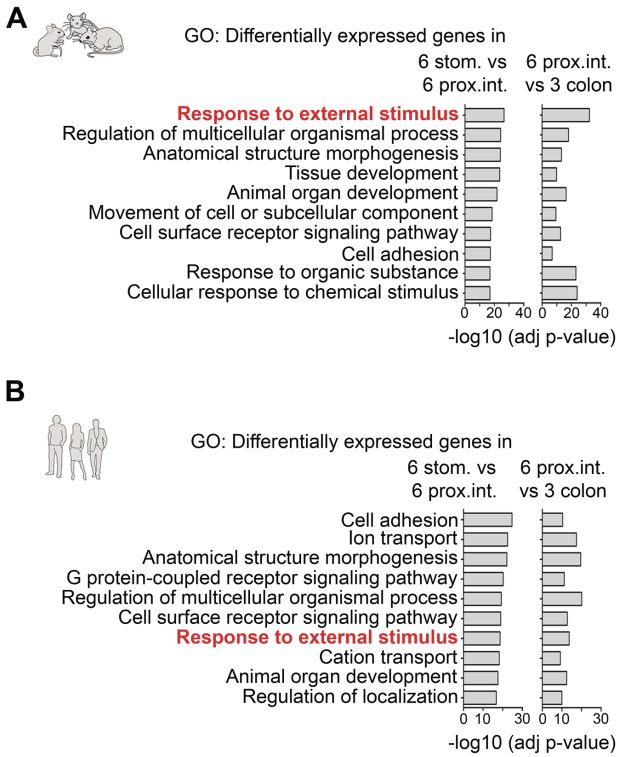


Figure 17: GO-term analysis revealed the term "Response to external stimulus" to be differentially expressed between different segments in both human and mouse. For the murine **(A)** and human **(B)** organoids lists of differentially expressed genes were generated, comparing 6 gastric organoids (3 corpus and 3 pylorus organoids combined) with 6 proximal intestinal organoids (3 duodenal and 3 jejunal organoids combined) or 6 proximal intestinal with 3 colonic organoids. GO enrichment analyses for biological processes were performed with the lists of DEGs between different GI segments. 6 gastric, 6 proximal intestinal and 3 colonic organoids were compared using the website g:Profiler. Scores of the top GO-terms indicate the enrichment p-value.

For a magnified view, genes of various PRRs and related proteins were selected from this GO-term, based on their biological relevance. These heatmaps also revealed extensive cephalocaudal organization of expression of the selected genes, again more pronounced in mice (**Figure 19**). These analyses showed that, expression of immune-related genes in the GI tract was neither uniform, nor followed a single pattern along the cephalocaudal axis. Instead, each segment expressed its specific set of immune-related genes.

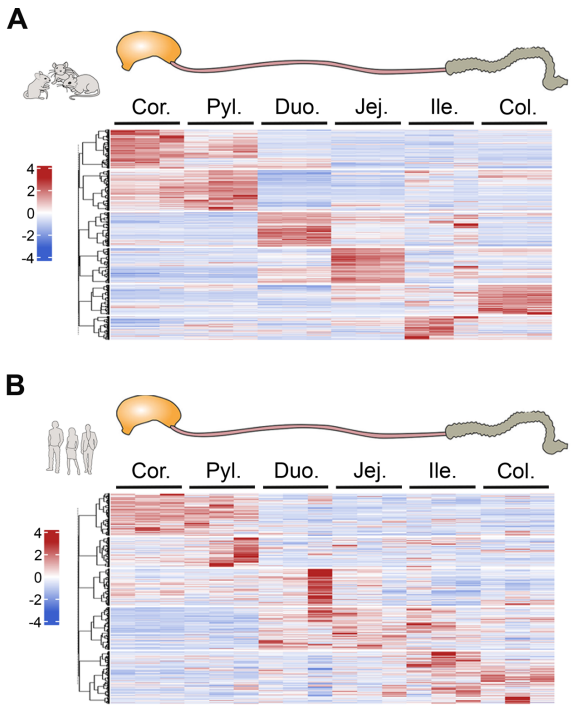


Figure 18: Genes in the GO-term "Response to external stimulus" cluster in segment-specific manner. Heatmaps depicting hierarchical clustering of 1939 murine **(A)** and 2012 human genes **(B)** in GO-term GO:0009605 – "Response to external stimulus".

Comparing the expression profiles of human and murine organoids revealed that some of the profiles were highly different between the species. A prominent example for this is *Tlr2*, which was most highly expressed in human gastric organoids but highest in murine colonic organoids. For other profiles, for example *Nod2*, *Nlrp6* and *Tlr4*, expression levels were comparable between mouse and human. Among these, *Nod2* was highly expressed in gastric organoids but expression was lower in intestinal organoids. *Nlrp6* was most highly expressed in organoids from the distal small intestine, and *Tlr4* was highly expressed in organoids from the stomach and colon but expression was lower in organoids from the proximal small intestine (**Figure 19**, **supplementary Figure S 2**). These data indicate that the expression profiles of immunity-related genes were highly species-specific and specific to each region of the GI tract.

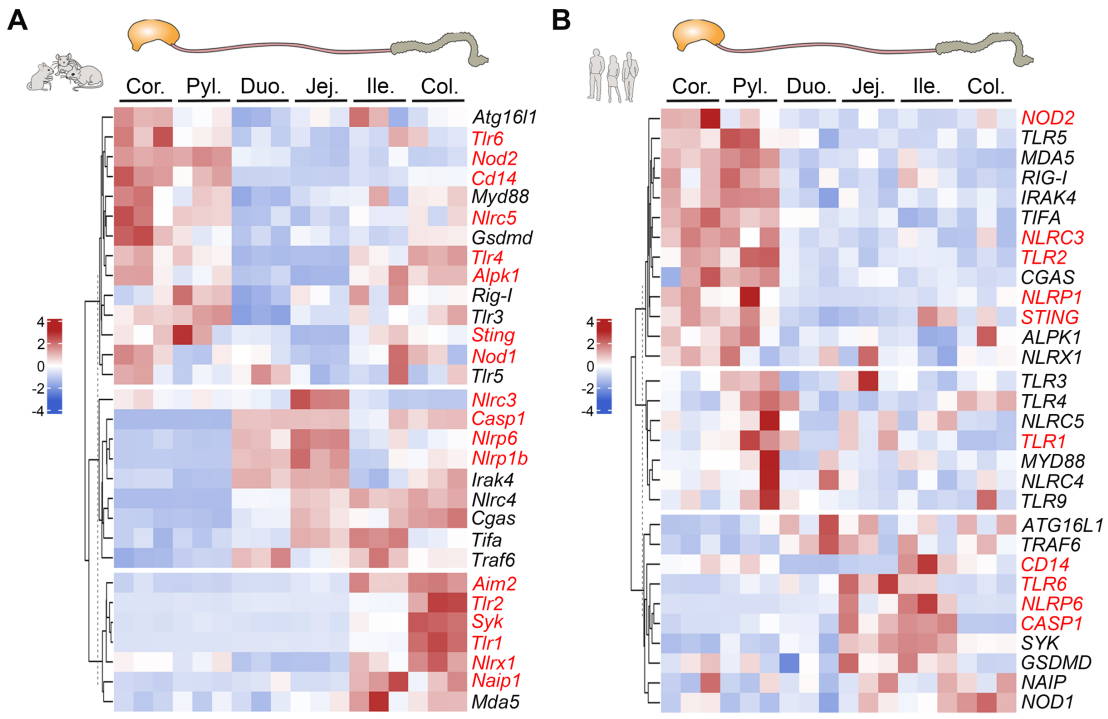


Figure 19: Many genes related to innate immunity reveal segment- and species-specific expression patterns. Heatmaps displaying normalized and scaled gene counts of selected genes from GO:0009605 "Response to external stimulus" for murine **(A)** and human **(B)** organoids. Genes identified as differentially expressed by the analysis ($p<0.05$ and log2Fold Change > ±1 (2-fold change)) are marked in red. (For bar graphs of expression values, see **supplementary Figure S 2**.)

3.2.2 Functionality of TLR2, 4 and 5 is GI segment- and species-specific

To analyze, whether the RNA expression translated into patterns of functionality, three exemplary TLRs were selected: TLR2 recognizes di- or tri-acetylated lipoproteins, TLR4 identifies LPS and TLR5 binds to flagellin.

In the sequencing data, *Tlr4* showed a peculiar expression pattern with high expression in the stomach and colon, but lower expression in the small intestine (**Figure 19**). This was in principle conserved in humans and mice, and was confirmed by conventional PCR and quantitative reverse transcription PCR (RT-qPCR) in organoids as well as primary isolated tissue (**Figure 20**). Expression of *Tlr4* showed similar profiles in *ex vivo* cells and organoids, although expression was not as high in organoids as in the original tissue (**Figure 20A**). Similarly, data comparing human *ex vivo* material with human organoids confirmed the specific expression profile of TLR4 (**Figure 20B**).

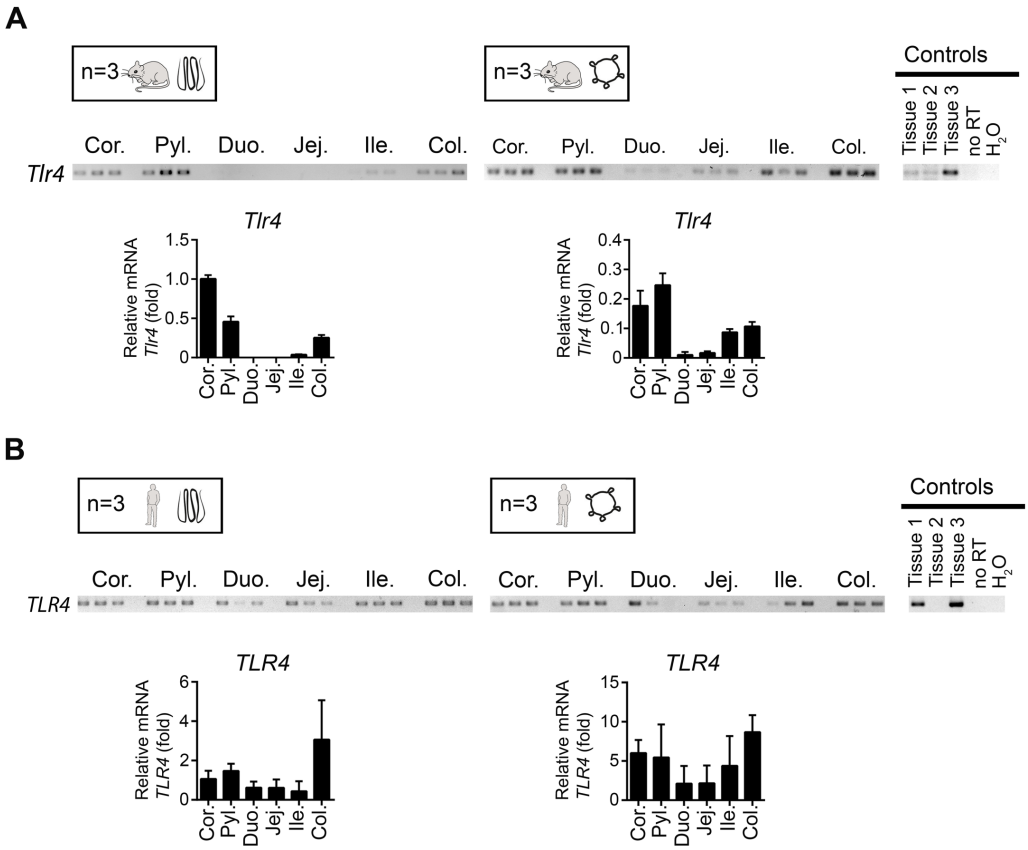


Figure 20: Expression of *Tlr4* in human and murine tissues and organoids. Crypts and glands were isolated from 3 mice **(A)** or 3 patients **(B)** per tissue from the segments of the GI tract. Organoids were grown from the tissue and expanded to a maximum of 5 passages. RNA was isolated and conventional PCR and RT-qPCR were performed for *Tlr4*. RT-qPCR results were normalized to glyceraldehyde-3-phosphate dehydrogenase *(Gapdh)* expression and then to murine or human corpus glands. Control tissues in **(A)** are murine liver, brain and RAW264.7 cell line. Control tissues in **(B)** are THP1 human monocyte-like cell line, HEK293T cell line and human organoid pool. Bars represent mean ± SD, separate organoid lines from the indicated gut segments of 3 individual mice or 3 patients.

However, a closer look at the TLR4 signaling pathway also revealed species-specificity. LPS sensing by TLR4 is either completely facilitated or at least augmented by the accessory proteins CD14, MD2 and LBP. RNA-seq data showed expression of *Cd14*, *Md2* and *Lbp* in mice, but only *CD14* was detected in human organoids, although in a much lower amount compared to mice (**Figure 21**). This led to the hypothesis that, although TLR4 is expressed in a similar pattern in both species, LPS stimulation may lead to NF-κB activation in murine gastric, ileal and colonic epithelium, but not in human GI epithelium.

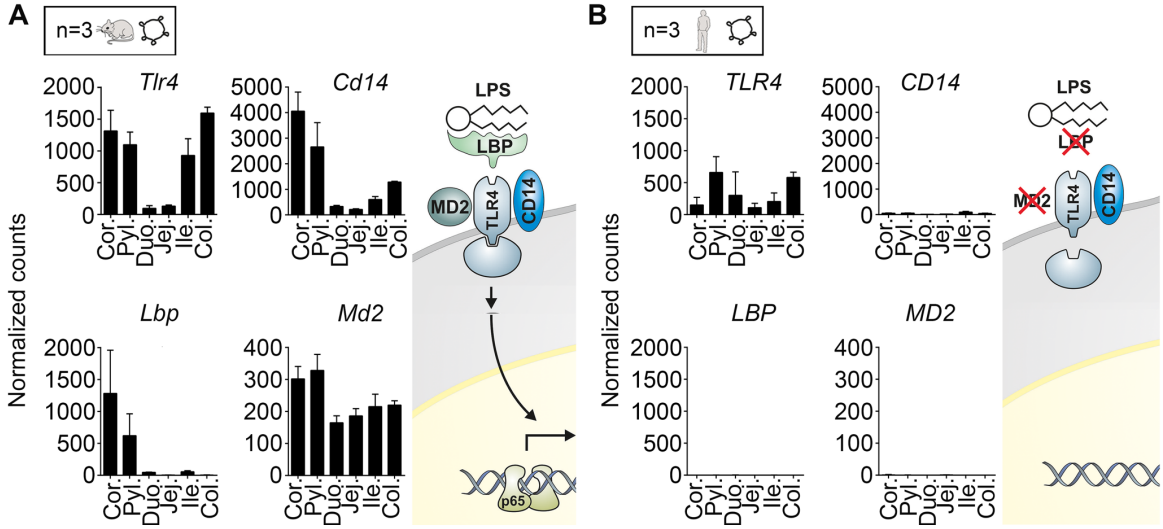


Figure 21: Expression of genes related to functionality of TLR4 signaling pathway shows segment- and species-specific patterns. Normalized gene counts of *Tlr4*, *Cd14*, *Lbp* and *Md2* in murine **(A)** and human **(B)** organoids and graphical summary of resulting hypothesis. Bars represent mean ± SD, separate organoid lines from the indicated gut segments of 3 individual mice or 3 patients.

To determine whether the TLR4 pathway is functional in the GI epithelium, expression of *IL-8* in humans and its analogue *Cxcl2* in mice were assessed by RT-qPCR, as well-known markers for inflammation. For practicability, 3 major GI regions: corpus, jejunum and colon were focused on; to represent the stomach, proximal intestine and colon. RT-qPCR analysis showed upregulation of *Cxcl2* in response to LPS, indicating that murine corpus and colon had functional TLR4, while jejunum did not (**Figure 22A**), mirroring the expression of *Tlr4*. In human organoids, *IL-8* was not induced, strengthening our hypothesis that TLR4 might need the accessory proteins CD14, LBP and MD2 to activate NF-κB (**Figure 22B**).

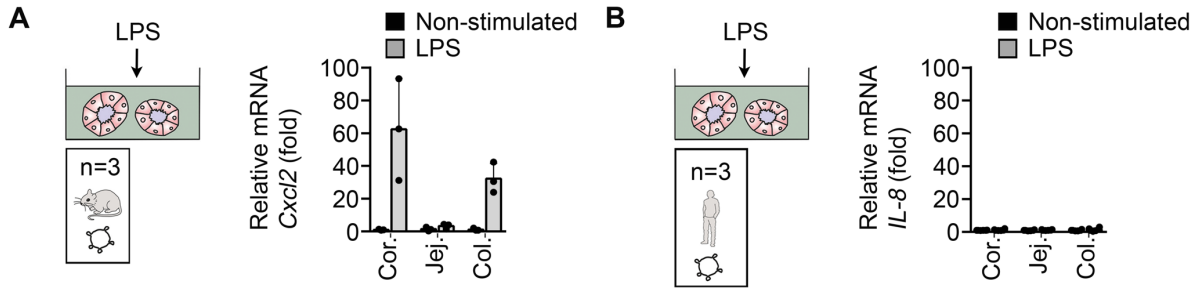


Figure 22: Functional analysis of TLR4 in murine and human GI tract. Separate lines of organoids from the indicated gut segments of 3 individual mice **(A)** or human patients **(B)** were exposed to 100 ng/ml LPS in the medium. After 2 h, cells were harvested, RNA was prepared and expression of *Cxcl2* **(A)** and its human analogue *IL-8* **(B)** were determined by RT-qPCR. Results were normalized to *Gapdh* and then to the average of non-stimulated controls. Bars represent mean ± SD, n = 3 biological samples for each group.

For murine *Tlr2*, the sequencing data showed a steep gradient with highest expression in the colon (**Figure 23A, left**), while human *TLR2* showed a much flatter gradient at lower expression levels with highest expression in the stomach (**Figure 23B, left**). Stimulation with the TLR2 agonist Pam3CSK4 demonstrated that both patterns were generally matched by *Cxcl2* or *IL-8* transcription (**Figure 23A and B, right**), but with a very low response in the human stomach (**Figure 23B, right**).

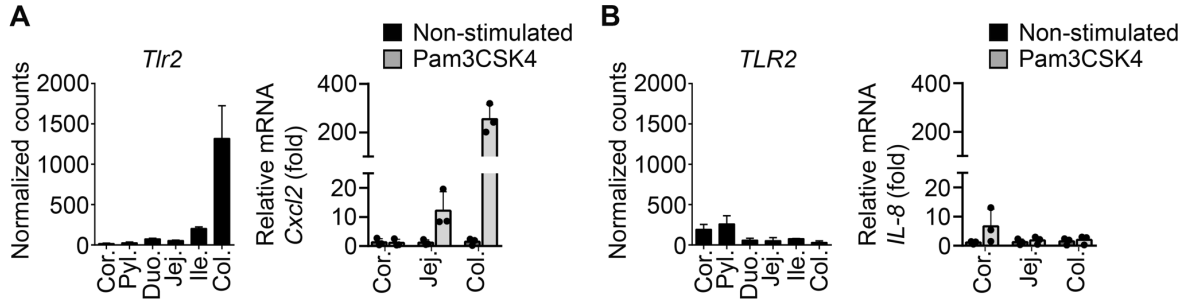


Figure 23: Expression and functionality of TLR2 in murine and human GI tract. Normalized gene counts of *Tlr2* in murine or human organoids **(A&B, left)** and functional test of TLR2 using 1 µg/ml Pam3CSK4 **(A&B, right)**. After 2 h, cells were harvested, RNA was extracted and expression of *Cxcl2* **(A, right)** or its human analogue *IL-8* **(B, right)** were determined by RT-qPCR. Results were normalized to *GAPDH* expression and then to the average of non-stimulated controls. Bars represent mean ± SD, separate organoid lines from the indicated gut segments of 3 individual mice **(A)** or 3 patients **(B)**.

Murine *Tlr5* and human *TLR5* were expressed in all gut segments, with much lower levels in the murine GI (**Figure 24A and B, left**). However, despite the similar gene expression, murine gastric organoids failed to respond to stimulation with flagellin, while jejunal and colonic organoids did. All human organoids responded readily (**Figure 24B, right**).

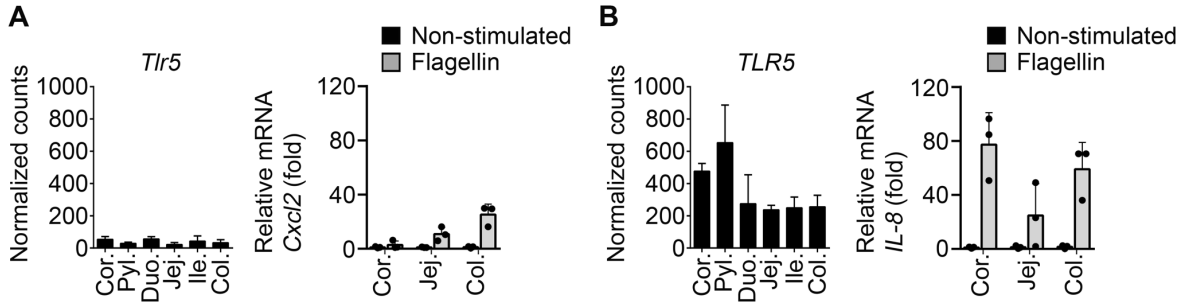


Figure 24: Expression and functionality of TLR5 in murine and human GI tract. Normalized gene counts of *Tlr5* in murine or human organoids **(A&B, left)** and functional test of TLR5 using 100 ng/ml flagellin **(A&B, right)**. After 2 h, cells were harvested, RNA was extracted and expression of *Cxcl2* **(A, right)** and its human analogue *IL-8* **(B, right)** were determined by RT-qPCR. Results were normalized to *GAPDH* and then to the average of non-stimulated controls. Bars represent mean ± SD, separate organoid lines from the indicated gut segments of 3 individual mice **(A)** or 3 patients **(B)**.

Taken together, RNA expression and functional testing of TLR2, 4 and 5 demonstrated gut segment- and species-specific patterning of these PRRs. While the functional data often mirrored the RNA expression data, the occasional absence of this congruency underlines the relevance of functional testing. Thus, not only expression but also function of PRRs in the gut is highly organized and segment-specific.

3.2.3 Murine gastric epithelial cells sense LPS from both apical and basal side

As explained before (Introduction-1.3, page 24 ff.) and also shown within the present chapter of this thesis, components of the innate immune system are distributed along the entire GI tract, but it has yet to be clearly identified how the body prevents continuous stimulation by apical stimuli in the gut. One prominent hypothesis states the body's mechanism to minimize the stimulation, is to confine pattern recognition to the basolateral or intracellular compartments [10, 95], which suggests that polarization of the cells plays a role in innate immune sensing. Because of the peculiar expression pattern, the species-specificity and general importance, further experiments were performed on TLR4 to characterize the polarity of LPS sensing in murine gastric epithelial cells.

To investigate stimulation at the basal side further, LPS was added to the medium of organoids in different concentrations (10 ng/ml, 100 ng/ml, 1000 ng/ml) for 2 hours or 100 ng/ml for differing durations (2, 4, 8 hours) as schematically depicted in **Figure 25A.** Following that, *Cxcl2* expression was measured by RT-qPCR. The peak of *Cxcl2* expression was measured after incubating with 100 ng/ml LPS for 2 hours (**Figure 25A**) and the expression started to decrease between 4 and 8 hours, suggesting a feedback mechanism for activated NF-κB pathway (**Figure 25B**). This suggests that murine gastric organoids reacted to the LPS stimulation in a concentration and time-dependent manner.

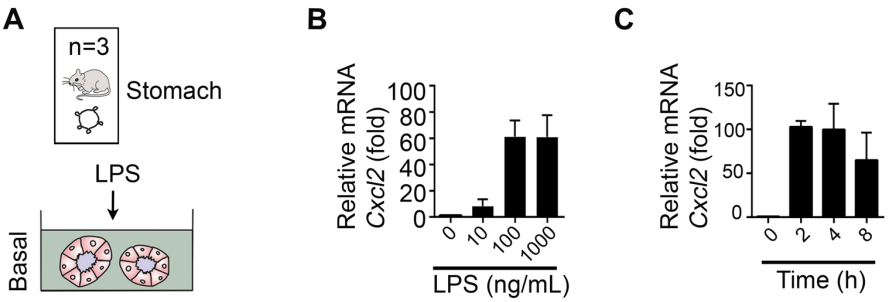


Figure 25: Murine gastric organoids are stimulated by LPS in a concentration and time-dependent manner. (A) Illustration depicting the LPS stimulation. Murine gastric organoids were stimulated with **(B)** indicated amount of LPS for 2 h or **(C)** 100 ng/ml LPS for indicated amount of time. Afterwards, cells were harvested, RNA was extracted and expression of *Cxcl2* was determined by RT-qPCR. Results were normalized to *Gapdh* and then to the average of non-stimulated controls. Bars represent mean ± SD, n = 3 biological samples for each group. *(Experiments were performed, and data were collected by Dr. Nina Wallaschek.)*

Since the expression of *Cxcl2* can be induced by many activators of NF-κB, organoids from two knockout mouse models were used to verify that the NF-κB activation indicated by *Cxcl2* expression was indeed induced by the LPS stimulation. Gastric organoids derived from a mouse lacking the expression of the TLR adaptor molecule, *Myd88* (*Myd88*$^{-/-}$), as well as gastric organoids derived from a *Tlr2* and *Tlr4* double knockout mouse (*Tlr2*/*Tlr4*$^{-/-}$) did not upregulate *Cxcl2* in response to LPS (**Figure 26**). This corroborates the regulation of *Cxcl2* in response to LPS stimulation of TLR4 in wild-type gastric organoids.

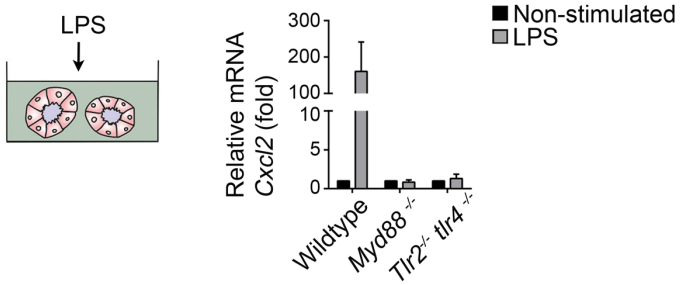


Figure 26: Murine gastric organoids from *Myd88*$^{-/-}$ and *Tlr2*/*Tlr4*$^{-/-}$ mice do not respond to LPS stimulation. Murine gastric organoids from wild-type, *Myd88*$^{-/-}$ and *Tlr2*/*Tlr4*$^{-/-}$ mice were stimulated with 100 ng/ml LPS. After 2 h, cells were harvested, RNA was extracted and expression of *Cxcl2* was determined by RT-qPCR. Results were normalized to *Gapdh* and then to the average of non-stimulated controls. Bars represent mean ± SD of cells from two technical replicates. *(Experiments were performed, and data were collected by Dr. Nina Wallaschek.)*

To evaluate whether LPS could be sensed from the apical side, various assays were set up. LPS-TLR4 signals via the NF-κB subunit p65, which is sequestered by its inhibitor in the cytoplasm when inactive, and is free to translocate to the nucleus upon activation [179]. To investigate apical stimulation, cells from murine gastric

organoids were seeded as 2D monolayers, stimulated with 100 ng/ml LPS for 2 hours, fixed and immunostained for p65. Image acquisition revealed nuclear translocation of p65 (**Figure 27A**). Quantification of cells with nuclear p65 showed dependency on LPS concentration (**Figure 27B**) and time, with a slight tendency for an oscillatory pattern (**Figure 27C**), similar to previous observations in cell lines [180]. Due to the oscillatory pattern, it is likely that all cells will have nuclear p65 at some point in time, but if cells are fixed and stained for p65, only a fraction shows nuclear p65 at that point of time.

To exclude an effect of the 2D or 3D growth conditions, polarity of the response was tested in two additional settings. For 3D cultures, microinjection of LPS was performed to either apical or basal side of the organoids *in situ* in 3D. Validation of the microinjection with FITC-labeled dextran showed that FITC-dextran mainly remained inside the injected organoids, whilst it quickly diffused throughout the extracellular matrix when it was placed outside of the organoids (**Figure 27D**). Following that, LPS was microinjected into the lumen of or outside of the organoids, in each case close to the epithelial layer. RT-qPCR data showed that stimulation at both the apical side (inside organoids) and basal side (outside organoids) led to expression of *Cxcl2* (**Figure 27E**). For 2D culture, cells from organoids were seeded in transwells and 100 ng/ml LPS was added either to the upper or lower compartment. Stimulation from either compartment induced upregulation of *Cxcl2* (**Figure 27F**). Due to differences in diffusion dynamics, it is not possible to compare the TLR4 activation between the two sides in absolute terms. However, it can be concluded that murine gastric TLR4 can sense both basally and apically administered LPS.

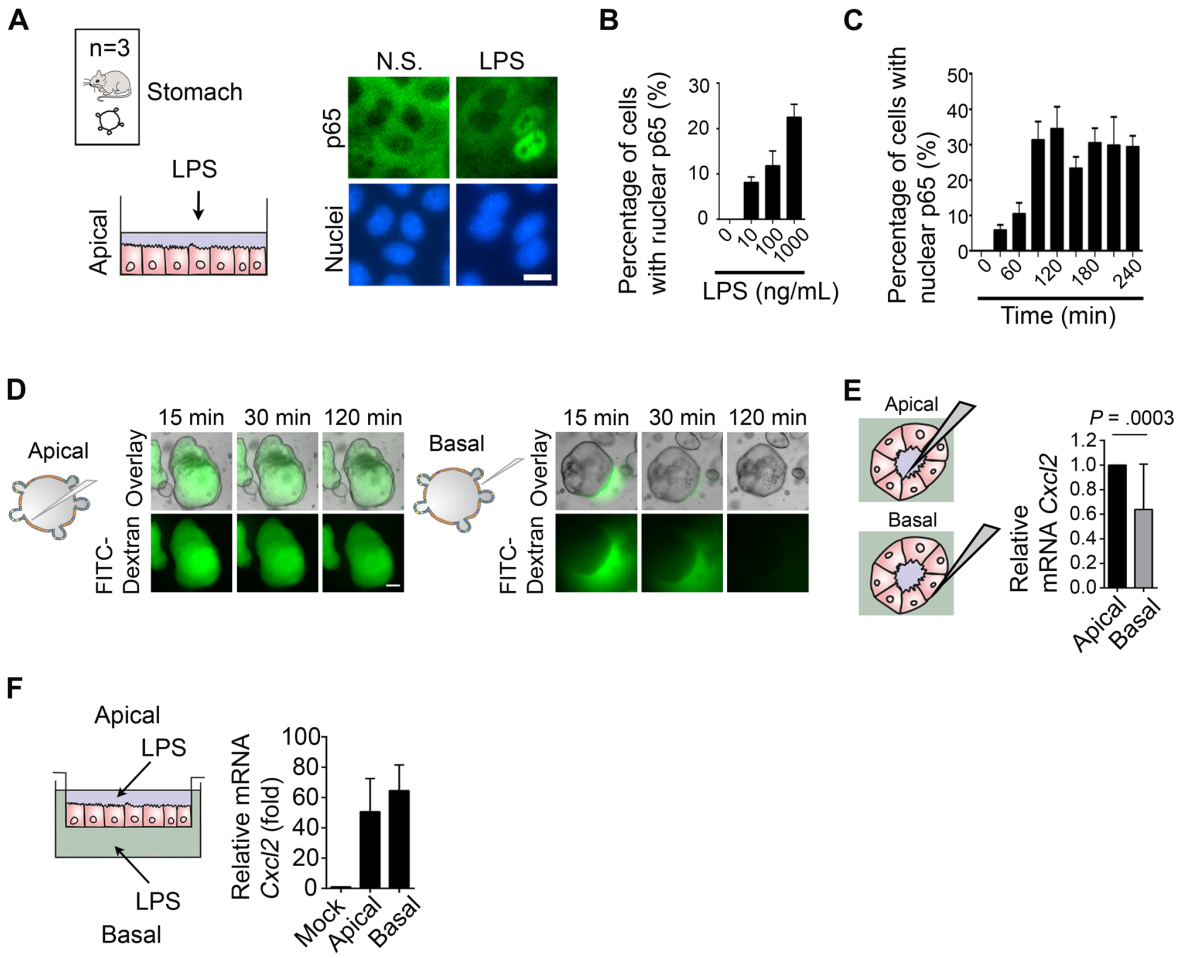


Figure 27: Murine gastric epithelial cells have responsive TLR4 at both the basal and apical side. (A-C) Cells from murine gastric organoids were seeded on conventional cell culture plates to form a 2D monolayer of cells, stimulated with LPS, fixed and stained for p65 and nuclei. **(A)** Representative image of p65 nuclear translocation in murine gastric epithelial cells. N.S.: non-stimulated. Scale bar: 10 µm. Cells with nuclear p65 were counted and are presented as percent of total cells after being stimulated with **(B)** indicated amounts of LPS or **(C)** for indicated durations with 100 ng/ml LPS. **(D)** FITC-dextran was microinjected either inside or outside of the murine gastric organoids and its diffusion was observed by live cell microscopy. Scale bar: 200 µm. **(E)** Organoids were microinjected with 100 ng/ml starting concentration of LPS either inside or outside of the murine gastric organoids. After 2 h, cells were harvested and *Cxcl2* expression was detected by RT-qPCR. Results were normalized to *Gapdh* and each outside injection was normalized to the paired inside injection on the same plate. **(F)** Cells from murine gastric organoids were seeded onto transwells and 100 ng/ml LPS was added either to the upper compartment (apical stimulation) or the lower compartment (basal stimulation). After 2 h, expression of *Cxcl2* was determined by RT-qPCR. Bars represent mean ± SD, n = 3 biological samples for each group. *(Experiments on A-C were performed, and data were collected by Isabella Pierotti, D&E by Michael Kern and E&F by Carolin Niklas.)*

3.2.4 Human gastric epithelial cells do not sense LPS from apical or basal side

To examine whether human gastric cells have functional TLR4 on the apical side, human gastric organoids were grown either in 3D or their cells were seeded as 2D monolayers and stimulated with 100ng/ml LPS for 2 hours. As a positive control TNF-α, a common NF-κB-activating stimulus, was used in a concentration of 10

ng/ml. By adding either LPS or TNF-α to the medium, it was observed that only basally administered TNF-α induced *IL-8* expression as measured by RT-qPCR (**Figure 28**). Stimulation with LPS on the other hand did not induce *IL-8* expression, neither on basal nor apical side. To exclude an effect of 3D versus 2D growth, apical versus basal stimulation was also tested in transwells where the apical side of the cells face the lumen of the insert [88]. Transepithelial electrical resistance (TEER) to ion migration was measured to assess the integrity and permeability of the monolayers and TEER values reached between 390 and 540 Ω*cm^2 (data not shown), which was consistent with previously published measurements of the intact gastric monolayers [88]. When the *IL-8* transcription is measured by RT-qPCR, it was observed that only TNF-α administered to the lower compartment induced transcription of *IL-8* (**Figure 28B**). Therefore, human gastric epithelium is unresponsive to stimulation with LPS from both sides. It senses TNF-α only from the basal, not the apical side, which shows that the assays established within the scope of this thesis can in principle discriminate between the two sides.

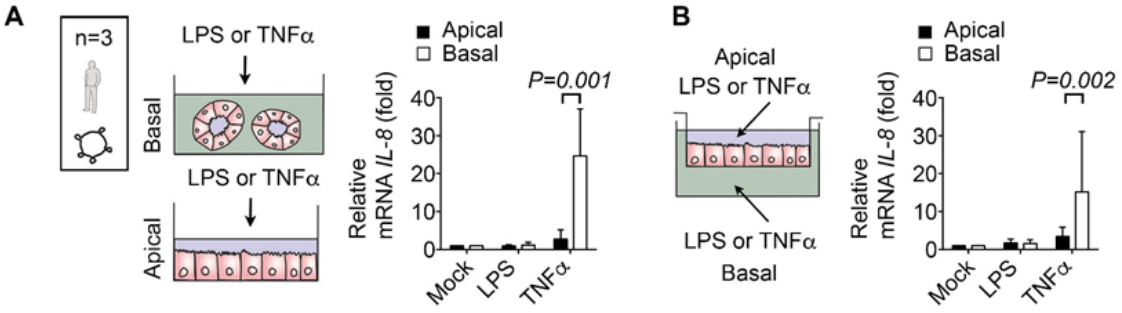


Figure 28: TLR4 does not sense LPS in human gastric epithelial cells regardless of the stimulation side. (A) Human gastric organoids were seeded in two conditions: For basal stimulation, organoids were cultured as 3D organoids; for apical stimulation, cells from organoids were seeded as 2D monolayers. Cells were treated with 100 ng/ml LPS or 10 ng/ml TNF-α. After 2 h, cells were harvested, and expression of *IL-8* was measured by RT-qPCR. **(B)** Cells from human gastric organoids were seeded into transwells and compounds as explained in **(A)** were added either to the upper compartment (apical stimulation) or the lower compartment (basal stimulation). After 2 h, cells were processed as in **(A)**. Bars represent mean ± SD, n = 3 biological samples for each group. *(Experiments were performed, and data were collected by Franziska Weiss.)*

3.3 Analysis of the factors regulating epithelial innate immunity

3.3.1 Investigating the role of exposure to microbial components

Intrigued by the distinct patterning of PRR expression, which did not follow the density of microbial colonization and possessed functional TLR4 receptors on the apical side, which would in principle allow sensing of gut luminal bacteria, it was hypothesized that PRR expression could be defined independently of contact to the microbiota. To address this question, a model that had never been in contact with microbial products was required. Germ-free mice were not ideal, since their food still contains LPS after sterilization by autoclaving. Therefore, organoids from murine embryos were generated, which – because of the sterile prenatal environment – had never been exposed to compounds of GI bacteria [181]. Moreover, such organoids have been shown to mature *in vitro* over time [91, 182, 183], thus with this model it was possible to address whether PRR patterning depends on stimulation with bacterial components while providing a purely epithelial model which had never been in contact with bacterial products.

3.3.1.1 Tissue identity is already encoded in tissue-resident stem cells in the embryo and retained in embryo-derived organoids

Since the embryonic GI tract is very small, it was not separated into six regions, but organoid lines were generated only from stomach and proximal intestine of embryonic mice at stage of embryonic day 12 (E12) and E16 (**Figure 29A**). Phenotypically, they resembled adult organoids (**Figure 29B**). Organoids grown from these regions were maintained in culture for a maximum of 5 passages and RNA was isolated for RNA-seq.

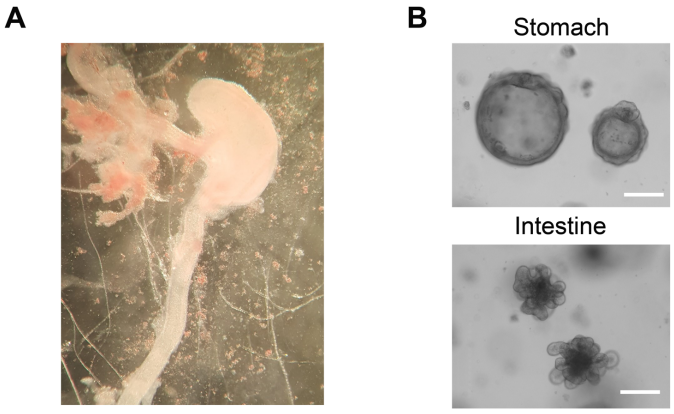


Figure 29: Gastric and proximal intestinal organoids were generated from murine embryos. (A) GI tract of a murine embryo at E16, depicting the stomach and proximal intestine. **(B)** Representative images of gastric and proximal intestinal organoids derived from murine E16 embryos. Scale bar: 200 µm.

When the embryo-derived organoids were analyzed together with adult-derived organoids, principal component analysis showed that embryo-derived gastric organoids clustered together with adult-derived gastric organoids and embryo-derived proximal intestinal organoids clustered together with adult-derived duodenal and jejunal organoids, demonstrating established regional identity and the expected *in vitro* maturation (**Figure 30A**). Particularly, embryo-derived stomach was closest in expression to adult pylorus and embryo-derived proximal intestine was closest in expression to duodenum (**Figure 30B**).

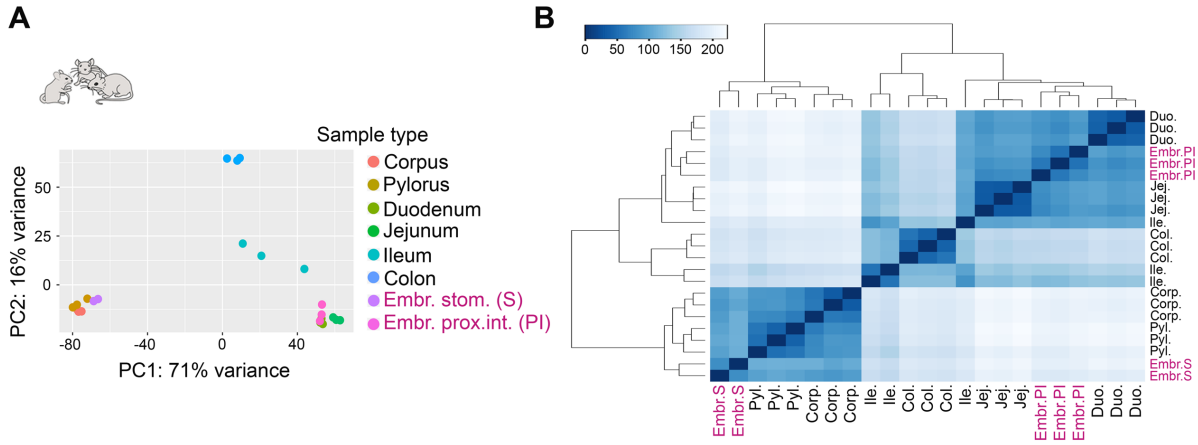


Figure 30: Murine embryo-derived gastric and proximal organoids clustered together with their adult counterparts. RNA was isolated from 2 organoid lines generated from embryonic stomach and 3 lines generated from proximal intestine (all E16). RNA was sequenced and analyzed with murine adult-derived organoids. **(A)** Principal component analysis and **(B)** hierarchical clustering of total mRNA sequencing, highlighting the closeness of the embryo-derived organoids to the adult-derived organoids.

Tissue-specific markers expressed in embryo-derived organoids confirmed regional identity and showed similar expression patterns to their adult counterparts (**Figure 31**, **supplementary Figure S 3**). Gastric markers such as *Muc1*, *Tff2* or *Cldn18* were expressed only in embryo-derived gastric organoids, while intestinal markers such as *Cdx1* or *Cdx2* were expressed in proximal intestinal organoids. The expression of *Lyz1*, a Paneth cell marker encoding lysozyme, in embryo-derived proximal intestinal organoids suggested that the organoids matured in culture and contained Paneth cells.

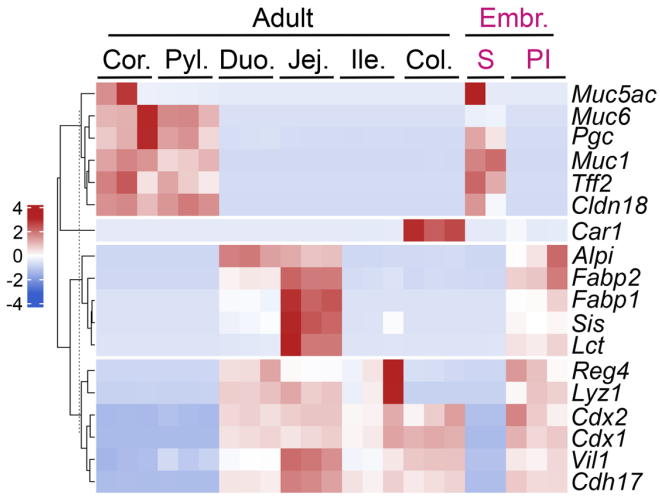


Figure 31: Murine embryo-derived organoids express tissue identity markers in a segment-specific manner. Heatmap displaying normalized and scaled gene counts of selected gastric and intestinal cell markers. For these genes, **supplementary Figure S 3** shows the bar graphs of expression values. (S: Stomach, PI: proximal intestine)

3.3.1.2 The epithelial innate immune barrier is primarily determined independent of microbial colonization

To identify the genes that characterize the GI border in the adult-derived organoids as well as embryo-derived organoids, differentially expressed genes (log2Fold Change > ±1 (2-fold change) and $p<0.05$) between embryonic stomach and proximal small intestine were compared to their adult counterparts. Lists of differentially expressed genes, comparing 6 murine adult gastric organoids (stomach = combined corpus and pylorus) and 6 murine adult proximal intestinal organoids (proximal intestine = combined duodenum and jejunum) or 2 murine embryonic (at E16) gastric organoids and 3 murine embryonic proximal intestinal organoids were generated. 2988 genes were significantly up- or downregulated by at least 2-fold both in adult- and embryo-derived organoids between two segments. Within the list of these genes,

the expected well-known genes for the GI border, such as *Cdx1*, *Muc5ac* or *Lyz1* (**Figure 32A, black gene names**) were present. In addition, many PRRs including *Tlr1*, *Tlr4* and *Nlrp6* were differentially expressed between the segments in both groups (**Figure 32A, blue gene names**). While various PRRs, such as *Tlr2*, *Tlr6* and *Nod1* were differentially expressed only in the adult-derived organoids, others such as *Tlr7*, *Tlr13* and *Nlrx1* were only expressed differentially in the embryo-derived organoids (**Figure 32A, blue gene names**). GO-term analysis of the 2988 common genes again highlighted the GO-term "Response to external stimulus" as well as many other GO-terms related to immune response in the top 15 (**Figure 32B**).

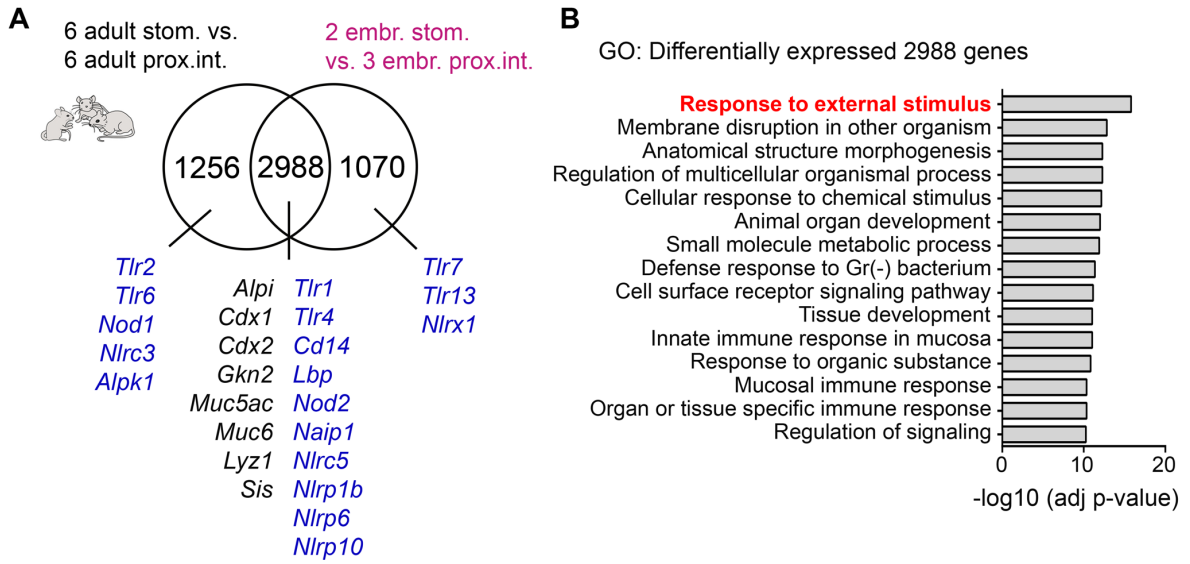


Figure 32: Expression of innate immune signaling components is segment-specific in murine embryo-derived organoids. (A) Lists of differentially expressed genes were generated, comparing 6 gastric organoids (combined corpus and pylorus) and 6 proximal intestinal organoids (combined duodenum and jejunum) or 2 embryonic gastric organoids and 3 embryonic proximal intestinal organoids (embryos at E16). Venn diagram displays number of differentially expressed genes ($p<0.05$ and log2Fold Change > ±1 (2-fold change)). Selected genes chosen for biological interest are listed. Black: Genes known for tissue identity. Blue: Genes known for immune function. **(B)** GO-term enrichment analysis (biological process) of 2988 overlapping adult versus embryonic differentially expressed genes, using the website g:Profiler. Scores of the top GO-terms, indicating the enrichment p-value.

Next, unsupervised clustering of all genes in the GO-term "Response to external stimulus" (1974 genes) placed the embryonic proximal intestinal organoids between adult duodenal and jejunal organoids, and embryonic gastric organoids next to the adult gastric organoids (**Figure 33A**). A heatmap depicting the same selected immune signaling components revealed that most of the genes including *Tlr4* were similarly expressed between adult- and embryo-derived organoids, but some PRRs, like *Tlr5* or *Nod2*, were expressed differently in embryo-derived organoids compared to their

adult counterparts (**Figure 33B, supplementary Figure S 4**). Taken together, expression patterns of most, but not all PRRs and immune signaling components are already developmentally encoded in the tissue-resident stem cells of the epithelium which had never been exposed to any microbiota.

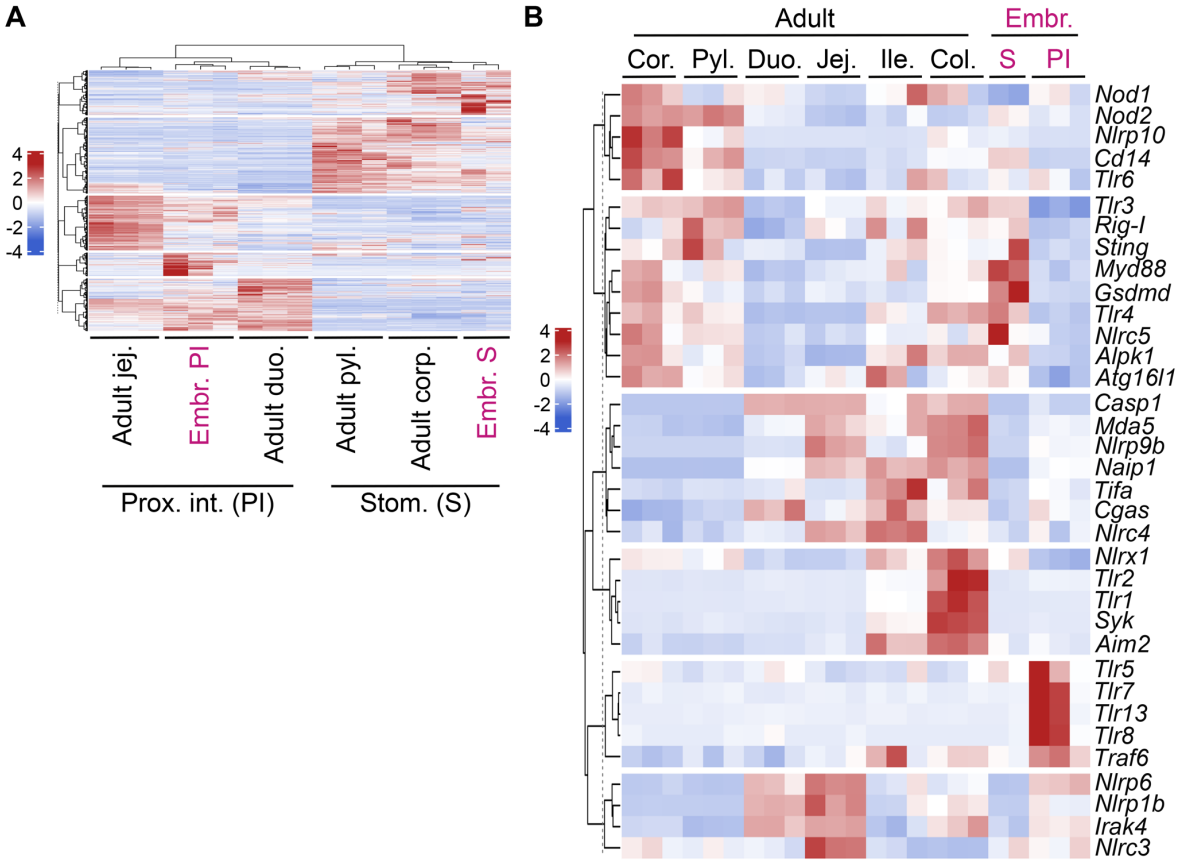


Figure 33: Many immune-related genes are defined developmentally. (A) Hierarchical clustering of all genes in the GO-term GO:0009605 "Response to external stimulus" (1974 genes). Heatmap depicts 2 embryonic gastric (S) organoids and 3 embryonic proximal intestinal (PI) organoids with their adult counterparts. **(B)** Heatmap displaying normalized and scaled gene counts of selected genes from GO:0009605 for adult and embryonic mouse- derived organoids. For these genes, **supplementary Figure S 4** shows the bar graphs of expression values.

3.3.1.3 Expression and function of TLR4 are developmentally and not environmentally defined

To analyze, whether this expression pattern also translates into the presence of functional proteins, once more TLR4 was chosen as an example. Expression levels of *Tlr4* in the RNA-seq were highly similar between adult- and embryo-derived organoids (**Figure 34A**). Conventional PCR confirmed that in all embryo-derived organoid lines, *Tlr4* was expressed in the stomach but not in the intestine (**Figure 34B**). Marker gene expressions showed the expected regional identity, with

organoids derived from the intestine expressing *Muc2* and organoids derived from the stomach expressing *Muc5ac* and *Muc6*.

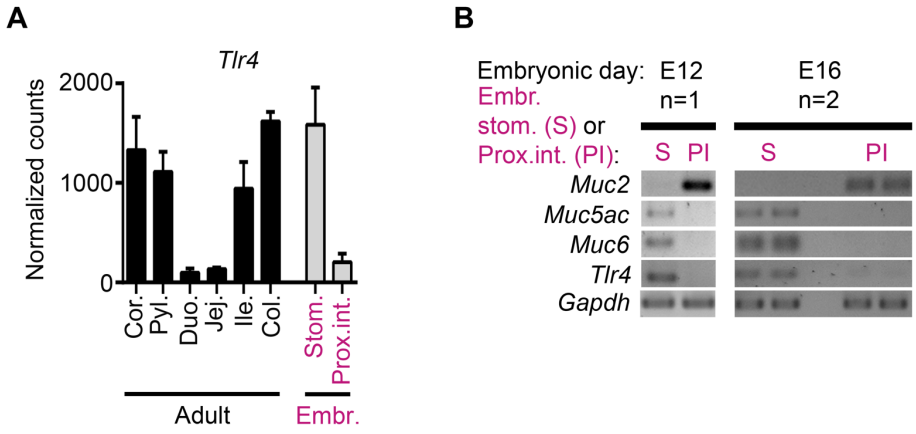
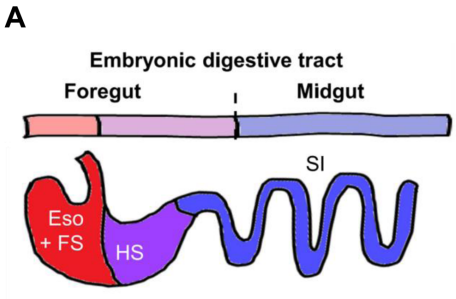


Figure 34: *Tlr4* expression is mainly developmentally programmed. (A) Normalized gene counts of *Tlr4* in murine organoids. Data from RNA-seq shown in **Figure 30** and **Figure 33**. Bars represent mean ± SD, n = 3 biological samples for each group, except embryonic stomach which is n=2. **(B)** RNA was prepared from embryo-derived organoids at E12 and E16. Expression of indicated genes was assessed by conventional PCR. Mucus genes are marker genes for stomach or proximal intestine. Each lane shows PCR results from the indicated organoid lines, n=3 stomach (S) and n=3 proximal intestine (PI).

After revealing that the *Tlr4* expression is regulated developmentally and connected to the tissue identity, I hypothesized that the segment-specific *Tlr4* expression would also be visible on the epigenomic level since it would be differently regulated in the developing stomach and small intestine. To address this question, one of the epigenome sequencing technologies, the assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) can be used to identify the accessible regions of the genomic DNA where transcription factors can bind for genes to be expressed [184] (for more detailed information, refer to **supplementary Figure S 9**).

For their published study, Banerjee *et al*. isolated the epithelial cells from three segments of the murine embryonic GI tract (esophagus + forestomach, hindstomach and small intestine) at different developmental stages as stated in **Figure 35A**, performed ATAC-seq and RNA-seq and provided the data to the Gene Expression Omnibus (GEO) repository (accession number: GSE115541) [185]. To visualize the chromatin accessibility at the *Tlr4* locus and *Tlr4* gene expression, these data were gathered and analyzed. Chromatin accessibility of the *Tlr4* locus in fore- and hind-stomach increased through developmental stages. In contrast in small intestine, the chromatin accessibility increased at E14, decreased at E16 and diminished at E18

(**Figure 35B**). Accordingly, the transcription level of *Tlr4* in fore- and hind- stomach increased to reach the maximum at E16, while the expression was not detected in small intestine at E18 and later (**Figure 35C**). Taken together, GI organ-specific chromatin accessibility and gene expression during development confirms the previous findings that the gene expression is regulated depending on the tissue identity and not on environmental cues in murine embryonic GI tract.



	Fore-stomach (FS)	Hind-stomach (HS)	Small intestine (SI)
ATAC-seq	E12	E12	E11
	E14	E14	E12
	E16	E16	E14
	P1	P1	E16
			E18
			P1
RNA-seq	E12	E12	E12
	E14	E14	E14
	E16	E16	E16
			E18

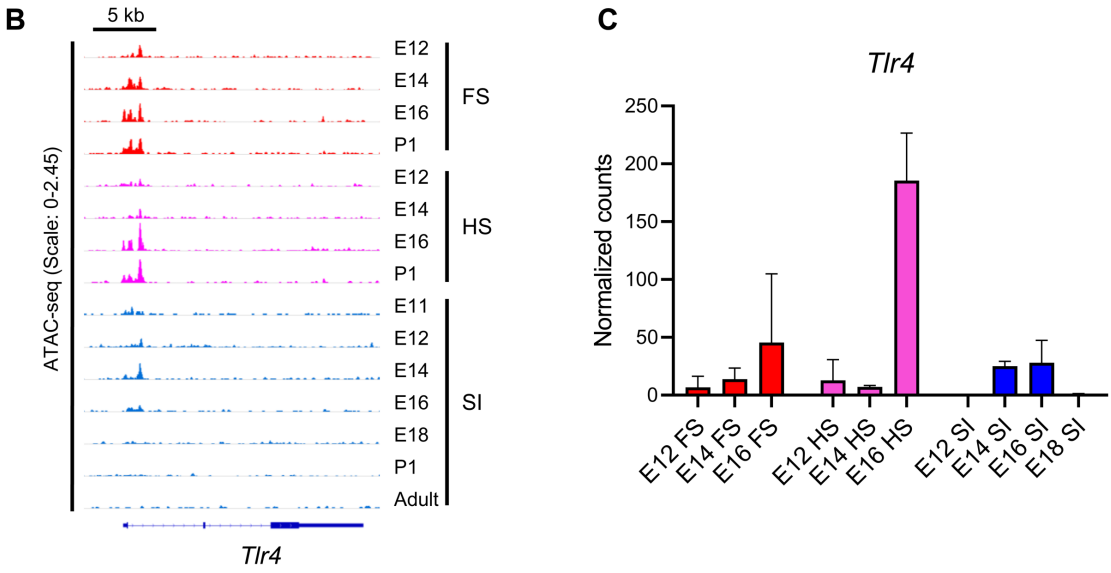


Figure 35: Chromatin access of *Tlr4* in murine GI tract is regulated prior to birth during development. All experiments were performed by Banerjee *et al*. [185]. The figures were curated from ATAC-seq and RNA-seq data under the accession number GSE115541 in GEO database. **(A)** Illustration of the embryonic digestive tract depicting different regions [185]. Epithelial cells were isolated and collected from forestomach (FS), hindstomach (HS) and entire small intestine (SI) of mice at different developmental stages from E11 to P1 for ATAC-seq and RNA-seq as depicted in the table. **(B)** Spatiotemporal chromatin accessibility at the 5′ locus of *Tlr4* gene, determined by ATAC-seq and visualized using IGV. For detailed information on ATAC-seq, refer to **supplementary Figure S 9**. **(C)** Normalized gene counts of *Tlr4* in different segments of embryonic digestive tract, determined by RNA-seq. Bars represent mean ± SD, n = 2 biological samples per timepoint.

Similar to their adult counterparts, the expression of *Tlr4* might not directly correspond to a functional protein. To assess whether expression of *Tlr4* was indicative of a functional receptor, organoids were stimulated by 100 ng/ml LPS for

2 h and *Cxcl2* expression was measured using RT-qPCR. The data revealed that embryonic proximal intestinal organoids did not react to stimulation with LPS, while embryonic gastric organoids expressed *Cxcl2* following stimulation with LPS, thus resembling their adult counterparts (**Figure 36**).

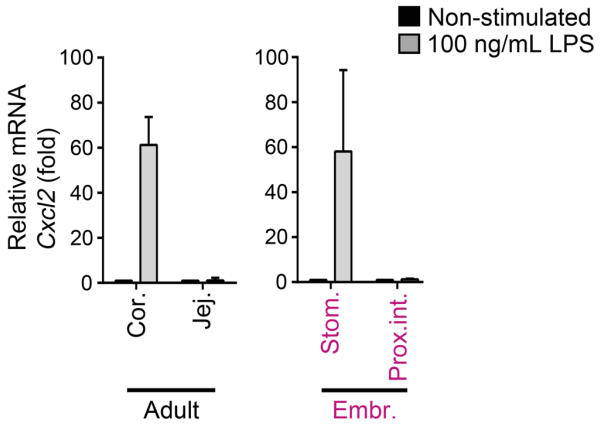


Figure 36: TLR4 is functional in embryo-derived gastric organoids and not proximal intestinal organoids. Embryo-derived and adult-derived organoids in culture were exposed to 100 ng/ml LPS for 2 h, cells were harvested, RNA was extracted and expression of *Cxcl2* was quantified by RT-qPCR. Results were normalized to expression of *Gapdh* and then to the paired non-stimulated control. Bars represent mean ± SD of 3 organoid lines for adult-derived organoids and 3 technical replicates for embryo-derived organoids.

Together, these results reveal that many, but not all, immune-related genes as well as functionality of TLR4 are determined as part of the tissue identity during development and this is independent of contact with bacterial components.

3.3.2 Investigating the role of cellular heterogeneity along the crypt-villus

The GI epithelium harbors many different epithelial cell types that self-organize into domains. For example, Paneth cells and stem cells are located at the bottom of the crypt, while the villus tip is populated with highly differentiated enterocytes. Several cell types, such as the Paneth cell, stem cells and sentinel goblet cells have been known to carry out specific innate immune cell functions [186, 187]. To analyze how the cellular differentiation in the GI tissue and the here used organoid models used in this thesis contribute to innate immunity, I analyzed RNA-seq data provided to open access GEO repository under the accession number GSE104803. This dataset was generated by Kozuka *et al*. [188], who compared the cells from the crypt and villus compartments of *ex vivo* isolations of small intestine (duodenum, jejunum and ileum) and colon (proximal and distal) to *in vitro* 2D monolayer cultures.

3.3.2.1 Expression patterns of innate immune genes are region-specific in isolated murine crypts and villi

To validate the identity of the sequenced crypts and villi, expression levels of known marker genes were analyzed (**Figure 37**) from the data provided by Kozuka *et al*. [188]. *Vil1*, a marker for differentiated enterocytes of villi, was expressed highly in small intestinal villi, while the stem cell marker *Lgr5* was mostly detected in the crypts of both small intestine and colon. Another stem cell marker olfactomedin 4 (*Olfm4*) which is specific to the small intestine [189] was detected only in the small intestinal and not in the colonic crypts, validating the identity of small intestinal samples. Carbonic anhydrase 1 (*Car1*), on the other hand, a marker for colonic epithelial cells was expressed only by the colonic segments (**Figure 37B**).

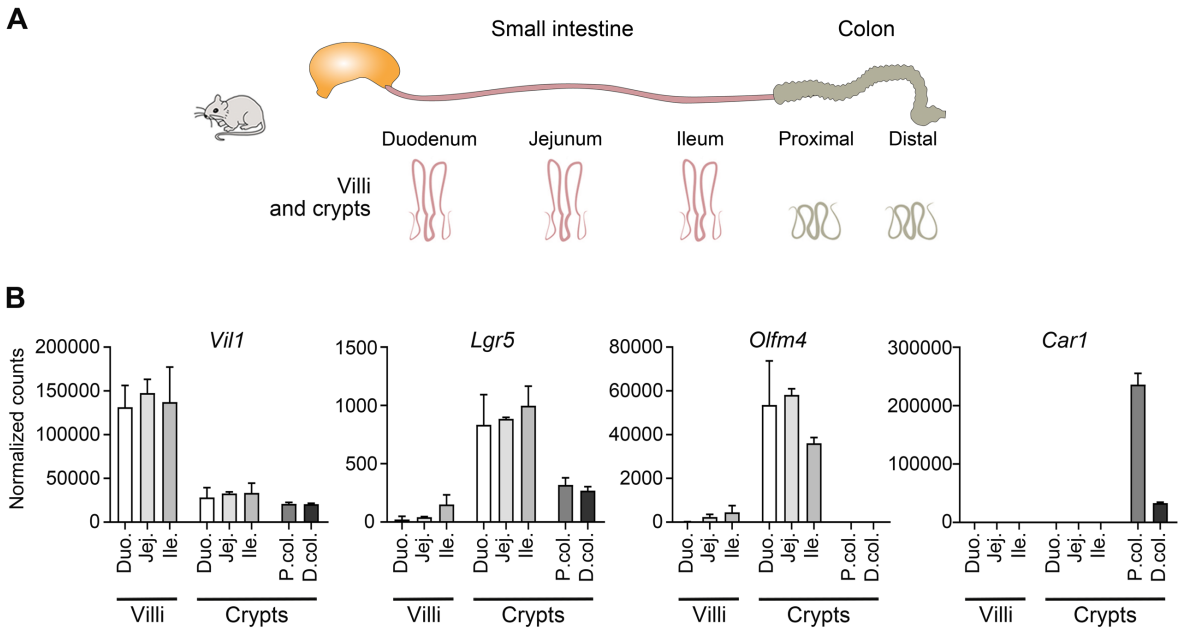


Figure 37: Crypts and villi of small intestine and crypts of colon expressed the tissue markers respectively. Experiments were performed by Kozuka *et al*. [188]. Murine small intestinal (duodenum, jejunum and ileum) villi and crypts and colonic (proximal and distal colon) crypts were freshly isolated, total RNA was extracted and sequenced. The data were generated from RNA-seq counts under the accession number GSE104803 in GEO database. **(A)** Illustration depicting the experiment design. **(B)** RNA-seq counts were normalized and graphs for respective genes were plotted. Bars represent mean ± SD, n = 2 biological samples for each group.

Having validated the tissue identity, the next step was the investigation of the genes related to innate immunity. The dataset here offered two levels of analyses: 1. Comparison along the cephalocaudal axis, which allows verification of the data obtained in the Results-3.2, (page 37 ff.) of this thesis; and 2. Comparison along the crypt-villus axis to understand possible heterogeneity within one gut segment.

The genes related to innate immunity chosen for this analysis were the genes previously shown to be region-specific along the cephalocaudal axis (Results-3.2, page 37 ff., refer to **Figure 19**). For the majority of the analyzed genes, cephalocaudal patterning and a decreased expression along the crypt-villus axis was observed, while an increase along the crypt-villus axis was rare (**Figure 38**, **supplementary Figure S 5)**. Among the genes with an increasing expression along the cephalocaudal axis were the cell surface *Tlrs*, *Tlr1*, *Tlr2*, *Tlr4*, *Tlr5;* several cytosolic receptors such as *Sting* and *Alpk1*; inflammasome making *Aim2* and *Naip1;* non-inflammasome making *Nod1*, *Nod2*, *Nlrp10* and *Nlrx1,* which were expressed more in the colon compared to the small intestinal segments (**Figure 38A**) and showed differences between proximal and distal colon. While *Nlrp10, Sting* and *Alpk1* were mainly expressed in the distal

colon, *Nod2* (similar to *Tlr2* and *Tlr5*) could be found mainly in the proximal colon. Conversely, the expression of the intracellular *Tlr*s – *Tlr3*, *Tlr9*, *Tlr12* and *Tlr13*, the *Rlr*s - *Rig-I* and *Mda5*, *Nlr*s not involved in the composition of inflammasomes - *Nlrc3* and *Nlrc5*, and inflammasome-related *Nlrp1b*, *Nlrp9b*, *Nlrc4*, *Nlrp6* and *Gsdmd* was highest in the small intestinal segments compared to colon, while *Nlrp6* did not show a substantial difference (**Figure 38B**). These data corroborated the overall finding from the first part of the thesis, that cell surface *Tlr*s were mostly expressed in the colon, while most of the inflammasome-related *Nlr*s were expressed mainly in the small intestine.

Regarding the heterogeneity within one gut segment, the data revealed that the mentioned intracellular *Tlr*s which had not been detected in the organoid samples prepared for the present thesis (see **Figure 19**) such as *Tlr9*, *Tlr12* and *Tlr13* were mostly expressed in the villi region of the small intestinal segments (**Figure 38B**). Similarly, inflammasome making *Nlr*s - *Nlrp1b*, *Nlrp9b*, *Naip1* and *Nlrc4* and the *Rlr*s - *Rig-I* and *Mda5* were expressed predominantly in the villus region of the small intestinal epithelial layer (**Figure 38B**). On the other hand, *Cgas* and *Pycard* were expressed mostly by the crypt region of the small intestinal segments (**Figure 38C**). This suggests that expression of the TLRs found in cell compartments and most of the inflammasome-related PRRs are expressed more in the villus region of the small intestinal segments.

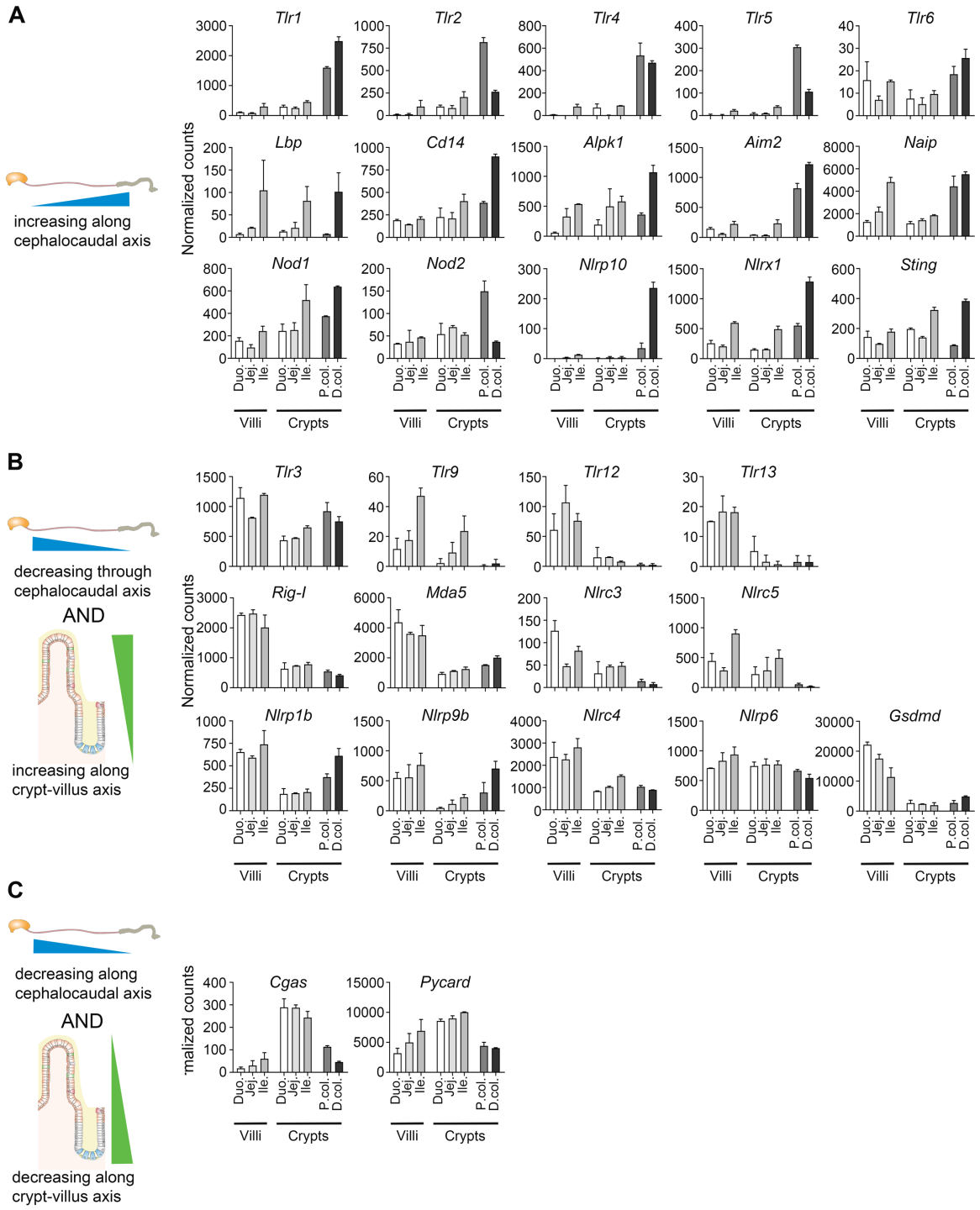


Figure 38: PRRs are expressed in different patterns along the GI tract and crypt-villus axis. Experiments were performed by Kozuka *et al*. [188]. RNA from murine small intestinal (duodenum, jejunum and ileum) villi and crypts as well as colonic (proximal and distal) crypts was sequenced and RNA-seq counts were submitted to GEO database (accession number GSE104803). The data presented here were generated after normalization of the RNA-seq counts. Various genes related to innate immunity are either expressed in a pattern **(A)** increasing or **(B&C)** decreasing along the GI tract, while **(B)** increasing or **(C)** decreasing along the crypt-villus axis of the small intestine. Bars represent mean ± SD, n = 2 biological samples for each group. For expression of all the genes related to innate immunity in **Figure 19**, refer to **supplementary Figure S 5.**

3.3.2.2 Cellular differentiation along the crypt-villus axis influences the expression of genes related to innate immunity

To model upper and lower parts of the intestinal crypts, Kozuka *et al*. established and characterized 2D monolayer epithelial cell cultures on transwells from murine small intestine (duodenum, jejunum and ileum) and colon (proximal and distal) [188]. Cells of the small intestinal segments were incubated with W(2,5x)ENRY medium (W: WNT, E: EGF, N: Noggin, R: R-spondin1 Y: Rho-kinase inhibitor) and colonic cells were incubated in WENR medium. Cells cultured in this medium containing the full set of growth factors comprised mostly the undifferentiated progenitor cells and thus modeled rather the lower part of the small intestinal villi and partially the crypts (except the highly differentiated Paneth cells found in the small intestinal crypts), and less differentiated cells of the colonic crypts. Differentiation of the cultured cells was achieved by changing the medium conditions. Small intestinal cells were cultured in ETB medium (E: EGF, T: thiazovivin, BMP-4: bone morphogenic protein 4), while WNT and R-spondin1 were removed from the medium of colonic cells. Cells cultured in this way represented the small intestinal villi and upper part of the colonic crypts with more differentiated cells (A schematic depiction of the experimental setup can be found in **Figure 39A and C**, for detailed culture conditions and durations refer to **supplementary Figure S 6).**

To confirm the differences in the differentiation state of the samples, well-known markers for highly differentiated cells were analyzed in this thesis. The undifferentiated cells of small intestinal and colonic segments expressed a higher amount of *Lgr5*, indicating stem cell identity. The expression of the differentiated enterocyte markers *Alpi*, *Apoa1* and *Vil1* on the other hand increased with differentiation of the 2D cultures (**Figure 39B and D**), confirming the differentiated character of the mature enterocytes.

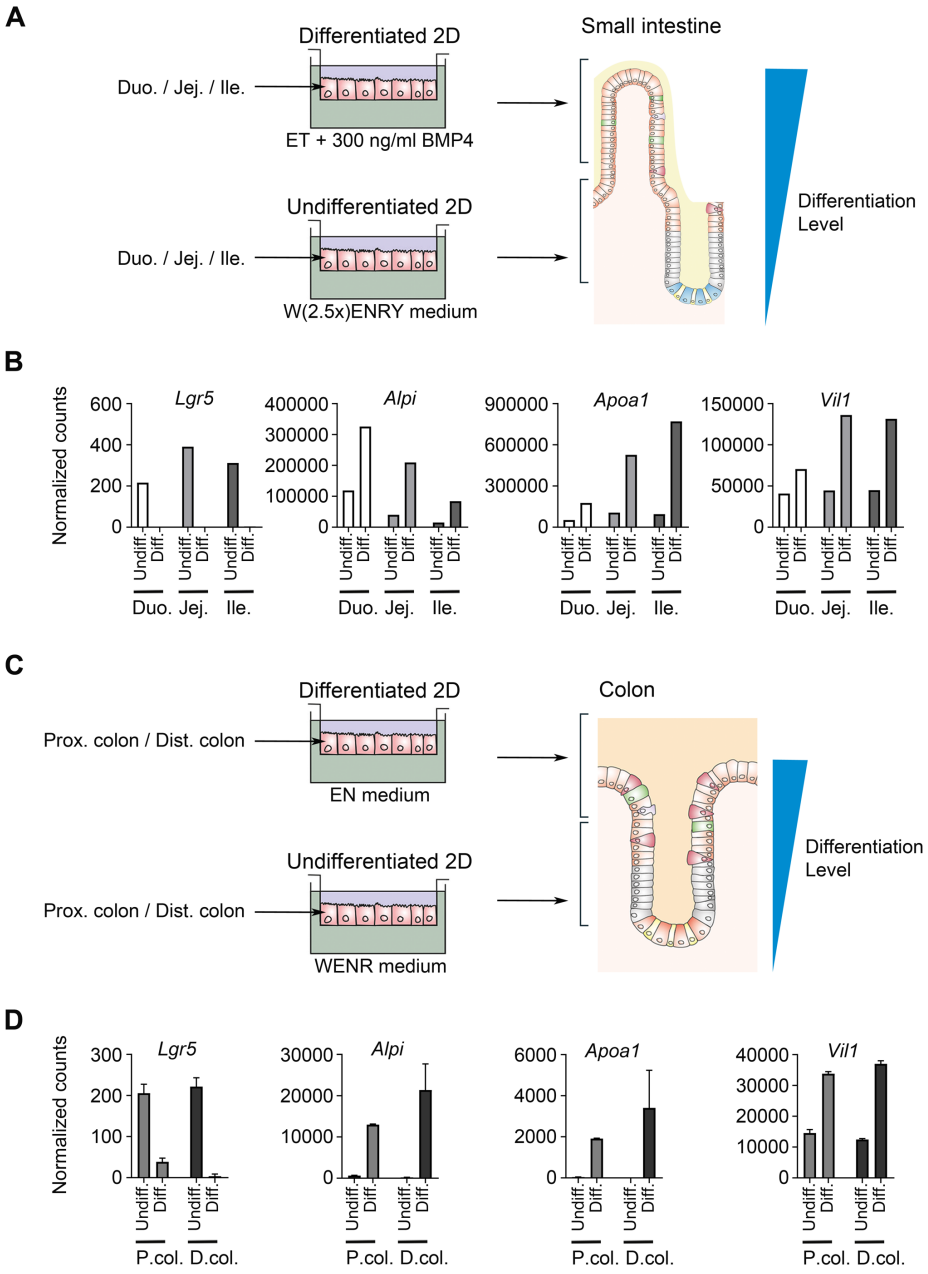


Figure 39: 2D monolayer cultures of undifferentiated and differentiated cells from murine small intestine and colon. All experiments were performed by Kozuka *et al*. [188]. The small intestinal and colonic cells were seeded on transwells and either differentiated or not differentiated by using different medium conditions. The cells were collected at day 7-8 after seeding and RNA was extracted to be sequenced. For this thesis, the RNA-seq read counts were accessed under the number GSE104803 in GEO database and normalized gene counts were plotted. **(A,C)** Illustration of 2D monolayer seeding of small intestinal (duodenum, jejunum, ileum) **(A)** and colonic (proximal and distal colon) **(C)** cells on transwells and the differentiation process. Undifferentiated 2D monolayers resemble the less differentiated crypt section (except the highly differentiated Paneth cells in small intestine), differentiated 2D monolayers represent the upper crypt section and villi. **(B,D)** Normalized gene counts of selected differentiation markers for small intestinal **(B)** and colonic **(D)** cells. Bars represent mean ± SD, n = 2 biological samples for each group except small intestinal samples which are n=1. W: WNT (1x for colon, 2.5x for small intestine), E: EGF, N: Noggin, R: R-spondin1, Y: Rho-kinase inhibitor, T: thiazovivin, B: BMP-4. Detailed culture conditions and duration can be found in **supplementary Figure S 6**.

Using this dataset, the expression of the previously identified genes related to innate immunity (see **Figure 19**, page 41) were analyzed (**Figure 40, supplementary Figure S 7**). Genes related to the inflammasome pathway increased upon differentiation of both small intestinal and colonic cells instead (**Figure 40A**). *Nlrp1b, Nlrp9b*, *Naip1*, *Gsdmd* and *Nlrx1* were expressed in undifferentiated small intestinal cells and increased in the differentiated state. In segments from the colon however, *Nlrp1b* and *Nlrp9b* were mainly expressed in differentiated cells, while the expression levels in undifferentiated cells were very low. In differentiated proximal and distal colon, the expression level of *Nlrp10* was the highest compared to every other segment and resembled the expression in the colon crypts (**Figure 38**). Conversely, the expression of *Nlrc3*, *Pycard*, *Cgas*, *Sting*, *Nod1* and *Nod2* decreased upon differentiation (**Figure 40B**). Although the decrease was slight for *Nlrc3* and *Nod2*, *Pycard* expression vastly decreased especially in differentiated jejunum and ileum. Similarly, expression levels of *Cgas* and *Sting* decreased highly upon differentiation of both proximal and distal colon (**Figure 40B**), corroborating with the expression levels in colonic crypts (**Figure 38**). Differentiation of the cells affected the expression levels of *Tlr5* and *Nlrp6* differently for small intestine and colon. While expression of these genes decreased drastically upon differentiation in small intestinal segments, it increased in differentiated colonic cells (**Figure 40C**). These data suggest that cellular differentiation and enrichment of certain cell populations has an impact on the expression of genes related to innate immunity in a GI segment-specific manner.

Taken together, the segment-specific expression patterns of innate immune genes observed in organoid samples in the previous chapter were also observed in freshly isolated murine tissues. Additionally, these genes showed differential expression along the crypt-villus axis of small intestinal segments where the expression of intracellular *Tlr*s and other cytosolic PRRs were concentrated mostly at the villus region. Furthermore, 2D monolayer cultures enriched for either less differentiated cells or more mature enterocytes show different expression levels of innate immune-related genes which might help to mimic the expression states in the actual tissue.

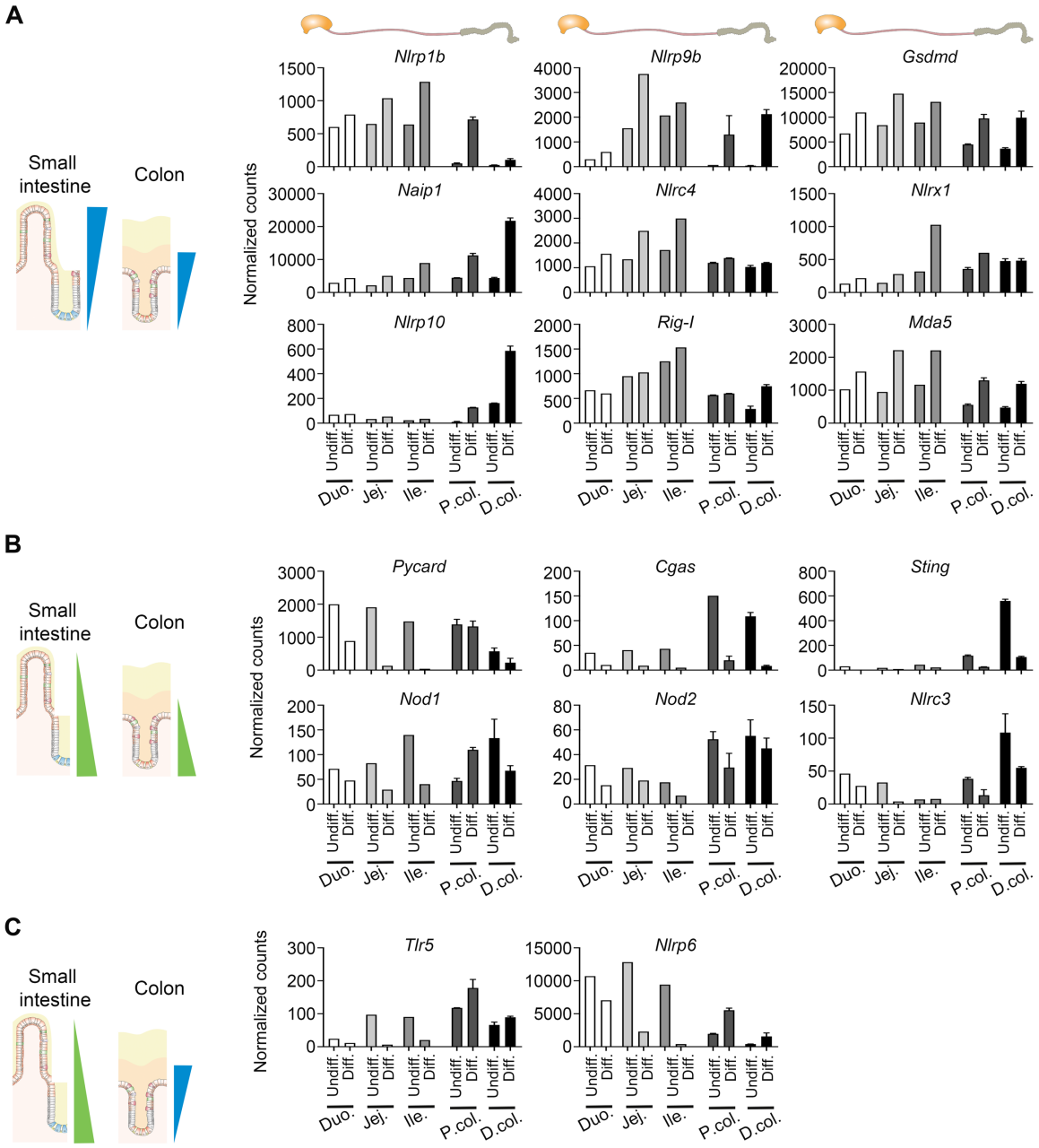


Figure 40: Expression levels of selected PRRs according to the differentiation state of the small intestinal and colonic 2D monolayers. Experiments were performed by Kozuka *et al*. [188] as explained in **supplementary Figure S 6**. The data were generated from RNA-seq counts under the accession number GSE104803 in GEO database. Small intestinal and colonic cells were seeded on transwells and cultured in different medium conditions. Selected genes were either expressed more in **(A)** differentiated state or **(B)** undifferentiated state. **(C)** Genes were expressed more in undifferentiated small intestinal monolayers and differentiated colonic monolayers. Bars represent mean ± SD, n = 2 biological samples for each group except small intestinal samples which are n=1. Expression of all the innate immune-related genes in **Figure 19** were displayed in **supplementary Figure S 7**.

3.3.3 Investigating the role of tissue-specific transcription factor CDX2

In the first part of this thesis, it was demonstrated that in 3D organoids, numerous genes related to innate immunity were expressed primarily in either stomach or intestine (Results-3.2, page 37 ff. **Figure 19**, **supplementary Figure S 2A**). These data suggest an organ-specific organization of immune-related genes along the cephalocaudal axis, which is a part of the regional patterning.

This regional patterning is established during development and controlled by genes regulating the identity of the GI segment. As explained in more detail previously (Introduction-1.1.2, page 11 ff.), SOX2 and CDX2 are two transcription factors which function antagonistically in the gut tube and shape the stomach-intestine boundary during the early development (**Figure 6**, page 13). SOX2 defines the anterior axis of the GI tract, while CDX2 is found in the posterior part in developing embryos. Hypothetically, the tissue-specific expression of the genes related to innate immunity might be related to the expression of these master regulators. Candidate genes to be regulated by CDX2 are genes not expressed in the stomach, but highly expressed in the intestine. Examples for these expression patterns are *Tlr2* and inflammasome-related *Naip1*, *Nlrp6*, *Aim2* and *Casp1* (**Figure 41**).

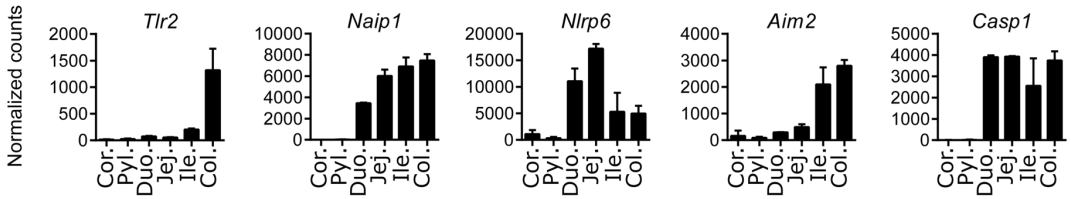


Figure 41: Expression of genes showing organ-specific patterning in murine organoids. Normalized gene counts of *Casp1*, *Tlr2*, *Naip1*, *Aim2* and *Nlrp6* in murine organoids. Bars represent mean ± SD, separate lines of organoids from the indicated gut segments of 3 individual mice.

Thus, to elucidate a possible contribution of CDX2 to the development of the expression patterns, the expression of such exemplary genes should be analyzed after *Cdx2*-knockout or *Cdx2* overexpression.

Within this section, published data provided by Simmini *et al*. [73] and Saxena *et al*. [190] to GEO repository were accessed and analyzed, to address the question how the lack of intestine-specific CDX2 would change the expression of the genes related to innate immunity detected in an organ-specific manner.

3.3.3.1 Expression of various intestine-specific PRRs is CDX2-dependent

In their study, Simmini and his colleagues demonstrated that organ-specific identity of GI stem cells is regulated by CDX2 [73], which is an intestine-specific transcription factor not expressed by other parts of the GI tract including the stomach. Genetic ablation of *Cdx2* in stem cells of small intestine collected from adult *Cdx2-/fl/Lgr5-EGFP-Ires-Cre^{ERT2}* mice led to their conversion to gastric lineage *in vitro* (**supplementary Figure S 8**). Cells started to express the markers of gastric lineage and the organoids derived from them could not grow in small intestinal medium, but only in the medium for gastric organoids. However, ectopic *Cdx2* expression in the murine gastric organoids by itself was not sufficient to induce an intestinal conversion. Nevertheless, this study suggested that CDX2 is a central regulator of stem cell identity in the murine small intestine [73]. To determine whether changes in the tissue identity also alter the expression of organ-specific genes related to innate immunity, the RNA-seq data Simmini *et al*. provided to GEO repository (accession number: GSE51751) were analyzed. This dataset includes the gene counts of the RNA-seq data of murine small intestinal organoids with genetically ablated *Cdx2* (*Cdx2^{null}* SI), control small intestinal organoids (Control SI), gastric organoids with overexpressed *Cdx2* (*Cdx2*+ Sto) and control gastric organoids (Control Sto) to test the regulation capacity of CDX2 as a transcription factor (**supplementary Figure S 8**).

To identify significant changes upon *Cdx2* loss in small intestinal organoids and gain in gastric organoids, DEG analysis was performed for this thesis. In comparison between Control SI and *Cdx2^{null}* SI organoids, 283 genes were significantly up- or downregulated by at least 2-fold. Similarly, between the Control Sto and *Cdx2*+ Sto organoids, 154 genes were found to be differentially expressed. 18 of the analyzed genes showed *Cdx2*-dependent alterations in the expression in both comparisons (**Figure 42A, see the gene list in supplementary Table S 1)**. Among these 18 genes, the PRRs *Nlrp6* and *Aim2* were also detected, and their expression decreased more than 2-fold in *Cdx2^{null}* SI organoids compared to the Control SI organoids and increased in *Cdx2*+ Sto organoids compared to the Control Sto organoids (**Figure 42B**). Similarly, *Naip1* expression decreased upon ablation of *Cdx2* in small intestinal organoids; however, unlike the other two genes, its expression decreased significantly, but less than 2-fold in *Cdx2*+ Sto organoids (**Figure 42B**). These three genes also showed organ-specific patterning as explained previously (**Figure 41**).

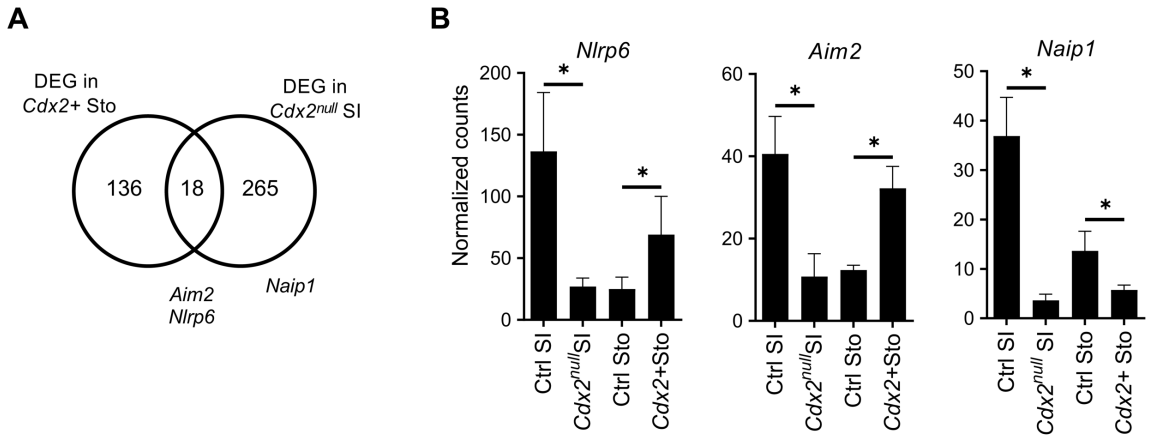


Figure 42: Expression of several PRRs is affected by the expression of *Cdx2*. Experiments were performed by Simmini *et al*. [73]. *Cdx2*^{null} small intestinal organoids (*Cdx2*^{null} SI), control small intestinal organoids (Control SI), control gastric organoids (Control Sto) and *Cdx2* expressing gastric organoids (*Cdx2*+ Sto) were generated as explained in **supplementary Figure S 8**. The data were curated from normalized RNA-seq counts provided to the GEO repository under the accession number GSE51751. **(A)** Lists of differentially expressed genes (DEGs) were generated, comparing 4 Control Sto and 4 *Cdx2*+ Sto organoids, or 4 Control SI to 4 *Cdx2*^{null} SI organoids. Venn diagram displays the numbers of DEGs in both groups ($p<0.05$ and log2Fold Change > ±1 (2-fold change)). For the complete gene list, see **supplementary Table S 1**). **(B)** Normalized gene counts of *Nlrp6*, *Aim2* and *Naip1*. For each gene, two-sided Mann-Whitney test was performed for statistical analysis. Asterisks denote significant changes in gene expression ($p<0.05$).

Taken together, these data indicated that CDX2 as a transcription factor might directly play a regulatory role in the expression of the intestine-specific PRRs *Nlrp6*, *Aim2* and *Naip1*.

3.3.3.2 Analysis of published datasets suggests a central role for CDX2 in regulation of Nlrp6 and Naip1

To gain a deeper mechanistic understanding of the role of *Cdx2* in the regulation of these genes, I searched for sequencing data of a chromatin immunoprecipitation of CDX2 (ChIP-Seq) together with the accessible chromatin regions (ATAC-seq). In such experiments, DNA-bound proteins such as transcription factors are immuno-precipitated by their specific antibodies and the sequence that the protein binds to can be coprecipitated to be identified by sequencing. ATAC-seq on the other hand, detects the accessible chromatin regions by sequencing the regions of the genome which are tagged by a transposase inserting sequencing adapters into accessible regions of the chromatin (for more detailed information, refer to **supplementary Figure S 9**). Since CDX2 is an essential transcription factor for intestinal specification and it maintains the chromatin accessibility at the loci of tissue-specific genes [69, 191], a combined analysis of ATAC-seq and ChIP-seq of CDX2 as well as histone

marks could fully elucidate the regulatory role of CDX2 on the corresponding DNA regions. This dataset was published in the GEO repository (accession number: GSE98724) [190] which I retrieved and re-analyzed within this thesis. To generate the data, Saxena *et al*. isolated murine jejunal villus cells from wild-type and tamoxifen-treated *Cdx2$^{f/f}$; Villin-Cre^{ERT2}* mice to induce loss of CDX2. On these cells, they performed ChIP-seq for CDX2 and several histone marks (including H3K4me2, H3K4me3 and H3K27ac), in addition to ATAC-seq and RNA-seq [190]. H3K4me3 and H3K27ac bound regions indicate the promoter region of actively transcribed genes, while high occupancy of H3K4me2 together with H3K27ac indicates the active enhancers. As a result of tamoxifen-treatment, expression of villus-specific genes such as sucrose isomaltase (*Sis*) and intestine-specific homeobox (*Isx*) were reduced considerably [190, 192].

To illustrate the following multi-omics data more clearly, *Sis* gene which is drastically affected by *Cdx2* ablation in transcriptional and epigenetic level was chosen as an example (**Figure 43A,** also refer to **supplementary Figure S 9** for more information on chromatin structure). Integrative Genomic Browser (IGV) was used to visualize the ChIP- and ATAC-seq data. Approximate location of the promoter was marked by high occupancy of H3K4me3 and H3K27ac in wild-type cells (**Figure 43A, WT orange and green tracks**). Similarly, the enhancer regions were characterized by strong ChIP-seq peaks for H3K4me2 and H3K27ac (**Figure 43A, WT magenta and green tracks**). ATAC-seq track of wild-type cells depicted the accessible parts of the chromatin with strong peaks which corresponded to these promoter and enhancer regions (**Figure 43A, WT blue track**). These regions also coincided with CDX2 binding sites determined by the CDX2 ChIP-seq (**Figure 43A, WT black track**). Upon ablation of *Cdx2*, the density of the active histone mark H3K27ac on both promoter and enhancer regions decreased compared to the wild-type villus cells, reflecting a change in the chromatin state (**Figure 43A, *Cdx2*-/- green track**). The promoter region occupied lower levels of H3K4me3 (**Figure 43A, *Cdx2*-/- orange track**), while the enhancers showed lower levels of H3K4me2 (**Figure 43A, *Cdx2*-/- magenta track**). Markedly lower signals were detected in ATAC-seq track of cells that lack CDX2 (**Figure 43A, *Cdx2*-/- blue track**), suggesting a decrease in the accessibility of chromatin in *Sis* gene locus. These observations were confirmed by the observation of a vast decrease in the gene expression in *Cdx2* inactivated cells which confirmed

the presumptions of change in the chromatin state in the *Sis* gene locus (**Figure 43B**). Thus, these observations suggested that CDX2 regulates the chromatin state and activity of *Sis* gene in villus cells.

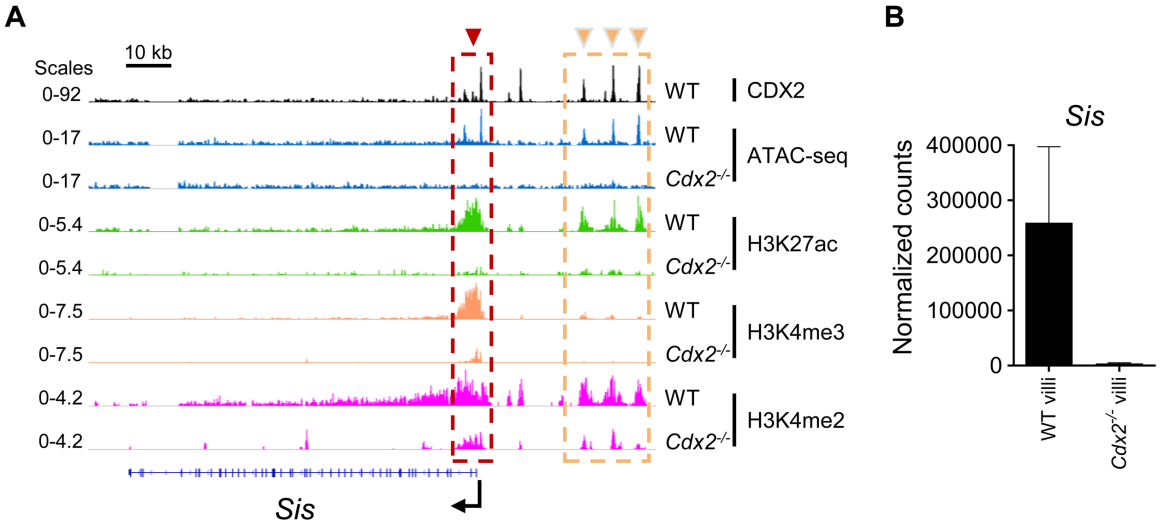


Figure 43: An example for strong *Cdx2*-dependent regulation: Expression of *Sis*. Experiments were performed by Saxena *et al*. [190]. Small intestinal villus cells of wild-type (WT) and *Cdx2*-/- mice were subjected to RNA-seq, ATAC-seq and ChIP-seq of different markers. For detailed information on these techniques, refer to **supplementary Figure S 9**. **(A)** The ATAC- and ChIP-seq tracks for *Sis* gene were generated for this thesis using IGV Browser, from the coverage files under the accession number GSE98724 in GEO database. CDX2 ChIP-seq (black) reveals the regions CDX2 binds. ATAC-seq tracks (blue) depict the accessible regions of the chromatin. Promoter regions are marked by high occupancy of H3K4me3 and H3K27ac (orange and green, red box). Enhancer regions are occupied by H3K4me2 and H3K27ac (magenta and green, yellow box). **(B)** Normalized gene counts of *Sis* in WT and Cdx2-/- villus cells. Bars represent mean ± SD, n = 2 WT, n=3 *Cdx2*-/- villus cells.

Next, the ChIP- and ATAC-seq data were analyzed to identify the chromatin state of the genes *Nlrp6*, *Naip1* and *Aim2* (**Figure 44**). CDX2 occupancy was observed at the promoter region of *Nlrp6* which coincided with the broad peaks of H3K4me3 and H3K27ac, as well as at an enhancer region upstream of the gene which was accompanied by histone marks H3K4me2 and H3K27ac. The accessible chromatin regions indicated by the ATAC peaks correlated with the histone marks (H3K4me1/H3K27ac on active enhancers, H3K4me3/H3K27ac on the promoter region), as well as with binding sites of CDX2, indicating a coordination between occupancy of CDX2 and chromatin accessibility (**Figure 44A**). To determine whether CDX2 is necessary for the active chromatin configuration, the consequence of its absence in tamoxifen-treated *Cdx2*$^{f/f}$; *Villin-Cre*^{ERT2} jejunal villi was examined. Although none of the histone marks diminished like in the *Sis* gene, the chromatin accessibility was slightly reduced in CDX2-binding regions in wild-type cells as indicated by the ATAC peaks upon loss of CDX2 (**Figure 44A**). Thus, the reduced

chromatin accessibility resulted in a decrease in the overall expression level of *Nlrp6* (**Figure 45A**), indicating that DNA accessibility and transcription of *Nlrp6* depends on *Cdx2*.

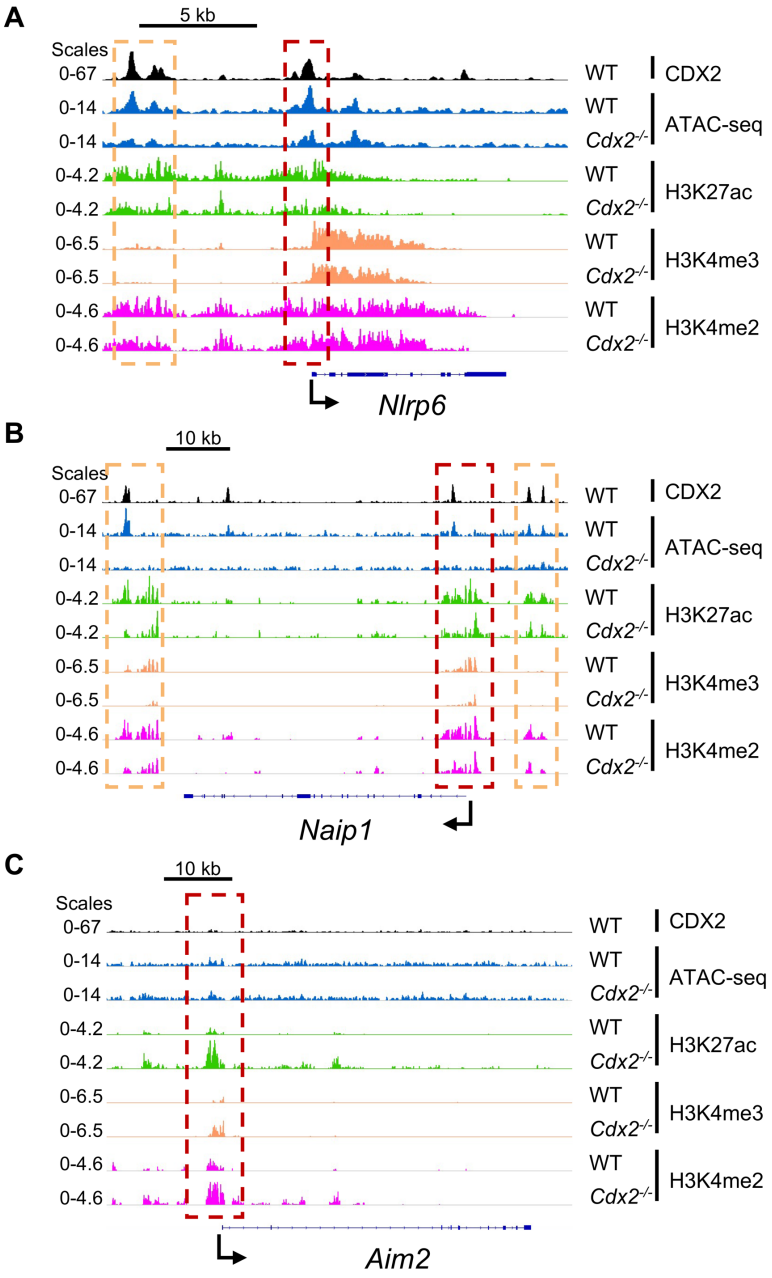


Figure 44: The transcription factor CDX2 affects the chromatin accessibility of *Nlrp6*, *Naip1* and *Aim2*. Experiments were performed by Saxena *et al*. [190]. Small intestinal villus cells of WT and *Cdx2*-/- mice were subjected to ATAC-seq and ChIP-seq of different markers. For detailed information on these techniques, refer to **supplementary Figure S 9**. The ATAC- and ChIP-seq tracks for **(A)** *Nlrp6* **(B)** *Naip1* and **(C)** *Aim2* were generated for this thesis using IGV Browser, from the coverage files under the accession number GSE98724 in GEO database. CDX2 ChIP-seq (black) reveals the regions CDX2 binds. ATAC-seq tracks (blue) depict the chromatin accessibility. Promoter regions are marked by high occupancy of H3K4me3 and H3K27ac (orange and green, red box). Enhancer regions are occupied by H3K4me2 and H3K27ac (magenta and green, yellow box).

ChIP-Seq data also revealed occupation of several enhancers upstream of the promoter region of *Naip1* with CDX2. These sites were flanked by the histone marks H3K4me3 and H3K27ac. Similar CDX2 occupancy and histone marks were observed downstream of the gene (**Figure 44A**). Although no clear CDX2 occupancy was observed in the promoter region, knockout of *Cdx2* clearly decreased the expression of *Naip1* (**Figure 45B**), indicating the expression of *Naip1* (**Figure 45B**), thereby indicating that the DNA accessibility and transcription of *Naip1* depends on *Cdx2*.

Lastly, the ChIP-seq data did not show a significant occupancy CDX2 on *Aim2* and inactivation of *Cdx2* did not lead an increase in the ATAC peaks. However, the occupancy of the promoter region by the histone marks H3K4me2, H3K4me3 and H3K27ac (**Figure 44C**) increased in the knockout mice, which might have been the cause of the increased expression of *Aim2* upon inactivation of *Cdx2* (**Figure 45B**). Hence, loss of *Cdx2* indirectly affected the *Aim2* and resulted into an increase in its expression.

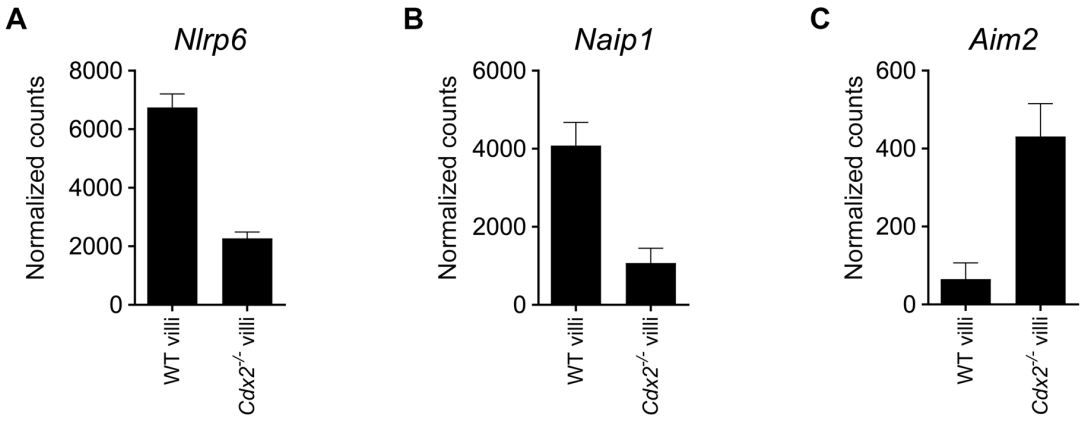


Figure 45: Expression of *Nlrp6, Naip1* and *Aim2*. All experiments were performed by Saxena *et al*. [190]. The data were generated from read counts of RNA-seq under the accession number GSE98724 in GEO database. Normalized gene counts of (A) *Nlrp6* (B) *Naip1* and (C) *Aim2* in WT and Cdx2-/- villus cells. Bars represent mean ± SD, n = 2 WT, n=3 *Cdx2*-/- villus cells.

Taken together, CDX2 as a master regulator of intestinal cell identity, might regulate the expression of various innate immune-related genes which are tissue-specific, such as *Nlrp6* and *Naip1*, by regulating the chromatin accessibility of these genes.

4 Discussion

4.1 Establishment of human and murine GI organoids

Tissue identity is conserved in the organoids

The organoids generated from tissue-specific stem cells of different segments of the GI tract contained genes that displayed the expected tissue specificity, demonstrating that tissue identity is encoded in the adult stem cells and maintained in adult stem cell-derived organoids. This is consistent with previous reports where organoids were generated from different GI segments and had specific expression profiles [81, 84]. The differences between the segments were noticeable during culturing as well. When murine organoid cultures were first established, the small intestine was isolated completely and organoids showing intestinal morphology could be grown in murine small intestinal medium containing EGF, R-spondin1 and Noggin but without WNT, whereas organoids from the stomach and colon could only grow in the presence of WNT in the medium [23, 82]. This was attributed to the presence of a relatively high amount of WNT-producing epithelial cells, specifically Paneth cells [193] present in small intestinal, but not colonic organoids [61]. Within this thesis, murine proximal intestinal organoids (derived from duodenum and jejunum) were able to grow in small intestinal organoid culture medium; however, it was only possible to grow murine ileal organoids with additional WNT in the medium, similar to colon organoids. A reason for this might be an intrinsically reduced number of Paneth cells, which, as mentioned above, provide the WNT ligands to the small intestinal epithelial cells. The presence of Paneth cells has been shown in murine SI organoids grown from whole SI [82] but has yet to be investigated for duodenal, jejunal and ileal organoids separately. Investigating the amount of Paneth cells differentiating in organoid cultures from different SI segments when exposed to different growth conditions could be a promising approach for future studies to elucidate the reason for the different media requirements.

The placement of murine ileal organoids on the PCA plot also suggests that they are different than proximal segments of the small intestine and closer to colon (see **Figure 14**, page 35), therefore there have to be other factors involved in shaping the regional identity. One of those factors could be reduced Paneth cell differentiation in ileal

organoids leading to the WNT addition to the culture medium. However, from just the data on murine organoids presented in this thesis it is impossible to differentiate whether the lack of Paneth cells and therefore the requirement for WNT ligands is due to an intrinsically different regional identity, or whether the placement of ileal organoids on the PCA plot is a result of adding WNT to the medium. On the other hand, human organoids require WNT in the culture medium, independent of the GI segment they were grown from [61, 194], yet also here ileal organoids cluster between gastric and colonic ones. This is a strong indication, that there really is an intrinsic regional identity for each GI segment, and that the ileal tissue identity is actually an intermediate between proximal SI and colonic regional identity.

Organoids provide a suitable model to address questions on innate immunity

Under *in vivo* conditions, epithelial cells can recruit immune cells, which in turn secrete factors that stimulate the epithelial cells, making it difficult to analyze one without the influence of the other. Before organoid technology, studies analyzing questions on immunity used primary isolated epithelial cells, which contain minor fractions of immune cells, including dendritic cells [169]. This is also reflected in the data presented here where the isolated primary epithelium also contained these minor fractions (see **Figure 12**, page 33). It can be speculated that even a low number of activated immune cells could potentially secrete proinflammatory stimuli such as TNF-α which could stimulate many epithelial cells and prevent to define the purely epithelial response, thereby making it difficult to draw definitive conclusions. Organoids on the other hand have been particularly helpful to analyze the epithelial immunity of the GI tract by allowing a concentrated view of purified epithelium without any crosstalk between different cells types, such as mesenchymal and immune cells. They grow from epithelial stem cells under conditions that support epithelial growth and after the first passage contain pure epithelium [82]. This is also reflected in the absence of immune and mesenchymal markers in the organoid cultures used for this study (see **Figure 12**, page 33). The concentrated view on the reaction of pure epithelium to stimulation with immune factors has already been exploited by others who have used organoids to test PRR signaling components [86, 195, 196]. Taken together, the data presented in the literature and this thesis show that organoid technology is likely to be a suitable model to address questions on

innate immunity without the crosstalk between different cell types impacting the analysis as shown in the following.

4.2 Analysis of innate immune signaling in the GI tract

Expression and function of innate immune signaling components is highly organized along the cephalocaudal axis of the GI tract

While it is expected that genes important for regional functions such as digestion and nutrient uptake follow spatial compartmentalization along the cephalocaudal axis [197], such an organization was not presumed for epithelial innate immune signaling. Previous studies had reported regulation of PRR signaling in response to stimulation with MAMPs. For example, TLR4 responsiveness decreases after birth, presumably because of the exposure to LPS during delivery and subsequent colonization of the gut [169]. Also, stimulation of TLR9 with its ligand CpG-DNA leads to a decrease of *Tlr4* expression and inhibits TLR4 signaling [170]. Thus, it was expected that contact with the microorganisms, their molecules and metabolites in the GI tract would lead to silencing of PRR expression towards the gut lumen [95]. However, contradicting results, caused by technical difficulties such as unreliable antibodies targeting TLRs, led to confusion in the field and for a long time it was unclear, whether a particular PRR was expressed or not [157].

In this thesis it was found that this highly complex regional organization of PRR signaling does not always follow the microbial load, instead, it turned out to be segment-specific (Results-3.2, page 37 ff.). Similar to the results presented here, early Northern blots for mRNA of *Tlr2* and *Tlr4* in *ex vivo* isolated epithelium already indicated that expression levels of these two *Tlr* molecules were segment-specific: *Tlr2* was expressed mainly in the colon, while *Tlr4* was mainly expressed in stomach and colon. The authors already termed this "strategic compartmentalization" of these TLRs [158]. Recently, to circumvent the technical problems with directly staining for TLRs, Price and colleagues used staining of genetically introduced HA-tags in reporter mice for *Tlr2*, *4*, *5*, *7* and *9*. By staining for HA, it was then possible to detect the tagged TLRs and analyze the expression of these genes in primary cells from small intestine and colon [159]. TLR2 and 5 were expressed in the small intestine and the proximal colon, TLR4 in colon, TLR7 and 9 were not expressed. Reporter

expression in organoids from these mice closely mimicked the *in vivo* expression, indicating that the expression is independent of contact with the microbiota or with immune cells [159]. The transcriptional data presented here confirmed the previously reported expression patterns for the TLRs, and additionally revealed a vast extent of differential expression of TLRs, NLRs, inflammasome components and other genes related to innate immunity.

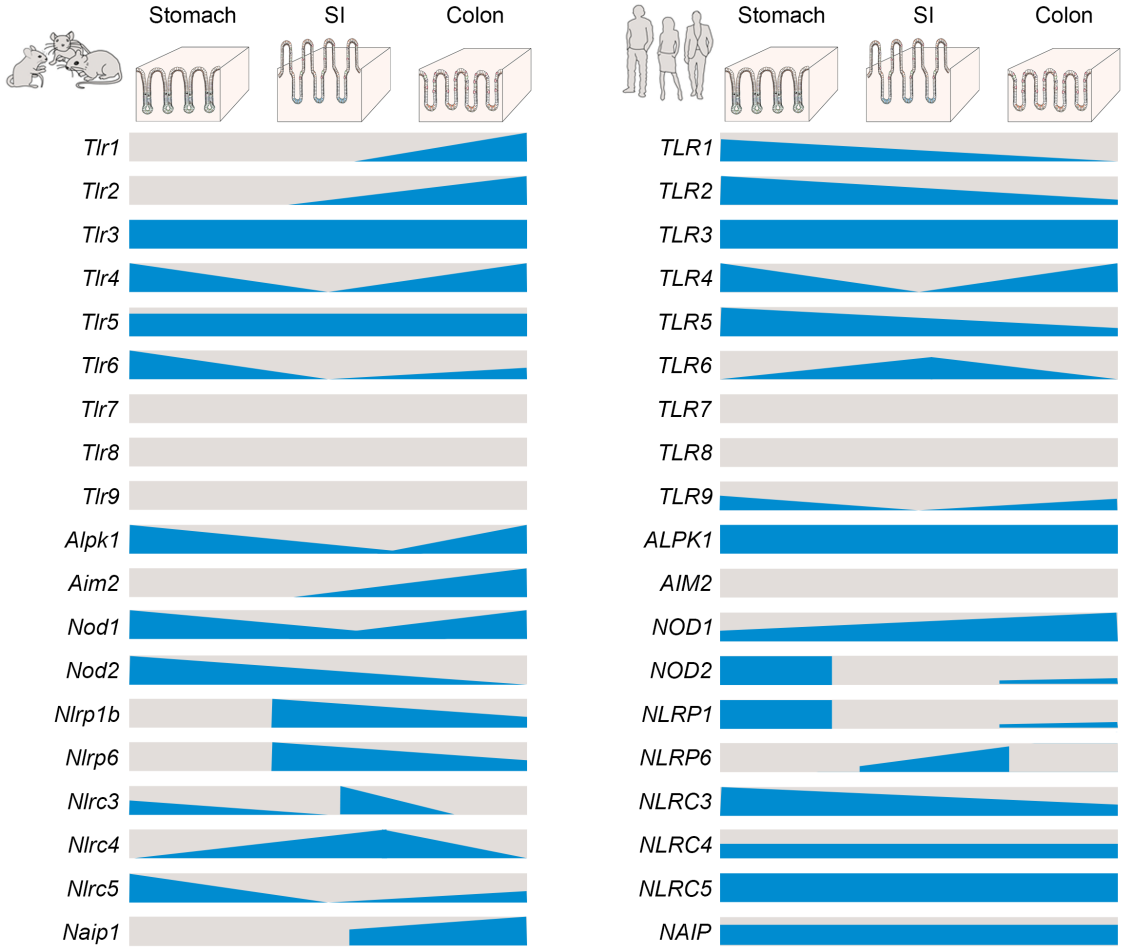


Figure 46: Distribution of various pattern recognition receptors along the murine (left) and human (right) GI tract. The graphic illustrates the relative level of RNA expression (blue), as measured by RNA-seq of the human and murine biobank generated for this thesis. Note the segment-specific expression and the differences between the species.

The patterning of many innate immune signaling components matches the location-specific tissue function

Although the mechanisms shaping the innate immune patterning are not yet fully understood, the need for specialized innate immune recognition might be evolutionarily determined by the complex environment, such as the physical or chemical barriers, of each site. In other words, the GI tract comprises anatomically

defined segments with vastly distinct physical functions (Karam, 1999). The main function of the stomach is the digestion of food and killing of incoming pathogens by gastric acid. Nutrients do not have to reach the epithelial cells, thus, the body heavily invests in a protective mucus barrier, shielding the epithelium not only from its own acid, but also from luminal content [198, 199]. By contrast, the main function of the small intestine is not only the continuation of digestion, but also the uptake of nutrients. Jejunum with the highest digestion and absorption, also contains the highest ratio of Paneth cells secreting antimicrobial peptides [200], which in turn decorate the rather loose mucus, thereby guarding the epithelium and keeping the crypts sterile [47]. In contrast, the ratio of the mucus-producing goblet cells increases in the ileum where the rate of digestion and absorption is also lower (Cheng, 1974b; Karam, 1999). Similarly, the goblet cell ratio in colon can be up to 25% of the epithelial layer, since the reabsorption is less extensive compared to small intestine and the organ can invest into an extensive, thick and bi-layered mucus cover to be able to safely harbor trillions of commensal bacteria [202]. Thus, the three gut segments have different strategies for maintaining a safe distance between the epithelial cell layer and the microbiota. MAMP recognition and activation of immune pathways is another layer of this interaction and therefore it is only reasonable to assume that they are also structured along the GI tract, which is further analyzed in the following.

TLR4 signaling is independent of polarization in the murine stomach

It is intuitively understood that the epithelium needs to minimize any unnecessary inflammation by microbiota, yet still enable the immediate sensing of pathogens and the initiation of a strong inflammatory response to infections [9, 10]. Thus, it has been postulated that this may be accomplished by a certain “blindness” of the epithelium toward the lumen of the gut, confining PRRs to the intracellular space or the basal side of the cells [7–10].

One example of this location-specific function of PRRs is TLR4. The data regarding the localization of TLR4 have been heterogenous, owing to difficulties with TLR-specific antibodies [157]. Price and colleagues presented the staining of the genetically introduced HA-tag at both apical and basolateral surfaces in their reporter mice for *Tlr2*, *4* and *5* [159]. Similarly, in the project presented here, apical and basal stimulation of TLR4 (Results-3.2.3, page 45 ff.) with its ligand LPS in murine gastric

epithelial cells induced an NF-κB response. It could still be the case that the reaction is not mounted by apically or basally located TLR4, but that LPS could be taken up by the cell to stimulate intracellular TLR4 instead [111], but the data presented here shows that murine gastric epithelial cells have the general capability to mount an inflammatory response to apical and basal LPS. However, it cannot be concluded that the luminal LPS also induces constant inflammation *in vivo*. It is probable that *in vivo*, the mucus in the stomach and the colon may pose a diffusion barrier to keep luminal LPS at a distance from epithelial TLR4. Indeed, LPS is present in the colon in a steep gradient with high luminal and very low mucosal concentrations [186]. In addition, gastric mucus and colonic mucus share physical properties and are both firmly attached to the epithelium, while the small intestine has a different, easily removable mucus [203]. Therefore, apical LPS sensing in the stomach and colon, but not the small intestine might indicate a breach of the mucus barrier. Low levels of LPS diffusing through the barrier at physiological conditions may induce the often proposed, low level of "physiological inflammation", important for epithelial homeostasis [97, 145, 146, 157]. Based on the results of this thesis, future studies have to verify these speculations.

In contrast to murine organoids, none of the human GI organoids mounted an inflammatory response to LPS (Results-3.2.4, page 48 ff.). Although expression of *TLR4* was observed and each segment had functional NF-κB pathway, this was not sufficient to activate the TLR4 signaling pathway either apically or basally. One explanation might be that mRNA expression does not always indicate a properly folded and functional protein and post-transcriptional modifications and regulations might play a role in this. Another explanation might be to assume a functional TLR4 protein but lack of MD2 and LBP since they were not expressed in human organoids, while CD14 was expressed in lower amounts compared to murine organoids with functional TLR4 signaling pathway. These components are crucial for functional stimulation and induction of NF-κB pathway. LBP binds and transfers the LPS to CD14 once LPS is released from vesicles of bacteria. CD14 splits LPS into monomeric molecules and presents them to the TLR4-MD2 complex which can lead to the activation of the NF-κB pathway [204]. In another study, colon cancer cell lines expressing low levels of *TLR4* and *MD2* did not respond to LPS stimulation, but only the transgenic expression of these proteins together enabled the recognition and

activation [205]. This underscores the importance of functional testing. Furthermore, LPS varies between bacterial species and also under different growth conditions within the same species [206]. Therefore, stimulation with another type of LPS might result in a different TLR4-mediated response, as long as it is purely isolated to ensure the activation of only TLR4. In addition, there might be a species-specific response to LPS purified from the same bacterial species [207, 208]. While murine TLR4/MD2 complex is activated by LPS from *E. coli*, human cells might not, since the E. coli is commensal in the gut.

Innate immune signaling in the GI epithelium is species-specific

A further surprising finding was the species-specificity of patterning (Results-3.2.2, page 41 ff.). It must be taken into account that the subjects the organoids were generated from were very different: laboratory mice had identical genotype, diet and probably also a highly similar gut microbiome, while the patients differed in these respects. It is likely that environmental factors also influence the expression patterns, and the higher diversity of expression patterns in the diverse human cohort can be taken as an indication for this. Beyond this, however, some PRRs, such as *TLR1*, *2* and *6* showed strikingly species-specific expression along the cephalocaudal axis. In addition, the functionality of the TLR2, 4 and 5 signaling cascades indicated species-specificity. Murine jejunum mounts the *Cxcl2* response to PAM3CSK4 and flagellin, but not to LPS, whereas murine colon expresses *Cxcl2* in response to all of these 3 tested ligands. In contrast, human organoids from all regions expressed the human *Cxcl2* analog *IL-8* in response to flagellin, but not in response to LPS and PAM3CSK. Again, the reasons for this are highly speculative at the moment. It may be environmental, as mentioned above. It is also possible that for each species tissue-specific sensors are beneficial in general, but which ones are present might not be important. For example, a species would benefit from a sensor indicating a breach of the mucus barrier, because this would enable sensing of pathogens that come in close proximity of the epithelium. It may not matter which of several possible sensors indicate such a threat. While TLR4 could assume this function in mouse, another PRR could assume it in human.

4.3 Analysis of the factors regulating innate immunity

The patterning of innate immune signaling components is largely independent of the contact to microbial compounds

Several studies have proposed a contribution of the environment, in particular microbial colonization, to the regulation of PRR expression after birth [95, 169, 170]. However, in previous studies germ-free vs specific pathogen-free mice did not show differences in TLR expression in either small intestine or colon, indicating that neither upregulation of TLR3 nor downregulation of TLR5 in this early period depend on the microbiota [159, 171]. Nevertheless, there is a possibility that the microbial compounds left on the autoclaved food pellets given to germ-free mice have an effect on microbial exposure in these animals [209]. To investigate the regulation of epithelial innate immunity and eliminate the effect of any microbial stimuli, gastric and proximal intestinal organoids generated from murine embryos were used. The embryo and the organoids derived from it are sterile and previous studies demonstrated that embryo-derived organoids mature over time in culture [81, 91, 182]. The data presented here also supported the expected *in vitro* maturation and regional identity which suggests that the embryo-derived organoids provide a beneficial alternative model to germ-free mice.

Comparing embryo-derived to adult-derived organoids, it was observed that some innate immune-related genes are differentially expressed, indicating that there might be an environmental influence and further development during or after birth. However, the expression of the majority of innate immune-related genes, including *Tlr4*, was already patterned in embryo-derived organoids independent of contact with bacteria at birth. Regarding TLR4, embryo-derived murine organoids show the same activation pattern as adult-derived organoids, with functional TLR4 signaling in the stomach, but not in the proximal intestine. These data support developmental definition of a large part of pattern recognition.

This indicates that a large part of the organization of the innate immune signaling pathways is defined independently of contact with the microbiota, and appears to be determined by default developmental processes, which shape the general tissue

identity along the GI tract. However, this does not exclude a further fine-tuning of PRR expression by environmental factors during adulthood.

Expression of several PRRs depend on cellular differentiation within the epithelial layer of GI segments

The analysis of the existing datasets of gene expression in freshly isolated epithelium again confirmed the patterning of the immune-related genes. As in the organoids (Results-3.2.1, page 37 ff.) and confirmed for TLR4 in own experiments (Results-3.2.2, page 41 ff.), also the analysis of the published datasets from a different study [188] showed that different segments of the gut each have a specific set of highly expressed innate immune-related genes. The general trend is either an increase or a decrease along the cephalocaudal axis, so that genes that are highly expressed in the proximal small intestine are only found in very low amounts in the colon, whereby the opposite is true for genes weakly or not expressed in the proximal SI. This is likely connected to the function, as can be illustrated on the example of *Tlr*s. As became apparent in the analysis performed here the expression of different TLRs varies along the cephalocaudal axis, whereby intracellular *Tlr*s (*Tlr3*, *9*, *12* and *13*) are highly expressed in the SI, decreasing along the cephalocaudal axis, whereas extracellular *Tlr*s (*Tlr1*, *2*, *4* and *5*) increase along said axis. While the expression of innate immunity genes does not exactly correlate with the overall microbial load (low in stomach and high in colon, refer to discussion 4.2) it does correlate with the exposure of GI cells to microbiota: A high amount of intracellular TLRs can be found in regions that cannot invest in a heavy mucus barrier, where the cells are generally exposed to microbiota (small intestine), and a high amount of extracellular TLRs in regions where a breach of the mucus barrier is generally a signal of danger (stomach and colon).

That this patterning along the cephalocaudal axis is very finely tuned is also visible in the fact, that separating the organ into smaller sections (such as duodenum, jejunum and ileum) oftentimes reveals even more differences than simply comparing the whole SI to the colon. This was emphasized in the analysis performed on the data of Kozuka et al., who additionally divided the colon into a proximal and a distal part, as other groups had also done in the past [159, 210], while in the first part of this thesis, I only analyzed the proximal colon to represent the whole organ. My analysis

on these data from Kozuka *et al*., did indeed reveal an even finer patterning: For example, while *Nod2* and *Tlr5* were mostly expressed in the proximal colon, *Nlrp10* was mainly expressed in the distal colon. This suggests that future studies could provide even finer descriptions of the innate immune patterning along the cephalocaudal axis.

In a second level of patterning, innate immune gene expression is also organized along the crypt-villus axis (Results-3.3.2, page 58 ff.). Most of the genes related to PRR signaling pathways which are expressed in the small intestine were expressed more in the villus region. Until now a thorough analysis of how genes related to innate immunity are located in crypt-villus axis has not been done. Price *et al*. observed that TLR5 was located only at the crypt bases of the small intestinal epithelium, specifically on the Paneth cells. They also did not detect TLR7 and TLR9 both at RNA and protein level, nor *Tlr6*, *8*, *11*, *12* and *13* at RNA level [159]. In the analysis performed as a part of this thesis, the division of the material into crypts and villi also allowed detection of *Tlr12* and *Tlr13* expression especially at the villi of the small intestine. Similarly, *Gsdmd*, *Nlrp1b*, *Nlrp9b*, *Naip1*, *Nlrc4* and *Rig-I* were detected more at the villous region as well. These data suggest that the PRRs are generally more abundantly expressed in small intestinal villus cells compared to the crypts. However, the opposite is true for some of the *Tlr*s, specifically those located intracellularly, which can be found more highly expressed in the villi. As also mentioned, concerning cephalocaudal patterning above, the intracellular *Tlr*s seem to be mainly found in regions that are generally exposed to the microbiota, where the immune response would otherwise be activated perpetually, whereas extracellular *Tlr*s can be found more highly expressed in regions where the contact to microbiota is a sign of danger rather than a normal occurrence.

Another study in the porcine intestine also found that crypts showed a lower amount of PRR expression [211]. This finding was confirmed for murine cells in this thesis by analysis of existing expression data of cells derived from organoids. For this, the group that performed the original experiments had used directed differentiation of 2D monolayers derived of organoids, which can be used to enrich for cell types [188]. The analysis presented in this thesis confirmed the expression of several innate immune signaling components: Similarly, as in the *ex vivo* isolated epithelium,

several genes such as *Gsdmd*, *Nlrp9b* and *Naip1* were expressed more highly in cell types typically present at the villous region i.e., enterocytes. Other genes such as *Cgas* and *Pycard*, were expressed more strongly in cell types present in the crypt, i.e., undifferentiated progenitor or stem cells. In the literature, this patterning was also observed for the receptor NOD2, which was detected mainly at the crypt region of the small intestine [172] and found to be restricted to stem cells [154], which is another prominent example suggesting that cells have their own repertoire along the crypt-villus axis. Contrary to that, in the analysis performed here no substantial difference could be observed in the expression of *Nod2* between villi and crypts. Furthermore, while *Nod2* was found to be more highly expressed in undifferentiated cells when the data from Kozuka *et al.* was analyzed here, it was still present in differentiated cells, not entirely restricted to stem cells. This controversy should be further investigated in future studies, but from literature and the data and analyses presented here there is a strong indication that the expression of several innate immune genes is likely cell type-specific. The distribution and differentiation of cell types along the crypt-villus axis (Paneth cells, stem cells and progenitors in the crypt versus differentiated goblet cells and enterocytes in the villus region) is likely to cause the observed patterning along the crypt-villus axis. Fully differentiated cells at the villus tip may play a role in epithelial innate immune recognition so far underestimated.

Tissue identity controlled by the intestinal master regulator CDX2 correlates with the expression of innate immune genes, such as Nlrp6, Naip1 and Aim2

Previous results presented in this thesis depicted the presence of various PRRs in either stomach or intestine, and that the expression of many genes related to innate immunity are determined independent of the microbial stimuli, but instead determined developmentally depending on the tissue identity. To approach a mechanistic understanding of the molecular basis of this cephalocaudal organization of innate immune signaling components, I searched for transcription factors, which are important in the cephalocaudal patterning of the GI tract. Prime candidates were CDX1 and CDX2, the well-known master regulators of intestinal development. Because of the importance of CDX2, several datasets of genetic analysis after knockout and overexpression of this transcription factor have been published. By re-

analyzing these published datasets, I investigated the possible impact of CDX2 on innate immune genes (Results-3.3.3, page 66 ff.).

The data from RNA-seq analysis after knockout or overexpression of *Cdx2* in organoids demonstrated that *Cdx2* expression levels regulate the expression of several PRRs. This observation is particularly clear for *Nlrp6*, *Naip1* and *Aim2* which are three cytosolic PRRs. For these genes, *Cdx2* inactivation in small intestinal organoids, which normally express *Cdx2* at high levels, causes their downregulation. Matching this, overexpression of *Cdx2* in gastric organoids, which normally do not express *Cdx2*, causes upregulation of these genes. Accordingly, in the analysis of CDX2 ChIP-seq, I found that CDX2 binds to the promoter and enhancer regions of the genes *Nlrp6* and *Naip1*, and ATAC-seq analysis revealed a decrease in the chromatin accessibility when it is inactivated.

These results need to be further confirmed by future experiments. Nevertheless, the observation that *Cdx2* not only regulates the typical intestine-specific genes relevant to its physiology, such as *Sis*, but also genes related to innate immunity, further solidifies the concepts that innate immune signaling in the GI epithelium is a part of the tissue identity.

5 Conclusion and Outlook

Pattern recognition and innate immune signaling play a crucial role in inflammatory responses and tissue homeostasis. Dysfunctional PRR signaling in the GI tract has been causally implicated in inflammatory diseases. However, the organization of the immune receptors in epithelial cells, as well as the factors promoting their localization and functionality were not clearly understood. In this thesis, a gene expression atlas of human and murine GI epithelial organoids was created. The results showed that the epithelial innate immune barrier follows a specific pattern per GI segment which is encoded in the tissue-resident stem cells and determined independent of bacteria. Innate immune receptors are not completely silenced in the GI tract, but instead, there is regional apical sensing of microbial molecules, supporting the possibility of regional physiological inflammation in homeostasis. Additional to the segment-specificity, the cellular heterogeneity in the epithelial layer and the differentiation state of the cells determine the expression of innate immune-related genes. The transcription factor CDX2 specific to the intestine determines the expression of at least some PRRs, which complements the concept of segment-specific innate immune profiles.

Previous studies have given an incomplete picture of expression of PRRs in the GI epithelium. One important study, published during the preparation of this thesis, used genetic labeling of TLR2, 3, 4 and 5, then showed that these TLRs are expressed in specific patterns along the cephalocaudal and crypt-villus axis of the GI tract [159]. In this present thesis, I have generated a biobank of GI epithelial organoids to cover 6 segments of the GI tract. RNA sequencing of this biobank indicated a vast extent of species- and segment-specific regional patterning of PRR signaling components throughout the GI tract. While it is difficult to summarize omics data to a single outcome, a picture emerged, that the stomach and colon both use extracellular PRRs, such as TLRs, while the small intestinal cells rather express inflammasome components. It is likely, that the three gut segments have different strategies for maintaining a safe distance between the epithelial cell layer and the microbiota. MAMP recognition and activation of immune pathways is another layer of this interaction and therefore it is only reasonable that they are also structured along the

GI tract [212]. However, it remains a fascinating riddle what exactly shapes the structure of the PRR organization.

Because the regional patterning did not follow the microbial load, I hypothesized that pattern recognition may be regulated independently of contact with microbial products. Transcriptional profiling in embryo-derived organoids showed that while expression of some PRRs may depend on environmental cues as expected, an unexpectedly large part of segment-specific expression of PRR signaling components is indeed independent of prior contact with microbial products. Taking TLR4 as one example, functional experiments show that sensing of LPS by TLR4 is also determined already prior to birth and therefore independent of feedback from microbiota. Therefore, the expression of not all, but a surprisingly large part of epithelial innate immunity is developmentally defined and conserved in tissue-resident stem cells. In the last part of this thesis, I identified direct regulation of CDX2 for three innate immune genes (*Nlrp6*, *Naip1* and *Aim2*). Because CDX2 is a master regulator of intestinal development, this represents a molecular mechanism how the epithelial innate immunity can be directly shaped by developmental processes independent of microbial compounds. This mechanism needs to be investigated further in the future.

The majority of knowledge about PRR signaling has been gathered from research on hematopoietic cells. It has been a challenge to discriminate the functions of PRR signaling in epithelial cells from those of the infiltrating immune cells. Major obstacles include difficulties in isolating pure epithelial cells and raising reliable antibodies against PRRs. With the advancement of organoids that derived from intestinal epithelial stem cells, a reductionist experimental model is now available that enables investigations into the innate immune response of the primary epithelial cells. In this thesis, I used adult stem cell-derived organoids as a model to define epithelial innate immunity. Importantly, these organoids consist exclusively of epithelial cells and therefore open up the possibility of studying expression and function of PRRs at the epithelial cell level, without the confounding effects of circulating immune cells. Thus, the study of organoids can shed light on the specific contribution of epithelial cells alone. In the present thesis, I used the stimulation of organoids with single PAMPs, such as LPS and flagellin. In the future, co-culture

experiments will provide information about the responses to or interactions with specific factors, such as inflammatory stimuli, microorganisms or immune cells. Using this reductionist model will help in the future to disentangle the intricate interplay of microbiota, pathogens, immune cells and epithelial cells in the future.

Lastly, it is important to remember that organoids are generated from patient tissue and each organoid reflects the genetic background of the patient. Accordingly, organoids generated from patients affected by GI disorders, such as IBD, maintain some of the characteristics seen in the corresponding tissue specimens they were derived from. This offers a unique opportunity to further unravel the epithelium- and disease-specific contribution to the pathogenesis of GI disorders – not only those involving inflammation-induced changes but also those involving changes in malignant diseases as well. Both of them may turn out to be linked to a specific contribution of epithelium-derived innate immunity. For this, the systematic establishment of "living biobanks" comprised of organoids is an important step. As a vision for the future, such living biobanks could augment already existing biobanks that currently provide only "dead" biological materials. This would assist not only in research but it could also facilitate individualized diagnostics and therapy for patients.

Taken together, in this thesis, organoid technology was used to illuminate differential and segment-specific expression and function of PRRs within the GI epithelium. A large part of the expression of PRRs was independent of contact with microbial compounds. In terms of mechanism, the transcription factor CDX2 was identified as likely to participate in the expression of three innate immune genes that are specific to the intestine. The overall functional consequences for the complex regulatory systems within the entire GI tract are still unclear and will need to be clarified in the future. This will hopefully contribute to a better understanding of innate immune responses in infection and inflammatory diseases such as IBD in the future.

6 Materials

6.1 Biological materials

6.1.1 Mouse strains

To limit the use of mice in accordance with the 3R principles, we used leftover material of mice that were used for experiments in other groups specified in more detail below and no additional mice were used for this study. Wild-type C57BL/6 adult mice were provided by Bernhard Nieswandt, Rudolf Virchow Zentrum, Würzburg, Germany. Embryos of C57BL/6 mice were provided by Manfred Gessler, Department of Developmental Biochemistry, Würzburg, Germany. *Myd88*$^{-/-}$ mice were provided by Mathias Hornef (Institute of Medical Microbiology, Uniklinik RWTH Aachen, Germany) [213], who originally obtained them from Jackson Laboratory (stock #009088), *Tlr2*/*Tlr4*$^{-/-}$ double knockout mice [214] by Arturo Zychlinsky (Max Planck Institute for Infection Biology Berlin, Germany) and *GFP-p65* knock-in mice [215] were provided by Manolis Pasparakis (Institute for Genetics, University of Cologne, Cologne, Germany).

6.1.2 Human samples

Human GI tissues were obtained from 18 patients to derive organoid cultures, and from the same patients plus an additional 19 patients to obtain tissue RNA. Tissue RNA was frequently contaminated with RNA from immune or mesenchymal cells, so we selected the best possible isolates. The 21 men and 16 women, aged between 22 and 87 years, underwent sleeve surgery, partial or total gastrectomy for corpus and pylorus samples, Whipple procedure (pancreaticoduodenectomy) for duodenum samples, gastric by-pass surgery for jejunum samples or ileocolectomy or hemicolectomy for ileum and colon samples. The procedures took place at the University Hospital of Würzburg. This study was approved by the ethical committee of the University of Würzburg's university clinic (approval 37/16).

Table 1: Patient information for the samples used for RNA sequencing.

Organoid type	Patient #	Gender	Age	Surgery
Corpus_org_1	Patient-01	F	32	Sleeve stomach
Corpus_org_2	Patient-10	F	37	Sleeve stomach
Corpus_org_3	Patient-32	F	47	Sleeve stomach
Pylorus_org_1	Patient-01	F	32	Sleeve stomach
Pylorus_org_2	Patient-36	M	55	Gastric carcinoma
Pylorus_org_3	Patient-42	M	58	Gastric carcinoma
Duodenum_org_1	Patient-43	F	75	Whipple
Duodenum_org_2	Patient-54	F	64	Whipple
Duodenum_org_3	Patient-55	M	60	Whipple
Jejunum_org_1	Patient-44	F	34	Gastric by-pass
Jejunum_org_2	Patient-48	M	50	Gastric by-pass
Jejunum_org_3	Patient-50	M	53	Gastric by-pass
Ileum_org_1	Patient-02	M	75	Hemicolectomy
Ileum_org_2	Patient-03	M	83	Colon carcinoma
Ileum_org_3	Patient-51	M	70	Colon carcinoma
Colon_org_1	Patient-34	M	52	Sigmoid colon carcinoma
Colon_org_2	Patient-35	F	76	Caecal carcinoma
Colon_org_3	Patient-37	F	74	Caecal carcinoma

6.2 Equipment and devices

Table 2: Equipment and devices

Equipment/Device	Producer
Electrophoresis gel system	Peqlab – PerfectBlue GelSystem Mini L
Fluorescence microscope	Thermo Fisher Scientific – EVOS FL Imaging System AMF4300
Freezing container	CoolCell™ LX Freezing Container
Hemocytometer	Hartenstein
Light microscope	Leica – DMi1
Microinjector	Eppendorf FemtoJet® 4i
Nanodrop 2000	Thermo Fisher Scientific – ND-2000
Quantitative thermal cycler	CFX96 Touch Real-Time PCR System
Thermal cycler	Bio-Rad Lab, inc. – C1000 Touch Thermal Cycler
UV Image maker	Intas – UV system

6.3 Disposable materials

Table 3: Disposable materials

Disposable material	Producer – Catalogue #
μ-Slide 8-well plate	ibidi GmbH – 80826
4-well multidish	Thermo Fisher Scientific
Cell culture multi-well plates (24 and 48-well)	Sarstedt Ag & Co., Standard, F
Freezing vials	Sarstedt Ag & Co. – CryoPure tube 72.380.002
Quantitative PCR plates	Bio-Rad inc. – Hard-Shell PCR Plates 96-well, thin-wall HSP9655
Transwells (Cell culture insert, 12 mm, polycarbonate, 3 μm)	Millicell – PITP01250

6.4 Kits

Table 4: Kits

Kits	Producer – Catalogue #
Rnase-Free Dnase Set	QIAGEN GmbH – 74106
Rneasy Mini Kit (250)	QIAGEN GmbH – M5510A

6.5 Chemicals, reagents and solutions

Table 5: Chemicals and solvents

Reagent	Supplier – Catalogue #
Dimethylsulfoxide (DMSO)	Sigma-Aldrich
DL-Dithiotreitol (DTT)	Sigma-Aldrich
D-Sorbitol	Carl Roth
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich
Glycerol	Sigma-Aldrich
KCl	Carl Roth
KH_2PO_4	Carl Roth
Na_2HPO_4	Carl Roth
NaCl	Carl Roth
Orange G	Carl Roth
Sucrose	Carl Roth
Tris base	Sigma-Aldrich

Table 6: Buffers and solutions

Buffer	Composition
6x orange loading dye	0.25% Orange G 30% glycerol in ddH2O
Chelating buffer (5x stock)	1.97 g Na_2HPO_4 2.7 g KH2PO4 14 g NaCl 0.3 g KCl 37.5 g Sucrose 25 g D-Sorbitol in 500 ml ddH2O After diluting to 1x, 0.04 g DTT is freshly added in 500 ml 1x buffer.
EDTA (0.5 M stock)	186.1 g EDTA.$2H_2O$ in 800 ml ddH2O pH = 8.0
Tris-acetate-EDTA (TAE) buffer (50x stock)	2 M Tris base 1 M acetic acid 0.05 M EDTA in ddH2O pH = 8.0
Paraformaldehyde (PFA) 4%	40g/l paraformaldehyde in PBS pH = 7.4
PBS (10x stock)	80 g NaCl 2 g KCl 14.4 g Na2HPO4 pH = 7.4 2.4 g KH2PO4 In 1 l ddH2O

Table 7: Reagents for cDNA synthesis and PCR

Reagent	Supplier – Catalogue #
10x DTT 0.1 M	NEB – B1034A
Agarose	Sigma-Aldrich – A0539-500G
Deoxynucleotide (dNTP) Solution Mix	NEB – N0447L
M-MuLV reverse transcriptase	NEB – M0253L
Protoscript II reverse transcriptase	NEB – M0368L
Quick-Load 2-Log DNA ladder (0.1-10 kb)	NEB – N0469S
Random primers	Invitrogen – 48190011
Sso Advanced Universal SYBR Green Supermix	Bio-Rad inc. – 1725271
Taq DNA polymerase	NEB – M0267X

Table 8: Reagents for immunofluorescence

Reagent	Supplier – Catalogue #
Alexa Fluor 488-conjugated anti-rabbit IgG	Cell Signaling – 4412
Anti-p65 rabbit monoclonal antibody	Cell Signaling – 8242
Bovine albumin fraction V (BSA)	Carl Roth – T844.2
Fluorescein Isothiocyanate (FITC)-dextran	Santa Cruz Biotechnology – sc-263323
Hoechst	Thermo Fisher Scientific – H3570
Triton X-100	Sigma-Aldrich – X100
Vectashield Mounting Medium	Vector Laboratories – H-1000-10

Table 9: Reagents for organoids culturing

Reagent	Supplier – Catalogue #
WNT3A conditioned medium	Produced by stable cell line, kind gift of Hans Clevers
R-Spondin1 conditioned medium	Produced by stable cell line, kind gift of Calvin Kuo
Noggin conditioned medium	Produced by stable cell line, kind gift of Hans Clevers
A83-10 – TGF-β-inhibitor	TOCRIS Bioscience – 2939
Advanced Dulbecco's Modified Eagle Medium F12 (Adv-DMEM)	Invitrogen – 12634-028
B-27 supplement (50x) without vitamin-a	Gibco – 12587-010
Cell culture freezing medium	Invitrogen – 12648-010
Cell recovery solution	Corning – 354253
Collagen I, bovine	Gibco – A10644-01
Epidermal growth factor (EGF) – Human	Peprotech – AF-100-15
Fibroblast growth factor-10 (FGF10)	Peprotech – 100-26
Gastrin-I	Peprotech – 1003377
GlutaMAX-I	Invitrogen – 35050-079
HEPES	Invitrogen – 15630-056
Matrigel	Corning – 356231
N-acetyl-L-cysteine (N-Ac)	Sigma-Aldrich – A9165-5G
Nicotinamide	Sigma-Aldrich – 72340
Primocin	Invivogen – Ant-pm-1
SB-202190 – p38-inhibitor	Sigma-Aldrich – 1264
TrypLE Express	Gibco – 12605-028
Y-27632 – RhoK-inhibitor	Abmole – M1817

Table 10: Ligands for NF-κB activation

Name	Supplier – Catalogue #
Flagellin (FLA-ST)	Invivogen – tlrl-epstfla
LPS-EK (*E.coli* K12 strain)	Invivogen – tlrl-eklps
Pam3CSK4	Invivogen – tlrl-pms
Recombinant human TNF-α	BD Pharmingen

6.6 Primers

Table 11: Primers for murine genes

Gene	Forward	Reverse	Product (bp)	Reference
Gapdh	5′-GTGCCAGCCTCGTCC-3′	5′-ACCCCATTTGATGTTAGTGG-3′	283	This study
Cd45	5′-GACCCTATTTCTTAGGGGCA-3′	5′-CTCTGTTGTGCTCAGTTCATC-3′	150	This study
Cdx2	5′-CCTAGGAAGCCAAGTGAAAA-3′	5′-TGCGGTTCTGAAACCAAAT-3′	185	This study
Cxcl2	5′-AAGTTTGCCTTGACCCTGAA-3′	5′-AGGCACATCAGGTACGATCC-3′	180	This study
Gli1	5′-AAGGGGACATGTCTAGCCCC-3′	5′-ACAGCCTTCAAACGTGCACT-3′	338	[63]
Muc2	5′-GTGTGTTGCTCAATGAGATG-3′	5′-TCTAGGCCATTGAAGTTTCC-3′	211	This study
Muc6	5′-AACCTGCAATCCTCCCCAGAA-3′	5′-GCTGGATGCTAAAGGTGGCG-3′	157	This study
Tlr4	5′-ATCCCTGCATAGAGGTAGTT-3′	5′-CAAGTTTGAGAGGTGGTGTA-3′	242	This study

Table 12: Primers for human genes

Gene	Forward	Reverse	Product (bp)	Reference
CD45	5′-CTTACCTACTCACACCACTG-3′	5′-GGGACATCTGAGATAGCATT-3′	300	This study
CDX1	5′-CGCCCTACGAGTGGAT-3′	5′-ATTGTGATGTAACGGCTGTA-3′	151	This study
FN1	5′-CAGTGGGAGACCTCGAGAAG-3′	5′-TCCCTCGGAACATCAGAAAC-3′	168	[216]
GAPDH	5′-GTTTCTATAAATTGAGCCCGC-3′	5′-TGTAAACCATGTAGTTGAGGT-3′	246	This study
IL-8	5′-ACACTGCGCCAACACAGAAAT-3′	5′-ATTGCATCTGGCAACCCTACA-3′	241	[86]
MUC2	5′-GCTGTACGTTGGAGTTCTAT-3′	5′-TGGTAGCTGTAGTAGAGTCC-3′	403	This study
MUC6	5′-AACATCGAAGGCTGCTACAA-3′	5′-GGGTTGGTAGTGTCATTGTG-3′	484	This study
TLR4	5′-TGGATACGTTTCCTTATAAG-3′	5′-GAAATGGAGGCACCCCTTC-3′	507	[217]

6.7 Open access data from GEO repository

Table 13: Accession numbers of the data used in this thesis

Accession number	Reference
GSE62784	Simmini *et al.* [73]
GSE98724	Saxena *et al.* [190]
GSE104803	Kozuka *et al.* [188]
GSE115541	Banerjee *et al.* [185]
GSE127938	This study

6.8 Software

Table 14: Software

Name	Developer
Bio-Rad CFX Manager	Bio-Rad inc.
CorelDRAW Graphics Suite	Corel Corporation
Fiji	ImageJ
GraphPad Prism 6	Prism
IGV	Broad Institute
Intas Gel Docu software	Intas
Microsoft Office	Microsoft
Nanodrop 2000/2000C	Thermo Fisher Sci.
Photoshop CS5	Adobe
R	CRAN
Rstudio	Rstudio, PBC

6.9 Websites

Table 15: Websites

Name	Website address
BioRender	https://biorender.com
g:Profiler	https://biit.cs.ut.ee/gprofiler/gost
Primer-Blast	https://www.ncbi.nlm.nih.gov/tools/primer-blast/
Venny 2.1.0	https://bioinfogp.cnb.csic.es/tools/venny/

7 Methods

7.1 Organoid culturing

7.1.1 Gland and crypt isolation

In order to generate organoid lines gastric glands, small intestine and colon crypts were isolated. Chelating buffer was prepared as 5x to be stored at 4°C and diluted with autoclaved water prior to the procedure. DTT was added freshly to the diluted buffer to help with dissolving the mucus. Only for the murine small intestine DTT was not added to the chelating buffer, since it accelerated the disintegration of the tissue.

Murine abdominal region was dissected and whole GI tract was extracted. From each tissue, crypts and glands were extracted using EDTA in cold chelating buffer. Optimized EDTA concentrations and incubation times according to the organism and segment of the GI tract are provided in **Table 16**.

Table 16: Isolation conditions for crypts and glands.

Tissue type		EDTA concentration	Incubation time	Temperature
Mouse	Corpus	5 mM	5 min	Room temp. (on bench)
	Pylorus	5 mM	5 min	Room temp. (on bench)
	Duodenum	1-2 mM	5 min	On ice
	Jejunum	1-2 mM	5 min	On ice
	Ileum	1-2 mM	5 min	On ice
	Colon	2 mM	5 min	On ice
Human	Corpus	10 mM	5 min	37°C (in hands)
	Pylorus	10 mM	5 min	37°C (in hands)
	Duodenum	2 mM	5 min	Room temp. (on bench)
	Jejunum	2 mM	5 min	Room temp. (on bench)
	Ileum	2 mM	5 min	Room temp. (on bench)
	Colon	2 mM	5 min	Room temp. (on bench)

Murine stomach was placed in a 10 cm petri dish and cut along the curvature with sterile scissors. After it was washed in chelating buffer, corpus (darker) and pylorus (lighter) regions were defined based on coloration. With the help of forceps, mucus and muscle layers were removed as possible. On another dry petri dish, the tissue was cut into small pieces of approximately 5 mm^2 size. The pieces were collected in a 50 ml tube with 10 ml chelating buffer. Sterile 10 ml pipettes were pre-wet in chelating buffers for the tissue pieces not to attach to the walls and the pieces were washed by

pipetting up and down. After tissue pieces settled, liquid was removed, and the washing step was repeated until the buffer was clear. EDTA was added to the tube (for murine stomach 5 mM EDTA for 5 minutes at room temperature). Tissue pieces then were transferred to a clean petri dish and extra liquid was removed. A sterile microscope slide was placed on the tissue pieces and pressure was applied on the slide. Cloudy appearance under the slide indicated the released glands from the tissue. With 10 ml Advanced DMEM/F12 (AD) everything on the petri dish was collected in a 15 ml tube. The supernatant containing the glands was transferred to another 15 ml tube to remove it from the tissue pieces. The tube was centrifuged for 5 minutes at 300 x g at 4°C. Supernatant was discarded and the tube with pellet of isolated glands was placed on ice. Depending on the size of the pellet, it was diluted with AD and centrifuged again to obtain a pellet containing around 500 glands per well. With 50 µl/well of Matrigel the pellet was resuspended and placed into wells of pre-warmed 24-well plate (overnight, 37°C). After allowing the Matrigel to solidify in the incubator in 5% CO_2 and 37°C for 15 minutes, each well was supplemented with 500 µl growth medium containing the necessary growth factors (**Table 17**).

Murine small intestine was cut into three equal parts and duodenum (proximal 2 cm of the proximal part), jejunum (proximal 2-4 cm of the middle part) and ileum (distal 2-4 cm of the distal part) were obtained. Of the murine colon, 2 cm of proximal colon was used. The pieces were cut open longitudinally and washed in chelating buffer (for small intestine, chelating buffer without DTT for the whole procedure) briefly. The villi of the small intestine parts were scraped using a microscope slide cover slip and the tissue was cut into smaller pieces. The procedure continued as for stomach, differing at the EDTA incubation step (**Table 16**) and growth media of the corresponding tissue (**Table 17**).

Dissections of the GI tract of embryonic mice were performed in a sterile environment. The GI tracts were separated into stomach and the proximal small intestine, which directly starts after the stomach, whereas the distal part of the tract was discarded. All chosen parts were dissected, cut into smaller pieces and put to separate falcons containing PBS. After incubation in 2 mM EDTA for 5 minutes in room temperature, the tissue pieces were collected in a petri dish and pressed with a microscope slide. With 10 ml AD, all tissue pieces were collected in a new 15 ml tube.

In contrast to adult tissue, where only the supernatant was collected, in case of embryonic mice the supernatant and additionally several pieces of tissue collected at the bottom of the falcon were transferred to another 15 ml falcon tube. The tube was centrifuged for 5 minutes at 300 x g at 4°C, the supernatant was discarded and the pellet still containing small tissue pieces was placed in 24-well plate in Matrigel as explained above.

Human tissues were obtained from the University Clinic of Würzburg. After a brief wash in chelating buffer, the tissue was placed on the petri dish and the main muscle tissue was separated from the epithelial layer. With forceps, remaining muscle and endothelial tissues were cleaned as possible. From this step on, the procedures continued as explained for murine segments, differing at the EDTA incubation step (**Table 16**) and growth medium was added according to the segment (**Table 17**).

7.1.2 Maintenance of organoids

Organoids were grown in basement matrix, Matrigel, supplemented with media containing growth factors appropriate for the species and GI segment as specified in **Table 17**. For basal medium (AD++), AD supplemented with 10 mmol/L HEPES and GlutaMAX 1x was used. RhoK-inhibitor was added only after the initial seeding and passaging of the organoids. The medium was changed every 2-3 days. Murine gastric organoids were passaged at a ratio of 1:5 every week and murine intestinal organoids at the ratio of 1:3 every 5 days. All human organoids were passaged at the ratio of 1:5 to 1:8 depending on the density of the organoids every 10-14 days. For the mRNA sequencing, all adult organoids were expanded for a maximum of 5 passages.

Passaging was performed by mechanical disruption (shredding) through a Pasteur pipette. After removing the medium from the well, 800 µl cold AD++ was added on the Matrigel with organoids. To break the Matrigel, AD++ was pipetted vigorously several times until the Matrigel was dissolved and transferred to a 15 ml tube. Following that, Pasteur pipettes were narrowed using flame under the hood and they were used to shred the organoids themselves physically by repeated pipetting. The tube was then filled with AD++ until 8 ml and inverted for a couple of times and centrifuged for 5 minutes at 300 x g at 4°C. After the supernatant was discarded, the pellet was resuspended with Matrigel (50 µl/well for a 24-well plate, 25 µl/well for a 48-well plate) and drops of Matrigel with pieces of organoids were placed in the wells

of pre-warmed plates (overnight, 37°C). The plates were then transferred to the incubator for 15 minutes and medium with RhoK-inhibitor was added in each well (500 µl/well for 24-well plate, 250 µl/well for 48-well plate).

For freezing the organoids, a similar procedure to passaging was followed. When a well of organoids (in 24-well plate, for one vial to freeze) was ready to freeze, they were collected and shredded to be centrifuged. The pellet was resuspended with 500 µl/vial cell freezing medium and transferred into a cryo-tube. The tubes were placed in a -80°C freezer within a freezing container. After at least 48 hours, vials were transferred to liquid nitrogen for long term storage.

Table 17: Organoid media composition for human and murine organoids. Organoid media were specifically tailored to the species and GI segment. CM: conditioned medium; inh.: inhibitor; N-Ac: N-acetylcysteine; EGF: epithelial growth factor; FGF10: fibroblast growth factor-10; TGF-β: transforming growth factor-β; RhoK: Rho-associated coiled-coil forming protein serine/threonine kinase.

Reagent	Mouse			Human	
	Adult sto./ Embr. sto.	Adult duod. jej./ Embr. int.	Ile. / Col.	Stomach	Sm. int. / Col.
AD++	30%	80%	30%	30%	30%
Wnt CM	50%	-	50%	50%	50%
R-Spondin CM	10%	10%	10%	10%	10%
Noggin CM	10%	10%	10%	10%	10%
Primocin	100 ng/ml	100 ng/ml	100 ng/ml	100 ng/ml	100 ng/ml
B27	1x	1x	1x	1x	1x
N-Ac	1.25 nM	1.25 nM	1.25 nM	1.25 nM	1.25 nM
EGF	50 ng/ml	50 ng/ml	50 ng/ml	50 ng/ml	50 ng/ml
FGF10	100 ng/ml	-	-	100 ng/ml	-
Gastrin-I	10 nM	-	-	1 nM	10 nM
TGF-β-inh.	-	-	-	2 µM	0.5 µM
p38-inh.	-	-	-	-	10 µM
Nicotinamide	-	-	-	-	10 mM
RhoK-inh.	10 µM	10 µM	10 µM	10 µM	10 µM

To thaw the organoids, vials were collected from the liquid nitrogen and quickly added to 9 ml AD++ in 15 ml tubes. After inverting the tubes couple of times, they were centrifuged for 5 minutes at 300 x g at 4°C. Supernatant was discarded and the pellet was resuspended with 50 µl Matrigel per vial to be placed on a pre-warmed 24-well plate. On the solidified Matrigel drops after 15 minutes of incubation at 37°C, 500 µl organoid type-specific media with RhoK-inhibitor were added.

7.1.3 Seeding the cells from 3D organoids to generate 2D monolayers

Growth of murine gastric cells as 2D required Matrigel coating. Therefore, prior to seeding cells for 2D cultures, glass slides, wells of 24-well plates and transwells were coated with Matrigel by incubating them in 1:10 diluted Matrigel in AD++ for 30 minutes at 37°C. Murine gastric organoids were mechanically disrupted as described in 7.1.2 to be seeded. After centrifugation, the organoid fragments were resuspended in previously prepared growth medium supplemented with RhoK-inhibitor and seeded in a ratio of 1:5 as pieces or as single cell suspensions. The remaining Matrigel-AD++ solution was removed before adding the cells.

Human gastric cells grew on the plastic surface of 48-well plates, while collagen coating was required for culturing 2D monolayers on transwells. Per well, solution of 1.6 µl collagen I in 150 µl of 2 M acetic acid was added and incubated for 1 hour at 37°C prior to seeding. After the human gastric organoids were shredded and centrifuged as described in 7.1.2, the supernatant was removed and the pellet was resuspended in 1 ml TrypLE Express. First after 5 minutes, then 10 minutes in 37°C water bath, cells were again pipetted with the narrowed Pasteur pipette to obtain single cells. Using the Neubauer counting chamber, cells were counted, and the cell suspension was diluted according to the required cell amount. On 48-well plates $1x10^5$, on transwells $2x10^5$ cells were seeded for the 2D culture to reach the confluency at day 7. The remaining solution was discarded before seeding the cells. In transwells, TEER to ion migration reached 390-540 $\Omega*cm^2$, which is in the range of published measurements [88].

7.2 Gene expression analyses

7.2.1 Total RNA extraction

RNA was isolated from organoids, isolated gastric glands or intestinal crypts according to the manufacturer's recommendations (RNeasy Mini Kit). For RNA isolation from organoids, Matrigel was first removed from the cultures by gentle disruption with cold AD++ and centrifuging 5 minutes at 300 x g at 4°C.

7.2.2 cDNA synthesis

Complementary DNA (cDNA) was generated using established reverse transcriptase protocols. 500 ng RNA were mixed with nuclease-free water up to 10 µl. After

addition of 2 µl of random primer mix (60 µM), the solution was incubated in a thermal cycler for 5 minutes at 65°C and cooled down at 4°C for 5 minutes. Subsequently, reaction mix was prepared for each tube as explained below and incubated in the thermal cycler as indicated in **Table 18** and **Table 19**, for two different reverse transcriptases used in this project.

Table 18: Reaction mix and duration protocol used for cDNA synthesis by ProtoScript II reverse transcriptase.

Component	1x
ProtoScript II reaction buffer (5x)	4 µl
10x DTT (0.1 M)	2 µl
dNTP mix (10 mM)	1 µl
ProtoScript II RT (200 U/µl)	0.5 µl
Nuclease-free water	0.5 µl

Temperature	Time
25°C	5 min
42°C	60 min
65°C	20 min
12°C	∞

Table 19: Reaction mix and duration protocol used for cDNA synthesis by M-MuLV reverse transcriptase.

Component	1x
M-MuLV reaction buffer (10x)	2 µl
dNTP mix (10 mM)	1 µl
M-MuLV RT (200 U/µl)	0.25 µl
Nuclease-free water	4.75 µl

Temperature	Time
25°C	5 min
42°C	60 min
65°C	20 min
12°C	∞

7.2.3 Conventional PCR and RT-qPCR

Conventional and quantitative reverse transcription PCR (RT-qPCR) were performed using the same primers (diluted to 10 µM), which are listed in **Table 11** and **Table 12** (page 93). PCR was performed with the reaction mix and cycling protocol as indicated in **Table 20** below.

Table 20: Reaction mix and cycling protocol used for conventional PCR.

Component	1x
cDNA	1 µl
10x Standard Taq reaction buffer	2.5 µl
dNTP mix (10 mM)	0.5 µl
Taq DNA polymerase (5000 U/ml)	0.2 µl
Forward primer	0.5 µl
Reverse primer	0.5 µl
Nuclease-free water	19.8 µl

Temperature	Time	
94°C	5 min	
94°C	30 sec	34x
55°C	30 sec	
72°C	30 sec	
72°C	7 min	
12°C	∞	

RT-qPCR was performed using SYBR Green reaction mix, as indicated in **Table 21**. Results were calculated using the ΔΔCt method. Relative quantification was achieved by normalizing results to the values obtained for *GAPDH*.

Table 21: Reaction mix and cycling protocol used for RT-qPCR.

Component	1x
SYBR Green reaction mix	5 µl
Forward primer	0.5 µl
Reverse primer	0.5 µl
Nuclease-free water	3 µl

Temperature	Time	
95°C	3 min	
95°C	10 sec	39x
55°C	30 sec	
95°C	10 sec	
95°C	5 sec	
95°C	0.5 °C	

7.2.4 Agarose gel electrophoresis

10 µl PCR product was mixed with 2 µl 6x loading dye and was loaded on a 2% TAE agarose gel. The gel electrophoresis was performed at 120 V for 30 minutes.

7.3 Stimulation assays

7.3.1 Stimulation of 3D and 2D cultures by ligand addition to media

To investigate activation of the NF-κB pathway, organoids were incubated for the indicated times with medium containing recombinant human TNF-α, LPS from *Escherichia coli* (*E. coli*), flagellin or Pam3CSK4 in the final concentration as stated in the respective section of the results. 3D organoids or 2D monolayers were stimulated shortly before the day on which they were due to be passaged to allow their maturation (murine organoids on day 5, human organoids on day 10-12). For microinjection, 3D organoids were seeded in 50 µL of Matrigel in 4-well multi-dishes, for isolation of RNA, or 8-well µ-slides, for microscopy.

7.3.2 Stimulation of 3D organoids by microinjection

Organoids were microinjected with LPS (100 ng/ml) or Fluorescein Isothiocyanate FITC-dextran (0.5 ng/ml) on day 5 using a micromanipulator and microinjector together with a stereomicroscope within a sterile safety cabinet. FITC-dextran was imaged after injection by live cell microscopy using a Leica DMI6000B with a 10 X objective. For LPS stimulation, 50 organoids were injected within each well. Injections were performed in a paired manner inside and outside the organoids on the same

plate. 2 h after injection of LPS, organoids were removed from Matrigel and lysed for isolation of RNA. For transwell experiments, approximately 200,000 cells were seeded into collagen-coated transwell inserts placed into a 24-well plate.

7.3.3 Quantification of nuclear translocation of NF-κB

Dissociated murine gastric organoids were seeded onto Matrigel-coated glass coverslips to form 2D monolayers for immunofluorescence staining. Cells were grown for 5 days to reach approximately 80% confluency and stimulated with LPS as indicated in individual figures. Cells were fixed with 2% PFA for 20 min at room temperature. Cells were washed 3 times with PBS and permeabilized in 1 X PBS supplemented with 0.3% Triton X-100 and 1% fresh BSA for 1 h. Cells were stained with anti-p65 rabbit monoclonal antibody (8242, Cell Signaling), 1:400 in 1 X PBS supplemented with 0.3% Triton X-100 and 1% BSA overnight at 4 °C followed by Alexa Fluor 488-conjugated anti-rabbit IgG (Cell Signaling), 1:250 in 1 X PBS supplemented with 0. % Triton X-100 and 1% BSA for 3 h at room temperature. DNA was stained with Hoechst 33342. After washing 3 times with PBS, cells were mounted on glass slides in mounting medium. Stained cells were visualized using a fluorescence microscope (EVOS). To quantify the percentage of cells with p65 localized in the nucleus, 9 frames per well were imaged and p65 nuclear translocation was manually quantified for all 9 images, resulting in a total quantification of several hundred (usually approximately 1,500) cells.

7.4 RNA sequencing

RNA was isolated from organoids and was delivered to Core Unit Systems Medicine (CU SysMed), University of Würzburg, Germany for library preparation and sequencing. Libraries for RNA sequencing were prepared using the Illumina TruSeq Stranded mRNA Library Prep Kit (Cat#20020594) with 500 ng of total RNA for the construction of sequencing libraries. RNA sequencing was performed using single-end reads on the Illumina NextSeq 500 platform. Quality control was also performed by the core unit using FastQC (Babraham Bioinformatics) and the reads were obtained for further analysis.

7.4.1 Bioinformatic analysis

Reads were aligned and mapped to the mouse reference genome GRCm38 or the human reference genome GRCh38 with segemehl [218] as a part of RNA-seq pipeline READemption (V.0.4.3) [219] which includes aligning the reads to the reference genomes. Raw read alignments were counted using htseq-count [220]. Bioinformatic analysis was performed in the R software package V.3.4.4 using Bioconductor V.3.6 packages. Low quality reads and adapter sequences were removed using cutadapt (V.1.16) [221]. Using RUVseq batch correction was performed by calling the *ruvs* function [222]. Raw counts were normalized by DESeq2's [223] median of ratios method and these values are depicted in the bar graphs. Differential gene expression analysis was performed on normalized counts using DESeq2 with a cut-off of adjusted $p<0.05$. The Benjamini-Hochberg method was used to adjust the p-values. For the heatmaps, the normalized counts were scaled for Z-score with row means of 0, and heatmaps were created using ComplexHeatmap [224]. Gene ontology enrichment analysis was performed using the web application of g:Profiler [178]. $p<0.05$ was considered as statistical significance and GO-terms were filtered to less than 3,000 genes for more definitive results.

7.5 Statistical analysis

Data represent mean ± standard deviation (SD). $p=0.05$ was taken as the maximum value for significance. The applied statistical tests and the level of significance are indicated in the figure legends.

8 References

1. Sekirov I, Russell SL, Caetano M Antunes L, Finlay BB (2010) Gut microbiota in health and disease. Physiol Rev 90:859–904. https://doi.org/10.1152/physrev.00045.2009

2. Simon GL, Gorbach SL (1986) The human intestinal microflora. Dig Dis Sci 31:147–162. https://doi.org/10.1007/BF01295996

3. Deplancke B, Gaskins HR (2001) Microbial modulation of innate defense: Goblet cells and the intestinal mucus layer. Am J Clin Nutr 73:1131S--1141S. https://doi.org/10.1093/ajcn/73.6.1131s

4. Payen D (2020) the Gut As a Hidden Source of Sepsis. Minerva Anestesiol 86:662–669. https://doi.org/10.23736/S0375-9393.20.14302-5

5. Burgueño JF, Abreu MT (2020) Epithelial Toll-like receptors and their role in gut homeostasis and disease. Nat Rev Gastroenterol Hepatol 17:263–278. https://doi.org/10.1038/s41575-019-0261-4

6. Farhadi A, Banan A, Fields J, Keshavarzian A (2003) Intestinal barrier: An interface between health and disease. J Gastroenterol Hepatol 18:479–497. https://doi.org/10.1046/j.1440-1746.2003.03032.x

7. Peterson LW, Artis D (2014) Intestinal epithelial cells: Regulators of barrier function and immune homeostasis. Nat Rev Immunol 14:141–153. https://doi.org/10.1038/nri3608

8. Pott J, Hornef M (2012) Innate immune signalling at the intestinal epithelium in homeostasis and disease. EMBO Rep 13:684–698. https://doi.org/10.1038/embor.2012.96

9. Abreu MT, Fukata M, Arditi M (2005) TLR Signaling in the Gut in Health and Disease. J Immunol 174:4453–4460. https://doi.org/10.4049/jimmunol.174.8.4453

10. Pédron T, Sansonetti P (2008) Commensals, Bacterial Pathogens and Intestinal Inflammation: An Intriguing Ménage à Trois. Cell Host Microbe 3:344–347. https://doi.org/10.1016/j.chom.2008.05.010

11. Karam SM (1999) Lineage commitment and maturation of epithelial cells in the gut. Front Biosci 4:d286. https://doi.org/10.2741/karam

12. Ross, Michael H. & Pawlina W (2011) Digestive system II: Esophagus and gastrointestinal tract. In: Histology. A text and atlas, 7th ed. Wolters Kluwer Health, Philadelphia, pp 568–627

13. Loe AKH, Rao-Bhatia A, Kim JE, Kim TH (2020) Mesenchymal Niches for Digestive Organ Development, Homeostasis, and Disease. Trends Cell Biol

xx:1–14. https://doi.org/10.1016/j.tcb.2020.11.010

14. Treuting PM, Arends MJ, Dintzis SM (2018) Upper Gastrointestinal Tract. In: Treuting PM, Dintzis SM (eds) Comparative Anatomy and Histology. Elsevier, San Diego, pp 191–211

15. Hounnou G, Destrieux C, Desmé J, et al (2002) Anatomical study of the length of the human intestine. Surg Radiol Anat 24:290–294. https://doi.org/10.1007/s00276-002-0057-y

16. Karam SM, Leblond CP (1993) Dynamics of epithelial cells in the corpus of the mouse stomach. III. Inward migration of neck cells followed by progressive transformation into zymogenic cells. Anat Rec 236:297–313. https://doi.org/10.1002/ar.1092360204

17. Choi E, Roland JT, Barlow BJ, et al (2014) Cell lineage distribution atlas of the human stomach reveals heterogeneous gland populations in the gastric antrum. Gut 63:1711–1720. https://doi.org/10.1136/gutjnl-2013-305964

18. Karam SM, Leblond CP (1993) Dynamics of epithelial cells in the corpus of the mouse stomach. V. Behavior of entero-endocrine and caveolated cells: General conclusions on cell kinetics in the oxyntic epithelium. Anat Rec 236:333–340. https://doi.org/10.1002/ar.1092360206

19. Longman RJ, Douthwaite J, Sylvester PA, et al (2000) Coordinated localisation of mucins and trefoil peptides in the ulcer associated cell lineage and the gastrointestinal mucosa. Gut 47:792–800. https://doi.org/10.1136/gut.47.6.792

20. Hanby AM, Poulsom R, Singh S, et al (1993) Spasmolytic polypeptide is a major antral peptide: Distribution of the trefoil peptides human spasmolytic polypeptide and pS2 in the stomach. Gastroenterology 105:1110–1116. https://doi.org/10.1016/0016-5085(93)90956-D

21. Ootani A, Toda S, Fujimoto K, Sugihara H (2003) Foveolar differentiation of mouse gastric mucosa in vitro. Am J Pathol 162:1905–1912. https://doi.org/10.1016/S0002-9440(10)64324-6

22. Karam SM (1993) Dynamics of epithelial cells in the corpus of the mouse stomach. IV. Bidirectional migration of parietal cells ending in their gradual degeneration and loss. Anat Rec 236:314–332. https://doi.org/10.1002/ar.1092360205

23. Barker N, Huch M, Kujala P, et al (2010) Lgr5+ve Stem Cells Drive Self-Renewal in the Stomach and Build Long-Lived Gastric Units In Vitro. Cell Stem Cell 6:25–36. https://doi.org/10.1016/j.stem.2009.11.013

24. Leushacke M, Ng A, Galle J, et al (2013) Lgr5+ Gastric Stem Cells Divide Symmetrically to Effect Epithelial Homeostasis in the Pylorus. Cell Rep 5:349–

356. https://doi.org/10.1016/j.celrep.2013.09.025

25. Stange DE, Koo B-K, Huch M, et al (2013) Differentiated Troy+ Chief Cells Act as Reserve Stem Cells to Generate All Lineages of the Stomach Epithelium. Cell 155:357–368. https://doi.org/10.1016/j.cell.2013.09.008

26. Sáenz JB, Mills JC (2018) Acid and the basis for cellular plasticity and reprogramming in gastric repair and cancer. Nat Rev Gastroenterol Hepatol 15:257–273. https://doi.org/10.1038/nrgastro.2018.5

27. Han S, Fink J, Jörg DJ, et al (2019) Defining the Identity and Dynamics of Adult Gastric Isthmus Stem Cells. Cell Stem Cell 25:342-356.e7. https://doi.org/10.1016/j.stem.2019.07.008

28. Bartfeld S, Koo BK (2017) Adult gastric stem cells and their niches. Wiley Interdiscip Rev Dev Biol 6:e261. https://doi.org/10.1002/wdev.261

29. Ye W, Takabayashi H, Yang Y, et al (2018) Regulation of Gastric Lgr5+ve Cell Homeostasis by Bone Morphogenetic Protein (BMP) Signaling and Inflammatory Stimuli. Cmgh 5:523–538. https://doi.org/10.1016/j.jcmgh.2018.01.007

30. Shinohara M, Mao M, Keeley TM, et al (2010) Bone morphogenetic protein signaling regulates gastric epithelial cell development and proliferation in mice. Gastroenterology 139:2050--2060.e2. https://doi.org/10.1053/j.gastro.2010.08.052

31. Kim T-H, Shivdasani RA (2011) Notch signaling in stomach epithelial stem cell homeostasis. J Exp Med 208:677–688. https://doi.org/10.1084/jem.20101737

32. Demitrack ES, Samuelson LC (2017) Notch as a Driver of Gastric Epithelial Cell Proliferation. Cmgh 3:323–330. https://doi.org/10.1016/j.jcmgh.2017.01.012

33. Spencer-Dene B, Sala FG, Bellusci S, et al (2006) Stomach Development Is Dependent on Fibroblast Growth Factor 10/Fibroblast Growth Factor Receptor 2b-Mediated Signaling. Gastroenterology 130:1233–1244. https://doi.org/10.1053/j.gastro.2006.02.018

34. Nyeng P, Norgaard GA, Kobberup S, Jensen J (2007) FGF10 signaling controls stomach morphogenesis. Dev Biol 303:295–310. https://doi.org/10.1016/j.ydbio.2006.11.017

35. Sigal M, Logan CY, Kapalczynska M, et al (2017) Stromal R-spondin orchestrates gastric epithelial stem cells and gland homeostasis. Nature 548:451–455. https://doi.org/10.1038/nature23642

36. Goodlad RA, Wilson TJG, Lenton W, et al (1987) Proliferative effects of urogastrone-EGF on the intestinal epithelium. Gut 28:37–43. https://doi.org/10.1136/gut.28.Suppl.37

37. Osuntokun B, Kocoshis SA (2006) Anatomy and Physiology of the Small and Large Intestine, Third Edit. Elsevier Inc.

38. Treuting PM, Arends MJ, Dintzis SM (2018) Lower Gastrointestinal Tract. In: Comparative Anatomy and Histology. Elsevier, pp 213–228

39. Kiela PR, Ghishan FK (2016) Physiology of intestinal absorption and secretion. Best Pract Res Clin Gastroenterol 30:145–159. https://doi.org/10.1016/j.bpg.2016.02.007

40. Snoeck V, Goddeeris B, Cox E (2005) The role of enterocytes in the intestinal barrier function and antigen uptake. Microbes Infect 7:997–1004. https://doi.org/10.1016/j.micinf.2005.04.003

41. Edelblum KL, Turner JR (2015) Epithelial Cells: Structure, Transport, and Barrier Function. Structure, Transport, and Barrier Function, Fourth Edi. Elsevier

42. Knoop KA, Newberry RD (2018) Goblet cells: multifaceted players in immunity at mucosal surfaces. Mucosal Immunol 11:1551–1557. https://doi.org/10.1038/s41385-018-0039-y

43. Cheng H, Leblond CP (1974) Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine I. Columnar cell. Am J Anat 141:461–479. https://doi.org/10.1002/aja.1001410403

44. Schneider C, O'Leary CE, Locksley RM (2019) Regulation of immune responses by tuft cells. Nat Rev Immunol 19:584–593. https://doi.org/10.1038/s41577-019-0176-x

45. Carlson BM (2019) The Digestive System. The human body. In: The human body. Elsevier, pp 321–355

46. Johansson MEV, Holmén Larsson JM, Hansson GC (2011) The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. Proc Natl Acad Sci U S A 108:4659–4665. https://doi.org/10.1073/pnas.1006451107

47. Johansson MEV, Ambort D, Pelaseyed T, et al (2011) Composition and functional role of the mucus layers in the intestine. Cell Mol Life Sci 68:3635–3641. https://doi.org/10.1007/s00018-011-0822-3

48. Sato T, Van Es JH, Snippert HJ, et al (2011) Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. Nature 469:415–418. https://doi.org/10.1038/nature09637

49. Eroschenko VP (2017) Chapter 15: Digestive System Part III: Small Intestine and Large Intestine. In: Atlas of Histology with Functional Correlations, 13th ed. Wolters Kluwer Health, pp 585–620

50. Rothenberg ME, Nusse Y, Kalisky T, et al (2012) Identification of a cKit+ colonic crypt base secretory cell that supports Lgr5+ stem cells in mice. Gastroenterology 142:1195--1205.e6. https://doi.org/10.1053/j.gastro.2012.02.006

51. Sasaki N, Sachs N, Wiebrands K, et al (2016) Reg4+ deep crypt secretory cells function as epithelial niche for Lgr5+ stem cells in colon. Proc Natl Acad Sci U S A 113:E5399–E5407. https://doi.org/10.1073/pnas.1607327113

52. Mabbott NA, Donaldson DS, Ohno H, et al (2013) Microfold (M) cells: Important immunosurveillance posts in the intestinal epithelium. Mucosal Immunol 6:666–677. https://doi.org/10.1038/mi.2013.30

53. Basak O, Born M, Korving J, et al (2014) Mapping early fate determination in <scp>L</scp> gr5 + crypt stem cells using a novel <scp> *K* </scp> *i67-* <scp> *RFP* </scp> allele. EMBO J 33:2057–2068. https://doi.org/10.15252/embj.201488017

54. Bjerknes M, Cheng H (1981) The stem-cell zone of the small intestinal epithelium. III. Evidence from columnar, enteroendocrine, and mucous cells in the adult mouse. Am J Anat 160:77–91. https://doi.org/10.1002/aja.1001600107

55. Haramis APG, Begthel H, Van Den Born M, et al (2004) De Novo Crypt Formation and Juvenile Polyposis on BMP Inhibition in Mouse Intestine. Science (80-) 303:1684–1686. https://doi.org/10.1126/science.1093587

56. He XC, Zhang J, Tong WG, et al (2004) BMP signaling inhibits intestinal stem cell self-renewal through suppression of Wnt-β-catenin signaling. Nat Genet 36:1117–1121. https://doi.org/10.1038/ng1430

57. De Lau W, Barker N, Low TY, et al (2011) Lgr5 homologues associate with Wnt receptors and mediate R-spondin signalling. Nature 476:293–297. https://doi.org/10.1038/nature10337

58. Bevins CL, Salzman NH (2011) Paneth cells, antimicrobial peptides and maintenance of intestinal homeostasis. Nat Rev Microbiol 9:356–368. https://doi.org/10.1038/nrmicro2546

59. Farin HF, Jordens I, Mosa MH, et al (2016) Visualization of a short-range Wnt gradient in the intestinal stem-cell niche. Nature 530:340–343. https://doi.org/10.1038/nature16937

60. Brügger MD, Valenta T, Fazilaty H, et al (2020) Distinct populations of crypt-associated fibroblasts act as signaling hubs to control colon homeostasis. PLOS Biol 18:e3001032. https://doi.org/10.1371/journal.pbio.3001032

61. Sato T, Stange DE, Ferrante M, et al (2011) Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. Gastroenterology 141:1762–1772.

https://doi.org/10.1053/j.gastro.2011.07.050

62. Kabiri Z, Greicius G, Madan B, et al (2014) Stroma provides an intestinal stem cell niche in the absence of epithelial Wnts. Dev 141:2206–2215. https://doi.org/10.1242/dev.104976

63. Farin HF, Van Es JH, Clevers H (2012) Redundant sources of Wnt regulate intestinal stem cells and promote formation of paneth cells. Gastroenterology 143:1518--1529.e7. https://doi.org/10.1053/j.gastro.2012.08.031

64. Clevers H (2013) The Intestinal Crypt, A Prototype Stem Cell Compartment. Cell 154:274–284. https://doi.org/10.1016/j.cell.2013.07.004

65. Sailaja BS, He XC, Li L (2016) The regulatory niche of intestinal stem cells. J Physiol 594:4827–4836. https://doi.org/10.1113/JP271931

66. San Roman AK, Kim T-H, Shivdasani RA (2016) The Alimentary Canal. In: Kaufman's Atlas of Mouse Development Supplement. Elsevier, pp 77–84

67. Zorn AM, Wells JM (2009) Vertebrate endoderm development and organ formation. Annu Rev Cell Dev Biol 25:221–251. https://doi.org/10.1146/annurev.cellbio.042308.113344

68. Sherwood RI, Maehr R, Mazzoni EO, Melton DA (2011) Wnt signaling specifies and patterns intestinal endoderm. Mech Dev 128:387–400. https://doi.org/10.1016/j.mod.2011.07.005

69. Sherwood RI, Chen TYA, Melton DA (2009) Transcriptional dynamics of endodermal organ formation. Dev Dyn 238:29–42. https://doi.org/10.1002/dvdy.21810

70. Gao N, White P, Kaestner KH (2009) Establishment of Intestinal Identity and Epithelial-Mesenchymal Signaling by Cdx2. Dev Cell 16:588–599. https://doi.org/10.1016/j.devcel.2009.02.010

71. Grainger S, Savory JGA, Lohnes D (2010) Cdx2 regulates patterning of the intestinal epithelium. Dev Biol 339:155–165. https://doi.org/10.1016/j.ydbio.2009.12.025

72. Stringer EJ, Duluc I, Saandi T, et al (2012) Cdx2 determines the fate of postnatal intestinal endoderm. Development 139:465–474. https://doi.org/10.1242/dev.070722

73. Simmini S, Bialecka M, Huch M, et al (2014) Transformation of intestinal stem cells into gastric stem cells on loss of transcription factor Cdx2. Nat Commun 5:5728. https://doi.org/10.1038/ncomms6728

74. Kim BM, Buchner G, Miletich I, et al (2005) The stomach mesenchymal transcription factor barx1 specifies gastric epithelial identity through inhibition

of transient Wnt signaling. Dev Cell 8:611–622. https://doi.org/10.1016/j.devcel.2005.01.015

75. Que J, Okubo T, Goldenring JR, et al (2007) Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm. Development 134:2521–2531. https://doi.org/10.1242/dev.003855

76. Rodriguez P, Da Silva S, Oxburgh L, et al (2010) BMP signaling in the development of the mouse esophagus and forestomach. Development 137:4171–4176. https://doi.org/10.1242/dev.056077

77. Barker N (2014) Adult intestinal stem cells: Critical drivers of epithelial homeostasis and regeneration. Nat Rev Mol Cell Biol 15:19–33. https://doi.org/10.1038/nrm3721

78. Muñoz J, Stange DE, Schepers AG, et al (2012) The Lgr5 intestinal stem cell signature: Robust expression of proposed quiescent ′+4′ cell markers. EMBO J 31:3079–3091. https://doi.org/10.1038/emboj.2012.166

79. Bartfeld S, Clevers H (2017) Stem cell-derived organoids and their application for medical research and patient treatment. J Mol Med 95:729–738. https://doi.org/10.1007/s00109-017-1531-7

80. Lancaster M, Takebe T, Lancaster M (2017) Advances in Organoid Technology: Hans Clevers, Madeline Lancaster, and Takanori Takebe. Cell Stem Cell 20:759–762. https://doi.org/10.1016/j.stem.2017.05.014

81. Kraiczy J, Nayak KM, Howell KJ, et al (2019) DNA methylation defines regional identity of human intestinal epithelial organoids and undergoes dynamic changes during development. Gut 68:49–61. https://doi.org/10.1136/gutjnl-2017-314817

82. Sato T, Vries RG, Snippert HJ, et al (2009) Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature 459:262–265. https://doi.org/10.1038/nature07935

83. Ootani A, Li X, Sangiorgi E, et al (2009) Sustained in vitro intestinal epithelial culture within a Wnt-dependent stem cell niche. Nat Med 15:701–706. https://doi.org/10.1038/nm.1951

84. Middendorp S, Schneeberger K, Wiegerinck CL, et al (2014) Adult stem cells in the small intestine are intrinsically programmed with their location-specific function. Stem Cells 32:1083–1091. https://doi.org/10.1002/stem.1655

85. Fukuda M, Mizutani T, Mochizuki W, et al (2014) Small intestinal stem cell identity is maintained with functional Paneth cells in heterotopically grafted epithelium onto the colon. Genes Dev 28:1752–1757. https://doi.org/10.1101/gad.245233.114

86. Bartfeld S, Bayram T, Van De Wetering M, et al (2015) In vitro expansion of human gastric epithelial stem cells and their responses to bacterial infection. Gastroenterology 148:126-136.e6. https://doi.org/10.1053/j.gastro.2014.09.042

87. Bartfeld S, Clevers H (2015) Organoids as model for infectious diseases: Culture of human and murine stomach organoids and microinjection of helicobacter pylori. J Vis Exp 2015:1–9. https://doi.org/10.3791/53359

88. Boccellato F, Woelffling S, Imai-Matsushima A, et al (2019) Polarised epithelial monolayers of the gastric mucosa reveal insights into mucosal homeostasis and defence against infection. Gut 68:400–413. https://doi.org/10.1136/gutjnl-2017-314540

89. Schlaermann P, Toelle B, Berger H, et al (2016) A novel human gastric primary cell culture system for modelling Helicobacter pylori infection in vitro. Gut 65:202–213. https://doi.org/10.1136/gutjnl-2014-307949

90. Stanifer ML, Mukenhirn M, Muenchau S, et al (2020) Asymmetric distribution of TLR3 leads to a polarized immune response in human intestinal epithelial cells. Nat Microbiol 5:181–191. https://doi.org/10.1038/s41564-019-0594-3

91. Navis M, Martins Garcia T, Renes IB, et al (2019) Mouse fetal intestinal organoids: new model to study epithelial maturation from suckling to weaning. EMBO Rep 20:e46221. https://doi.org/10.15252/embr.201846221

92. Elmentaite R, Ross ADB, Roberts K, et al (2020) Single-Cell Sequencing of Developing Human Gut Reveals Transcriptional Links to Childhood Crohn's Disease. Dev Cell 55:771-783.e5. https://doi.org/10.1016/j.devcel.2020.11.010

93. Liwinski T, Zheng D, Elinav E (2020) The microbiome and cytosolic innate immune receptors. Immunol Rev 297:207–224. https://doi.org/10.1111/imr.12901

94. Takeuchi O, Akira S (2001) Toll-like receptors; their physiological role and signal transduction system. Int Immunopharmacol 1:625–635. https://doi.org/10.1016/S1567-5769(01)00010-8

95. Abreu MT (2010) Toll-like receptor signalling in the intestinal epithelium: How bacterial recognition shapes intestinal function. Nat Rev Immunol 10:131–143. https://doi.org/10.1038/nri2707

96. Yu S, Gao N (2015) Compartmentalizing intestinal epithelial cell toll-like receptors for immune surveillance. Cell Mol Life Sci 72:3343–3353. https://doi.org/10.1007/s00018-015-1931-1

97. Zhang K, Hornef MW, Dupont A (2015) The intestinal epithelium as guardian of gut barrier integrity. Cell Microbiol 17:1561–1569. https://doi.org/10.1111/cmi.12501

98. Hutcheon C, Paulvannan P, Subramanian N (2016) Cytoplasmic Sensing in Innate Immunity. Elsevier Ltd.

99. Fitzgerald KA, Kagan JC (2020) Toll-like Receptors and the Control of Immunity. Cell 180:1044–1066. https://doi.org/10.1016/j.cell.2020.02.041

100. Wyllie DH, Kiss-Toth E, Visintin A, et al (2000) Evidence for an Accessory Protein Function for Toll-Like Receptor 1 in Anti-Bacterial Responses. J Immunol 165:7125–7132. https://doi.org/10.4049/jimmunol.165.12.7125

101. Takeuchi O, Sato S, Horiuchi T, et al (2002) Cutting Edge: Role of Toll-Like Receptor 1 in Mediating Immune Response to Microbial Lipoproteins. J Immunol 169:10–14. https://doi.org/10.4049/jimmunol.169.1.10

102. Takeuchi O, Kawai T, Mühlradt PF, et al (2001) Discrimination of bacterial lipoproteins by Toll-like recepttor 6. Int Immunol 13:933–940. https://doi.org/10.1093/intimm/13.7.933

103. Alexopoulou L, Holt AC, Medzhitov R, Flavell RA (2001) Recognition of double-stranded RNA and activation of NF-κB by Toll-like receptor 3. Nature 413:732–738. https://doi.org/10.1038/35099560

104. Gewirtz AT, Navas TA, Lyons S, et al (2001) Cutting Edge: Bacterial Flagellin Activates Basolaterally Expressed TLR5 to Induce Epithelial Proinflammatory Gene Expression. J Immunol 167:1882–1885. https://doi.org/10.4049/jimmunol.167.4.1882

105. Hayashi F, Smith KD, Ozinsky A, et al (2001) The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. Nature 410:1099–1103. https://doi.org/10.1038/35074106

106. Diebold SS, Kaisho T, Hemmi H, et al (2004) Innate Antiviral Responses by Means of TLR7-Mediated Recognition of Single-Stranded RNA. Science (80-) 303:1529–1531. https://doi.org/10.1126/science.1093616

107. Krüger A, Oldenburg M, Chebrolu C, et al (2015) Human TLR 8 senses UR / URR motifs in bacterial and mitochondrial RNA. EMBO Rep 16:1656–1663. https://doi.org/10.15252/embr.201540861

108. Hemmi H, Takeuchi O, Kawai T, et al (2000) A Toll-like receptor recognizes bacterial DNA. Nature 408:740–745. https://doi.org/10.1038/35047123

109. Yarovinsky F, Zhang D, Andersen JF, et al (2005) Immunology: TLR11 activation of dendritic cells by a protozoan profilin-like protein. Science (80-) 308:1626–1629. https://doi.org/10.1126/science.1109893

110. Hidmark A, von Saint Paul A, Dalpke AH (2012) Cutting Edge: TLR13 Is a Receptor for Bacterial RNA. J Immunol 189:2717–2721. https://doi.org/10.4049/jimmunol.1200898

111. Hornef MW, Normark BH, Vandewalle A, Normark S (2003) Intracellular Recognition of Lipopolysaccharide by Toll-like Receptor 4 in Intestinal Epithelial Cells. J Exp Med 198:1225–1235. https://doi.org/10.1084/jem.20022194

112. Lee J, Gonzales-Navajas JM, Raz E (2008) The "polarizing-tolerizing" mechanism of intestinal epithelium: Its relevance to colonic homeostasis. Semin Immunopathol 30:3–9. https://doi.org/10.1007/s00281-007-0099-7

113. Gay NJ, Symmons MF, Gangloff M, Bryant CE (2014) Assembly and localization of Toll-like receptor signalling complexes. Nat Rev Immunol 14:546–558. https://doi.org/10.1038/nri3713

114. Park BS, Song DH, Kim HM, et al (2009) The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. Nature 458:1191–1195. https://doi.org/10.1038/nature07830

115. Hailman E, Lichenstein HS, Wurfel MM, et al (1994) Lipopolysaccharide (LPS)-binding protein accelerates the binding of LPS to CD14. J Exp Med 179:269–277. https://doi.org/10.1084/jem.179.1.269

116. Tan Y, Kagan JC (2014) A cross-disciplinary perspective on the innate immune responses to bacterial lipopolysaccharide. Mol Cell 54:212–223. https://doi.org/10.1016/j.molcel.2014.03.012

117. Ryu JK, Kim SJ, Rah SH, et al (2017) Reconstruction of LPS Transfer Cascade Reveals Structural Determinants within LBP, CD14, and TLR4-MD2 for Efficient LPS Recognition and Transfer. Immunity 46:38–50. https://doi.org/10.1016/j.immuni.2016.11.007

118. Ting JPY, Lovering RC, Alnemri ES, et al (2008) The NLR Gene Family: A Standard Nomenclature. Immunity 28:285–287. https://doi.org/10.1016/j.immuni.2008.02.005

119. Martinon F, Burns K, Tschopp J (2002) The Inflammasome: A molecular platform triggering activation of inflammatory caspases and processing of proIL-β. Mol Cell 10:417–426. https://doi.org/10.1016/S1097-2765(02)00599-3

120. Schroder K, Tschopp J (2010) The Inflammasomes. Cell 140:821–832. https://doi.org/10.1016/j.cell.2010.01.040

121. Girardin SE, Tournebize R, Mavris M, et al (2001) CARD4/Nod1 mediates NF-KB and JNK activation by invasive Shigella flexneri. EMBO Rep 2:736–742. https://doi.org/10.1093/embo-reports/kve155

122. Inohara N, Ogura Y, Chen FF, et al (2001) Human Nod1 Confers Responsiveness to Bacterial Lipopolysaccharides. J Biol Chem 276:2551–2554. https://doi.org/10.1074/jbc.M009728200

123. Barbé F, Douglas T, Saleh M (2014) Advances in Nod-like receptors (NLR)

biology. Cytokine Growth Factor Rev 25:681–697. https://doi.org/10.1016/j.cytogfr.2014.07.001

124. Elinav E, Strowig T, Henao-Mejia J, Flavell RA (2011) Regulation of the Antimicrobial Response by NLR Proteins. Immunity 34:665–679. https://doi.org/10.1016/j.immuni.2011.05.007

125. Windheim M, Lang C, Peggie M, et al (2007) Molecular mechanisms involved in the regulation of cytokine production by muramyl dipeptide. Biochem J 404:179–190. https://doi.org/10.1042/BJ20061704

126. Abbott DW, Yang Y, Hutti JE, et al (2007) Coordinated Regulation of Toll-Like Receptor and NOD2 Signaling by K63-Linked Polyubiquitin Chains. Mol Cell Biol 27:6012–6025. https://doi.org/10.1128/mcb.00270-07

127. Agostini L, Martinon F, Burns K, et al (2004) NALP3 forms an IL-1β-processing inflammasome with increased activity in Muckle-Wells autoinflammatory disorder. Immunity 20:319–325. https://doi.org/10.1016/S1074-7613(04)00046-9

128. Broz P, Pelegrín P, Shao F (2020) The gasdermins, a protein family executing cell death and inflammation. Nat Rev Immunol 20:143–157. https://doi.org/10.1038/s41577-019-0228-2

129. Swanson K V., Deng M, Ting JPY (2019) The NLRP3 inflammasome: molecular activation and regulation to therapeutics. Nat Rev Immunol 19:477–489. https://doi.org/10.1038/s41577-019-0165-0

130. Hoving JC, Wilson GJ, Brown GD (2014) Signalling C-type lectin receptors, microbial recognition and immunity. Cell. Microbiol. 16:185–194

131. Zhou P, She Y, Dong N, et al (2018) Alpha-kinase 1 is a cytosolic innate immune receptor for bacterial ADP-heptose. Nature 561:122–126. https://doi.org/10.1038/s41586-018-0433-3

132. Milivojevic M, Dangeard AS, Kasper CA, et al (2017) ALPK1 controls TIFA/TRAF6-dependent innate immunity against heptose-1,7-bisphosphate of gram-negative bacteria. PLoS Pathog 13:e1006224. https://doi.org/10.1371/journal.ppat.1006224

133. Zimmermann S, Pfannkuch L, Al-Zeer MA, et al (2017) ALPK1- and TIFA-Dependent Innate Immune Response Triggered by the Helicobacter pylori Type IV Secretion System. Cell Rep 20:2384–2395. https://doi.org/10.1016/j.celrep.2017.08.039

134. Hornung V, Ablasser A, Charrel-Dennis M, et al (2009) AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. Nature 458:514–518. https://doi.org/10.1038/nature07725

135. Sun L, Wu J, Du F, et al (2013) Cyclic GMP-AMP synthase is a cytosolic DNA

sensor that activates the type I interferon pathway. Science (80-) 339:786–791. https://doi.org/10.1126/science.1232458

136. Shang G, Zhang C, Chen ZJ, et al (2019) Cryo-EM structures of STING reveal its mechanism of activation by cyclic GMP–AMP. Nature 567:389–393. https://doi.org/10.1038/s41586-019-0998-5

137. Ishikawa H, Ma Z, Barber GN (2009) STING regulates intracellular DNA-mediated, type i interferon-dependent innate immunity. Nature 461:788–792. https://doi.org/10.1038/nature08476

138. Wu J, Sun L, Chen X, et al (2013) Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. Science (80-) 339:826–830. https://doi.org/10.1126/science.1229963

139. Gui X, Yang H, Li T, et al (2019) Autophagy induction via STING trafficking is a primordial function of the cGAS pathway. Nature 567:262–266. https://doi.org/10.1038/s41586-019-1006-9

140. Brestoff JR, Artis D (2013) Commensal bacteria at the interface of host metabolism and the immune system. Nat Immunol 14:676–684. https://doi.org/10.1038/ni.2640

141. Abrams GD, Bauer H, Sprinz H (1963) Influence of the normal flora on mucosal morphology and cellular renewal in the ileum. A comparison of germ-free and conventional mice. Lab Invest 12:355–64

142. Round JL, Mazmanian SK (2009) The gut microbiota shapes intestinal immune responses during health and disease. Nat Rev Immunol 9:313–323. https://doi.org/10.1038/nri2515

143. Stecher B, Hardt WD (2008) The role of microbiota in infectious disease. Trends Microbiol 16:107–114. https://doi.org/10.1016/j.tim.2007.12.008

144. Wells JM, Rossia O, Meijerink M, Van Baarlen P (2011) Epithelial crosstalk at the microbiota-mucosal interface. Proc Natl Acad Sci U S A 108:4607–4614. https://doi.org/10.1073/pnas.1000092107

145. Abraham C, Medzhitov R (2011) Interactions between the host innate immune system and microbes in inflammatory bowel disease. Gastroenterology 140:1729–1737. https://doi.org/10.1053/j.gastro.2011.02.012

146. Nenci A, Becker C, Wullaert A, et al (2007) Epithelial NEMO links innate immunity to chronic intestinal inflammation. Nature 446:557–561. https://doi.org/10.1038/nature05698

147. Wullaert A, Bonnet MC, Pasparakis M (2011) NF-κB in the regulation of epithelial homeostasis and inflammation. Cell Res 21:146–158. https://doi.org/10.1038/cr.2010.175

148. Cario E, Gerken G, Podolsky DK (2007) Toll-Like Receptor 2 Controls Mucosal Inflammation by Regulating Epithelial Barrier Function. Gastroenterology 132:1359–1374. https://doi.org/10.1053/j.gastro.2007.02.056

149. Elinav E, Strowig T, Kau AL, et al (2011) NLRP6 inflammasome regulates colonic microbial ecology and risk for colitis. Cell 145:745–757. https://doi.org/10.1016/j.cell.2011.04.022

150. Rakoff-Nahoum S, Paglino J, Eslami-Varzaneh F, et al (2004) Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. Cell 118:229–241. https://doi.org/10.1016/j.cell.2004.07.002

151. Rhee SH, Im E, Riegler M, et al (2005) Pathophysiological role of Toll-like receptor 5 engagement by bacterial flagenllin in colonic inflammation. Proc Natl Acad Sci U S A 102:13610–13615. https://doi.org/10.1073/pnas.0502174102

152. Wlodarska M, Thaiss CA, Nowarski R, et al (2014) NLRP6 inflammasome orchestrates the colonic host-microbial interface by regulating goblet cell mucus secretion. Cell 156:1045–1059. https://doi.org/10.1016/j.cell.2014.01.026

153. Levy A, Stedman A, Deutsch E, et al (2020) Innate immune receptor NOD2 mediates LGR5+ intestinal stem cell protection against ROS cytotoxicity via mitophagy stimulation. Proc Natl Acad Sci U S A 117:1994–2003. https://doi.org/10.1073/pnas.1902788117

154. Nigro G, Rossi R, Commere PH, et al (2014) The cytosolic bacterial peptidoglycan sensor Nod2 affords stem cell protection and links microbes to gut epithelial regeneration. Cell Host Microbe 15:792–798. https://doi.org/10.1016/j.chom.2014.05.003

155. Gaya DR, Russell RK, Nimmo ER, Satsangi J (2006) New genes in inflammatory bowel disease: lessons for complex diseases? Lancet 367:1271–1284. https://doi.org/10.1016/S0140-6736(06)68345-1

156. Kaser A, Zeissig S, Blumberg RS (2010) Inflammatory bowel disease. Annu Rev Immunol 28:573–621. https://doi.org/10.1146/annurev-immunol-030409-101225

157. Bäckhed F, Hornef M (2003) Toll-like receptor 4-mediated signaling by epithelial surfaces: Necessity or threat? Microbes Infect 5:951–959. https://doi.org/10.1016/S1286-4579(03)00189-8

158. Ortega-Cava CF, Ishihara S, Rumi MAK, et al (2003) Strategic Compartmentalization of Toll-Like Receptor 4 in the Mouse Gut. J Immunol 170:3977–3985. https://doi.org/10.4049/jimmunol.170.8.3977

159. Price AE, Shamardani K, Lugo KA, et al (2018) A Map of Toll-like Receptor Expression in the Intestinal Epithelium Reveals Distinct Spatial, Cell Type-Specific, and Temporal Patterns. Immunity 49:560-575.e6.

https://doi.org/10.1016/j.immuni.2018.07.016

160. Cario E, Gerken G, Podolsky DK (2002) "For Whom the Bell Tolls!" - Innate Defense Mechanisms and Survival Strategies of the Intestinal Epithelium Against Lumenal Pathogens. Z Gastroenterol 40:983–990. https://doi.org/10.1055/s-2002-36159

161. Schlegel N, Boerner K, Waschke J (2021) Targeting desmosomal adhesion and signalling for intestinal barrier stabilization in inflammatory bowel diseases—Lessons from experimental models and patients. Acta Physiol 231:e13492. https://doi.org/10.1111/apha.13492

162. Grivennikov SI, Greten FR, Karin M (2010) Immunity, Inflammation, and Cancer. Cell 140:883–899. https://doi.org/10.1016/j.cell.2010.01.025

163. Kavanagh ME, O'Sullivan KE, O'Hanlon C, et al (2014) The esophagitis to adenocarcinoma sequence; the role of inflammation. Cancer Lett 345:182–189. https://doi.org/10.1016/j.canlet.2013.08.017

164. Barsouk A, Rawla P, Barsouk A, Thandra KC (2019) Epidemiology of Cancers of the Small Intestine: Trends, Risk Factors, and Prevention. Med Sci 7:46. https://doi.org/10.3390/medsci7030046

165. Lee J, Mo JH, Katakura K, et al (2006) Maintenance of colonic homeostasis by distinctive apical TLR9 signalling in intestinal epithelial cells. Nat Cell Biol 8:1327–1336. https://doi.org/10.1038/ncb1500

166. Hornef MW, Torow N (2020) 'Layered immunity' and the 'neonatal window of opportunity' – timed succession of non-redundant phases to establish mucosal host–microbial homeostasis after birth. Immunology 159:15–25. https://doi.org/10.1111/imm.13149

167. Renz H, Adkins BD, Bartfeld S, et al (2018) The neonatal window of opportunity—early priming for life. J Allergy Clin Immunol 141:1212–1214. https://doi.org/10.1016/j.jaci.2017.11.019

168. Torow N, Hornef MW (2017) The Neonatal Window of Opportunity: Setting the Stage for Life-Long Host-Microbial Interaction and Immune Homeostasis. J Immunol 198:557–563. https://doi.org/10.4049/jimmunol.1601253

169. Lotz M, Gütle D, Walther S, et al (2006) Postnatal acquisition of endotoxin tolerance in intestinal epithelial cells. J Exp Med 203:973–984. https://doi.org/10.1084/jem.20050625

170. Gribar SC, Sodhi CP, Richardson WM, et al (2009) Reciprocal Expression and Signaling of TLR4 and TLR9 in the Pathogenesis and Treatment of Necrotizing Enterocolitis. J Immunol 182:636–646. https://doi.org/10.4049/jimmunol.182.1.636

171. Pott J, Stockinger S, Torow N, et al (2012) Age-dependent TLR3 expression of the intestinal epithelium contributes to rotavirus susceptibility. PLoS Pathog 8:. https://doi.org/10.1371/journal.ppat.1002670

172. Kobayashi KS (2005) Nod2-Dependent Regulation of Innate and Adaptive Immunity in the Intestinal Tract. Science (80-) 307:731–734. https://doi.org/10.1126/science.1104911

173. Naito T, Mulet C, De Castro C, et al (2017) Lipopolysaccharide from crypt-specific core microbiota modulates the colonic epithelial proliferation-to-differentiation balance. MBio 8:1–16. https://doi.org/10.1128/mBio.01680-17

174. Neal MD, Sodhi CP, Jia H, et al (2012) Toll-like receptor 4 is expressed on intestinal stem cells and regulates their proliferation and apoptosis via the p53 up-regulated modulator of apoptosis. J Biol Chem 287:37296–37308. https://doi.org/10.1074/jbc.M112.375881

175. Tanabe H, Ayabe T, Bainbridge B, et al (2005) Mouse paneth cell secretory responses to cell surface glycolipids of virulent and attenuated pathogenic bacteria. Infect Immun 73:2312–2320. https://doi.org/10.1128/IAI.73.4.2312-2320.2005

176. Kayisoglu O, Weiss F, Niklas C, et al (2021) Location-specific cell identity rather than exposure to GI microbiota defines many innate immune signalling cascades in the gut epithelium. Gut 70:687–697. https://doi.org/10.1136/gutjnl-2019-319919

177. Geiszt M, Lekstrom K, Brenner S, et al (2003) NAD(P)H Oxidase 1, a Product of Differentiated Colon Epithelial Cells, Can Partially Replace Glycoprotein 91 phox in the Regulated Production of Superoxide by Phagocytes . J Immunol 171:299–306. https://doi.org/10.4049/jimmunol.171.1.299

178. Reimand J, Arak T, Adler P, et al (2016) g:Profiler-a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res 44:W83–W89. https://doi.org/10.1093/nar/gkw199

179. Hoffmann A, Levchenko A, Scott ML, Baltimore D (2002) The IκB-NF-κB signaling module: Temporal control and selective gene activation. Science (80-) 298:1241–1245. https://doi.org/10.1126/science.1071914

180. Bartfeld S, Hess S, Bauer B, et al (2010) High-throughput and single-cell imaging of NF-κB oscillations using monoclonal cell lines. BMC Cell Biol 11:21. https://doi.org/10.1186/1471-2121-11-21

181. Hornef MW, Penders J (2017) Does a prenatal bacterial microbiota exist? Mucosal Immunol 10:598–601. https://doi.org/10.1038/mi.2016.141

182. Fordham RP, Yui S, Hannan NRF, et al (2013) Transplantation of expanded

fetal intestinal progenitors contributes to colon regeneration after injury. Cell Stem Cell 13:734–744. https://doi.org/10.1016/j.stem.2013.09.015

183. Howell KJ, Kraiczy J, Nayak KM, et al (2018) DNA Methylation and Transcription Patterns in Intestinal Epithelial Cells From Pediatric Patients With Inflammatory Bowel Diseases Differentiate Disease Subtypes and Associate With Outcome. Gastroenterology 154:585–598. https://doi.org/10.1053/j.gastro.2017.10.007

184. Buenrostro JD, Giresi PG, Zaba LC, et al (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 10:1213–1218. https://doi.org/10.1038/nmeth.2688

185. Banerjee KK, Saxena M, Kumar N, et al (2018) Enhancer, transcriptional, and cell fate plasticity precedes intestinal determination during endoderm development. Genes Dev 32:1430–1442. https://doi.org/10.1101/gad.318832.118

186. Birchenough GMH, Nystrom EEL, Johansson MEV, Hansson GC (2016) A sentinel goblet cell guards the colonic crypt by triggering Nlrp6-dependent Muc2 secretion. Science (80-) 352:1535–1542. https://doi.org/10.1126/science.aaf7419

187. Ogura Y, Lala S, Xin W, et al (2003) Expression of NOD2 in Paneth cells: A possible link to Crohn's ileitis. Gut 52:1591–1597. https://doi.org/10.1136/gut.52.11.1591

188. Kozuka K, He Y, Koo-McCoy S, et al (2017) Development and Characterization of a Human and Mouse Intestinal Epithelial Cell Monolayer Platform. Stem Cell Reports 9:1976–1990. https://doi.org/10.1016/j.stemcr.2017.10.013

189. van der Flier LG, Haegebarth A, Stange DE, et al (2009) OLFM4 Is a Robust Marker for Stem Cells in Human Intestine and Marks a Subset of Colorectal Cancer Cells. Gastroenterology 137:15–17. https://doi.org/10.1053/j.gastro.2009.05.035

190. Saxena M, San Roman AK, O'Neill NK, et al (2017) Transcription factor-dependent 'anti-repressive' mammalian enhancers exclude H3K27me3 from extended genomic domains. Genes Dev 31:2391–2404. https://doi.org/10.1101/gad.308536.117

191. Kumar N, Tsai YH, Chen L, et al (2019) The lineage-specific transcription factor CDX2 navigates dynamic chromatin to control distinct stages of intestine development. Dev 146:dev172189. https://doi.org/10.1242/dev.172189

192. Verzi MP, Shin H, He HH, et al (2010) Differentiation-Specific Histone Modifications Reveal Dynamic Chromatin Interactions and Partners for the Intestinal Transcription Factor CDX2. Dev Cell 19:713–726.

https://doi.org/10.1016/j.devcel.2010.10.006

193. Gregorieff A, Pinto D, Begthel H, et al (2005) Expression Pattern of Wnt Signaling Components in the Adult Intestine. Gastroenterology 129:626–638. https://doi.org/10.1053/j.gastro.2005.06.007

194. Fujii M, Matano M, Toshimitsu K, et al (2018) Human Intestinal Organoids Maintain Self-Renewal Capacity and Cellular Diversity in Niche-Inspired Culture Condition. Cell Stem Cell 23:787-793.e6. https://doi.org/10.1016/j.stem.2018.11.016

195. Davies JM, Santaolalla R, Von Furstenberg RJ, et al (2015) The viral mimetic polyinosinic: Polycytidylic acid alters the growth characteristics of small intestinal and colonic crypt cultures. PLoS One 10:1–19. https://doi.org/10.1371/journal.pone.0138531

196. Farin HF, Karthaus WR, Kujala P, et al (2014) Paneth cell extrusion and release of antimicrobial products is directly controlled by immune cell-derived IFN-γ. J Exp Med 211:1393–1405. https://doi.org/10.1084/jem.20130753

197. Shaw-Smith CJ, Walters JRF (1997) Regional expression of intestinal genes for nutrient absorption. Gut 40:5–8. https://doi.org/10.1136/gut.40.1.5

198. Johansson MEV, Sjövall H, Hansson GC (2013) The gastrointestinal mucus system in health and disease. Nat Rev Gastroenterol Hepatol 10:352–361. https://doi.org/10.1038/nrgastro.2013.35

199. Koelz HR (1992) Gastric acid in vertebrates. Scand J Gastroenterol 27:2–6. https://doi.org/10.3109/00365529209095998

200. Cheng H (1974) Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine IV. Paneth cells. Am J Anat 141:521–535. https://doi.org/10.1002/aja.1001410406

201. Cheng H (1974) Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine II. Mucous cells. Am J Anat 141:481–501. https://doi.org/10.1002/aja.1001410404

202. Chang WWL, Leblond CP (1971) Renewal of the epithelium in the descending colon of the mouse. I. Presence of three cell populations: Vacuolated-columnar, mucous and argentaffin. Am J Anat 131:73–99. https://doi.org/10.1002/aja.1001310105

203. Ermund A, Schütte A, Johansson MEV, et al (2013) Studies of mucus in mouse stomach, small intestine, and colon. I. Gastrointestinal mucus layers have different properties depending on location as well as over the Peyer's patches. Am J Physiol - Gastrointest Liver Physiol 305:341–347. https://doi.org/10.1152/ajpgi.00046.2013

204. Park BS, Lee JO (2013) Recognition of lipopolysaccharide pattern by TLR4 complexes. Exp Mol Med 45:e66. https://doi.org/10.1038/emm.2013.97

205. Abreu MT, Vora P, Faure E, et al (2001) Decreased Expression of Toll-Like Receptor-4 and MD-2 Correlates with Intestinal Epithelial Cell Protection Against Dysregulated Proinflammatory Gene Expression in Response to Bacterial Lipopolysaccharide. J Immunol 167:1609–1616. https://doi.org/10.4049/jimmunol.167.3.1609

206. Bertani B, Ruiz N (2018) Function and Biogenesis of Lipopolysaccharides. EcoSal Plus 8:ecosalplus.ESP-0001-2018. https://doi.org/10.1128/ecosalplus.ESP-0001-2018

207. Akashi S, Nagai Y, Ogata H, et al (2001) Human MD-2 confers on mouse toll-like receptor 4 species-specific lipopolysaccharide recognition. Int Immunol 13:1595–1599. https://doi.org/10.1093/intimm/13.12.1595

208. Hajjar AM, Tsai JH, Wilson CB, Miller SI (2002) Human Toll-like receptor 4 recognizes host-specific LPS modifications. Nat Immunol 3:354–359. https://doi.org/10.1038/ni777

209. Fontaine CA, Skorupski AM, Vowles CJ, et al (2015) How free of germs is germ-free? Detection of bacterial contamination in a germ free mouse unit. Gut Microbes 6:225–233. https://doi.org/10.1080/19490976.2015.1054596

210. Ortega-Cava CF, Ishihara S, Rumi MAK, et al (2006) Epithelial toll-like receptor 5 is constitutively localized in the mouse cecum and exhibits distinctive down-regulation during experimental colitis. Clin Vaccine Immunol 13:132–138. https://doi.org/10.1128/CVI.13.1.132-138.2006

211. Gourbeyre P, Berri M, Lippi Y, et al (2015) Pattern recognition receptors in the gut: Analysis of their expression along the intestinal tract and the crypt/villus axis. Physiol Rep 3:e12225. https://doi.org/10.14814/phy2.12225

212. Mowat AM, Agace WW (2014) Regional specialization within the intestinal immune system. Nat Rev Immunol 14:667–685. https://doi.org/10.1038/nri3738

213. Hou B, Reizis B, DeFranco AL (2008) Toll-like Receptors Activate Innate and Adaptive Immunity by using Dendritic Cell-Intrinsic and -Extrinsic Mechanisms. Immunity 29:272–282. https://doi.org/10.1016/j.immuni.2008.05.016

214. Weiss DS, Takeda K, Akira S, et al (2005) MyD88, but not toll-like receptors 4 and 2, is required for efficient clearance of Brucella abortus. Infect Immun 73:5137–5143. https://doi.org/10.1128/IAI.73.8.5137-5143.2005

215. De Lorenzi R, Gareus R, Fengler S, Pasparakis M (2009) GFP-p65 knock-in mice as a tool to study NF-κB dynamics in vivo. Genesis 47:323–329.

https://doi.org/10.1002/dvg.20468

216. Sudo T, Iwaya T, Nishida N, et al (2013) Expression of mesenchymal markers vimentin and fibronectin: The clinical significance in esophageal squamous cell carcinoma. Ann Surg Oncol 20:324–335. https://doi.org/10.1245/s10434-012-2418-z

217. Kumar A, Shamsuddin N (2012) Retinal muller glia initiate innate response to infectious stimuli via toll-like receptor signaling. PLoS One 7:. https://doi.org/10.1371/journal.pone.0029830

218. Hoffmann S, Otto C, Kurtz S, et al (2009) Fast mapping of short sequences with mismatches, insertions and deletions using index structures. PLoS Comput Biol 5:e1000502. https://doi.org/10.1371/journal.pcbi.1000502

219. Förstner KU, Vogel J, Sharma CM (2014) READemption-a tool for the computational analysis of deep-sequencing-based transcriptome data. Bioinformatics 30:3421–3423. https://doi.org/10.1093/bioinformatics/btu533

220. Anders S, Pyl PT, Huber W (2015) HTSeq-A Python framework to work with high-throughput sequencing data. Bioinformatics 31:166–169. https://doi.org/10.1093/bioinformatics/btu638

221. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17:10. https://doi.org/10.14806/ej.17.1.200

222. Risso D, Ngai J, Speed TP, Dudoit S (2014) Normalization of RNA-seq data using factor analysis of control genes or samples. Nat Biotechnol 32:896–902. https://doi.org/10.1038/nbt.2931

223. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550. https://doi.org/10.1186/s13059-014-0550-8

224. Gu Z, Eils R, Schlesner M (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32:2847–2849. https://doi.org/10.1093/bioinformatics/btw313

225. Barski A, Cuddapah S, Cui K, et al (2007) High-Resolution Profiling of Histone Methylations in the Human Genome. Cell 129:823–837. https://doi.org/10.1016/j.cell.2007.05.009

226. Heintzman ND, Hon GC, Hawkins RD, et al (2009) Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature 459:108–112. https://doi.org/10.1038/nature07829

227. Heinz S, Romanoski CE, Benner C, Glass CK (2015) The selection and function of cell type-specific enhancers. Nat Rev Mol Cell Biol 16:144–154. https://doi.org/10.1038/nrm3949

228. Jiang S, Mortazavi A (2018) Integrating ChIP-seq with other functional genomics data. Brief Funct Genomics 17:104–115. https://doi.org/10.1093/bfgp/ely002

9 Supplementary Data

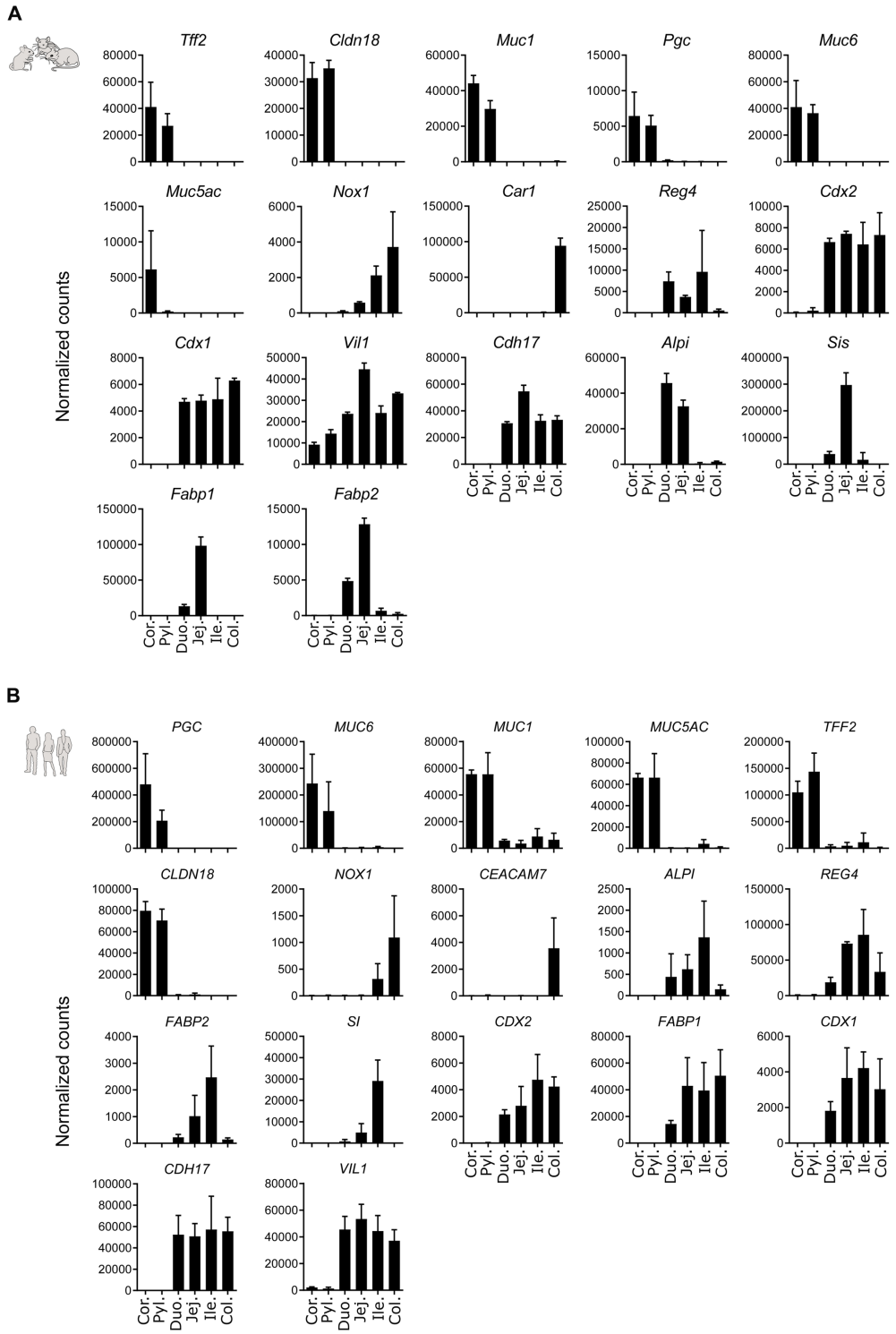


Figure S 1: Organoids express tissue identity markers in a segment-specific manner. The normalized gene counts of selected gastric and intestinal cell marker genes for murine **(A)** and human **(B)** organoids, which were depicted in the heatmaps (**Figure 15**) in the same order. Bars represent mean ± SD, n = 3 biological samples for each group.

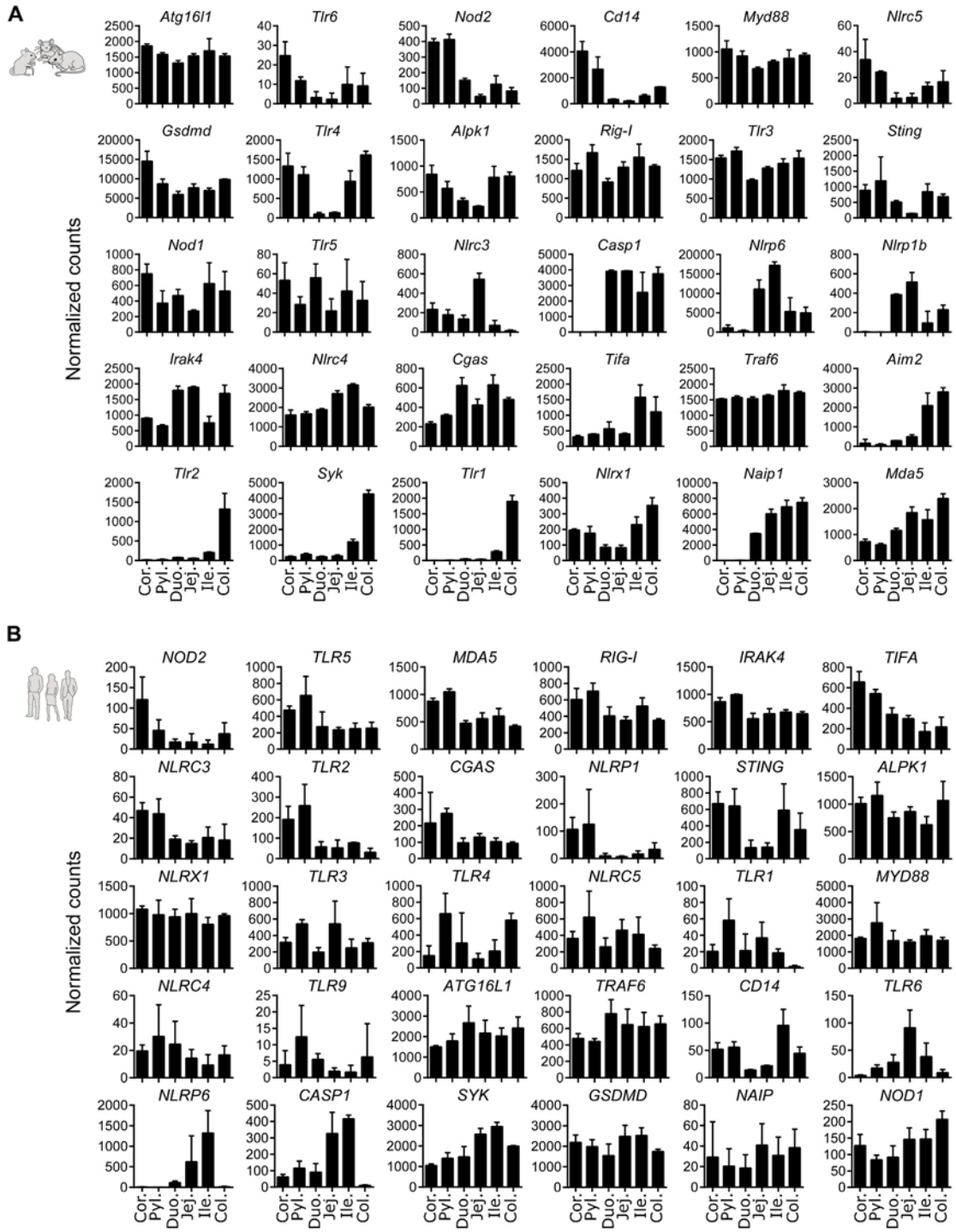


Figure S 2: **Transcriptome profiling in the human and murine GI tract reveals segment-specific patterning of PRR signaling components.** Graphs depict the normalized gene counts scaled in the heatmaps of **Figure 19** in the same order. **(A)** Murine genes related to PRR signaling pathways. **(B)** Human genes related to PRR signaling pathways. Bars represent mean ± SD, n = 3 biological samples for each group.

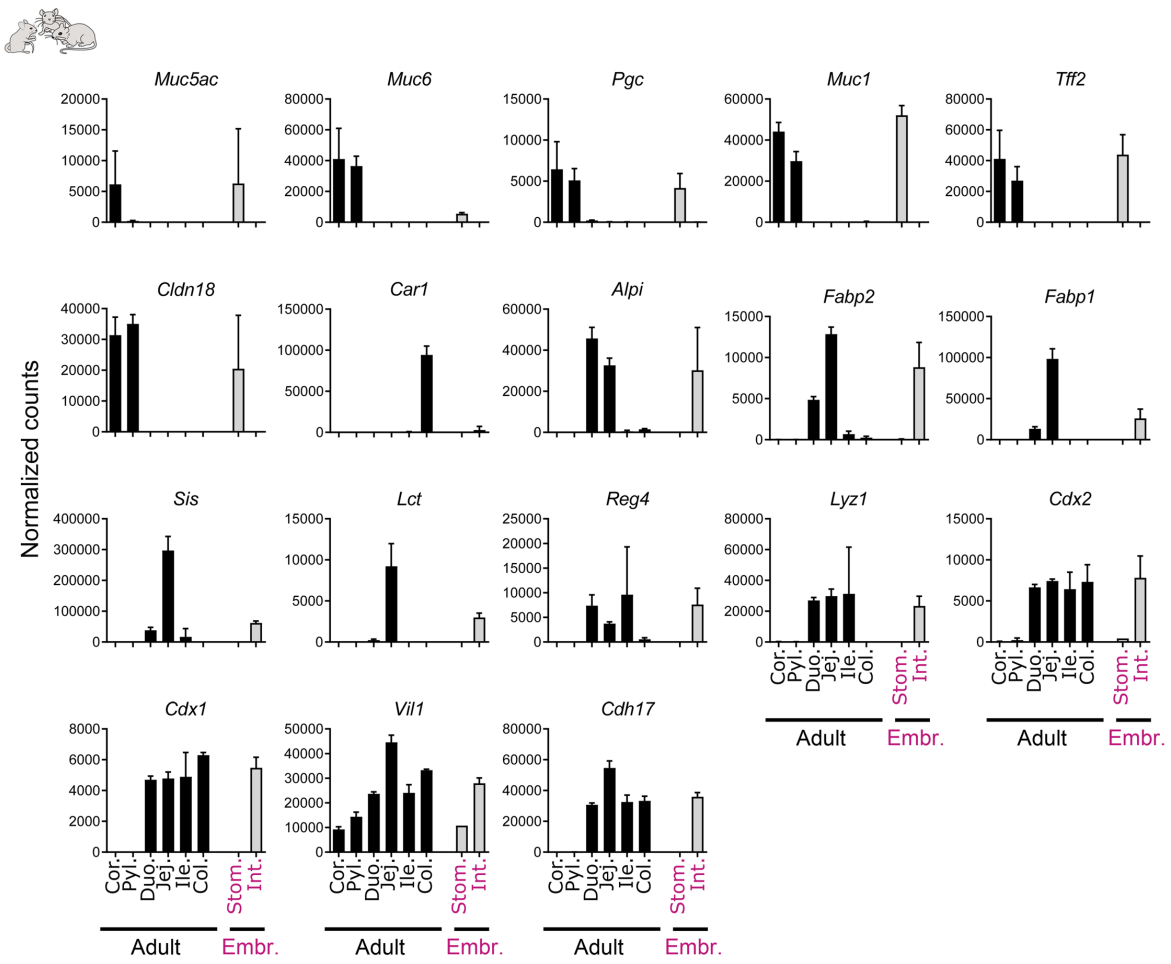


Figure S 3: Embryo-derived organoids express tissue identity markers in a segment-specific manner. Graphs depict the normalized gene counts scaled in the heatmap of **Figure 31** for the genes related to gastric (for embryo-derived organoids, Stom.) and intestinal (for embryo-derived proximal intestinal organoids, Int.) tissue-specific markers in the same order. Bars represent mean ± SD, n = 3 biological samples for each group, except embryonic stomach which is n=2.

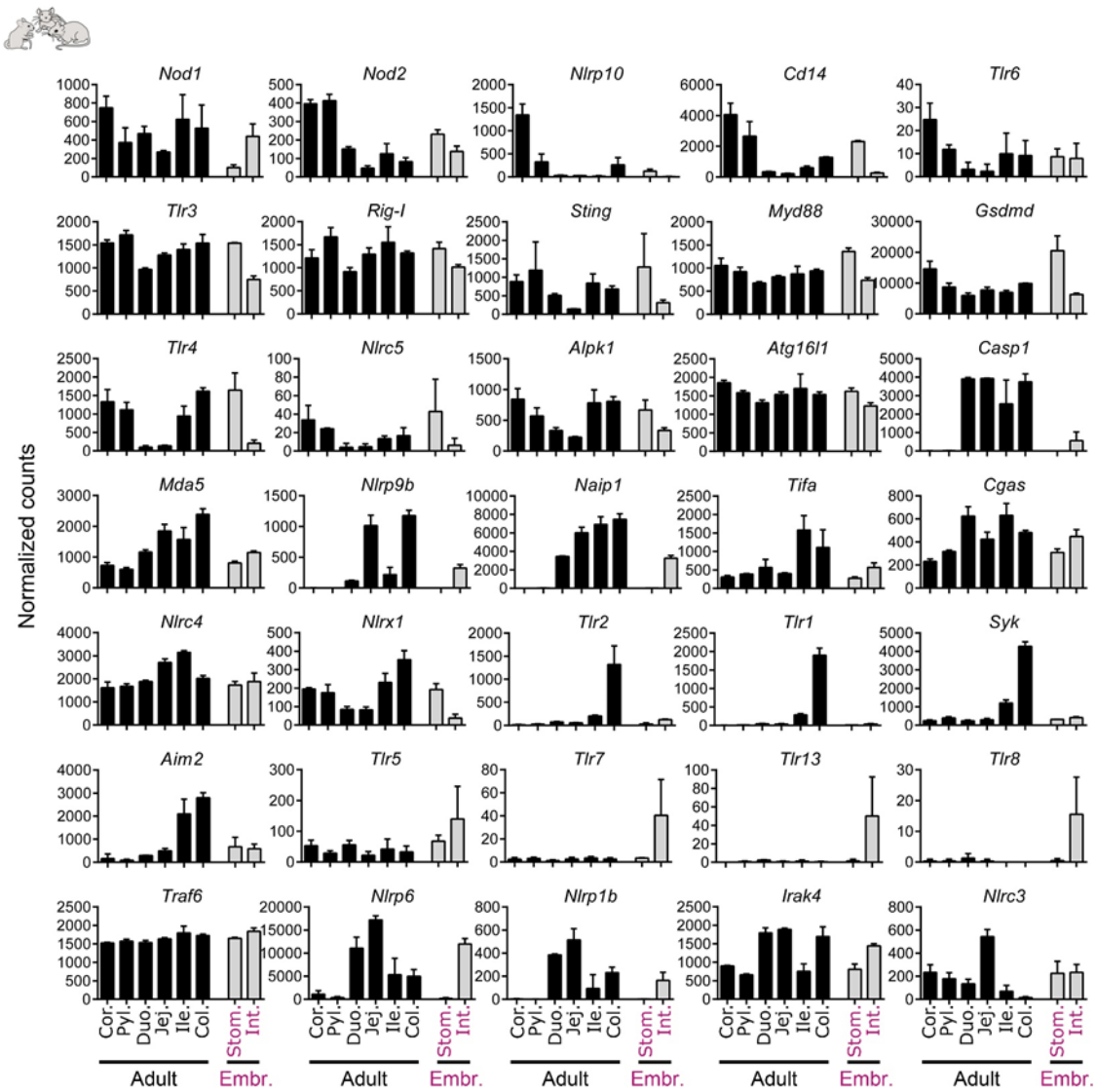


Figure S 4: Transcriptome profiling of the adult and embryonic mouse derived gastric and proximal intestinal organoids revealed patterned expression of genes related to PRR signaling pathways. Graphs depict the normalized gene counts scaled in the heatmaps of **Figure 33** for the genes related to pattern recognition receptor signaling pathways in the same order. Stom.: embryo-derived gastric organoids, Int: embryo-derived proximal small intestinal organoids. Bars represent mean ± SD, n = 3 biological samples for each group, except embryonic stomach which is n=2.

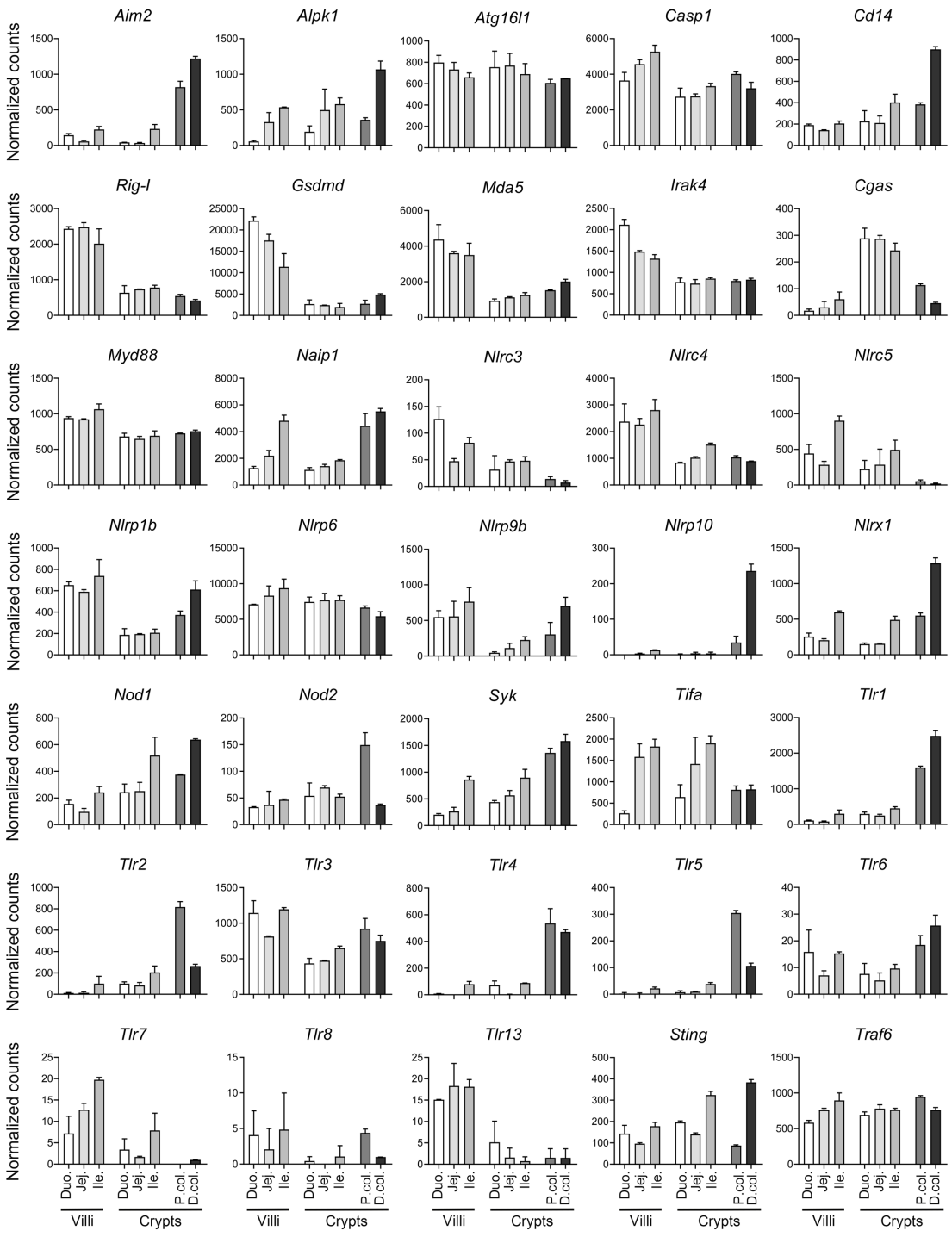


Figure S 5: PRRs are expressed in different patterns through murine crypt-villus axis and between crypts of small intestine and colon. All experiments were performed by Kozuka *et al*. [188]. Murine small intestinal (duodenum, jejunum and ileum) villi and crypts and colonic (proximal colon, distal colon) crypts were isolated, total RNA was extracted and sequenced. The data were generated from the RNA-seq counts under the accession number GSE104803 in GEO database. The normalized gene counts for the genes related the innate immunity (for the genes depicted in **Figure 19**). Bars represent mean ± SD, n = 2 biological samples for each group.

A

Supplemented basal medium (SBM)		
Advanced DMEM/F12		
HEPES (10 mM)		
GlutaMAX (1x)		
Penicillin/Streptomycin (1x)		
N2 (1x)		
B27 (1x)		
N-acetylcysteine (1 mM)		
Undifferentiated 2D	**Small intestinal 2D**	**Colonic 2D**
SBM	√	√
WNT (W)	250 ng/ml	100 ng/ml
R-spondin1 (R)	500 ng/ml	500 ng/ml
Noggin (N)	100 ng/ml	100 ng/ml
EGF (E)	50 ng/ml	50 ng/ml
RhoK inh. (Y)	20 µM	20 µM
Thiazovivin (T)	-	-
BMP-4 (B)	-	-
Differentiated 2D	**Small intestinal 2D**	**Colonic 2D**
SBM	√	√
WNT (W)	-	-
R-spondin1 (R)	-	-
Noggin (N)	-	100 ng/ml
EGF (E)	50 ng/ml	50 ng/ml
RhoK inh. (Y)	-	-
Thiazovivin (T)	2.5 µM	-
BMP-4 (B)	300 ng/ml	-

B

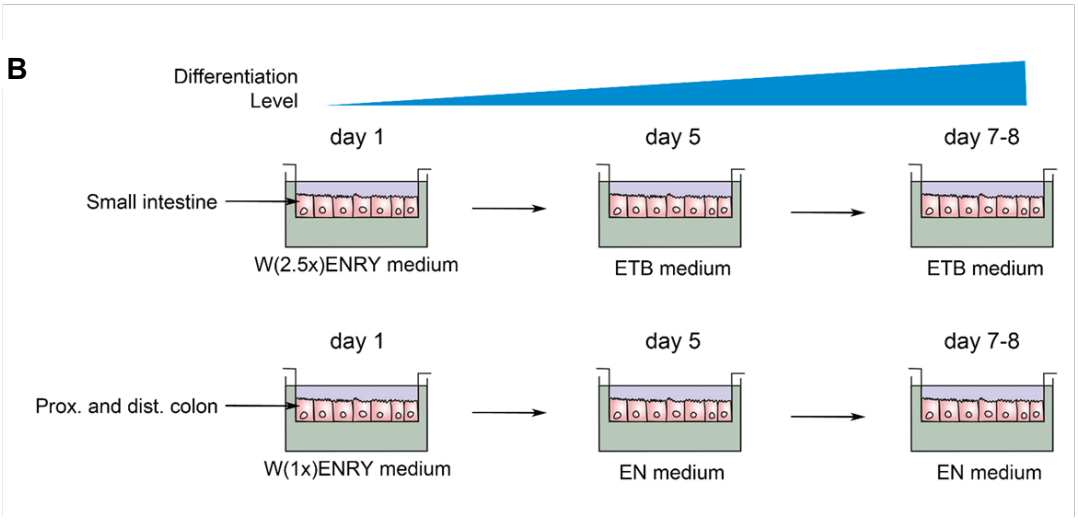


Figure S 6: Murine small intestinal and colonic cells were seeded on transwells as 2D monolayers and differentiated upon adjusting the medium conditions. All experiments were performed by Kozuka *et al*. [188]. The small intestinal and colonic cells were seeded on transwells and either differentiated or kept undifferentiated by different medium conditions. **(A)** Table depicting the ingredients of media for undifferentiated/differentiated cultures of small intestine and colon. **(B)** Type of medium added on cells and durations of culturing. At day 5, medium conditions were changed to drive the differentiation of cells and all samples were collected at day 7-8 for the RNA to be extracted and sequenced. SBM: Supplemented basal medium, W: WNT (1x for colon, 2.5x for small intestine), E: EGF, N: Noggin, R: R-spondin1, Y: Rho-kinase inhibitor, T: thiazovivin, B: BMP-4.

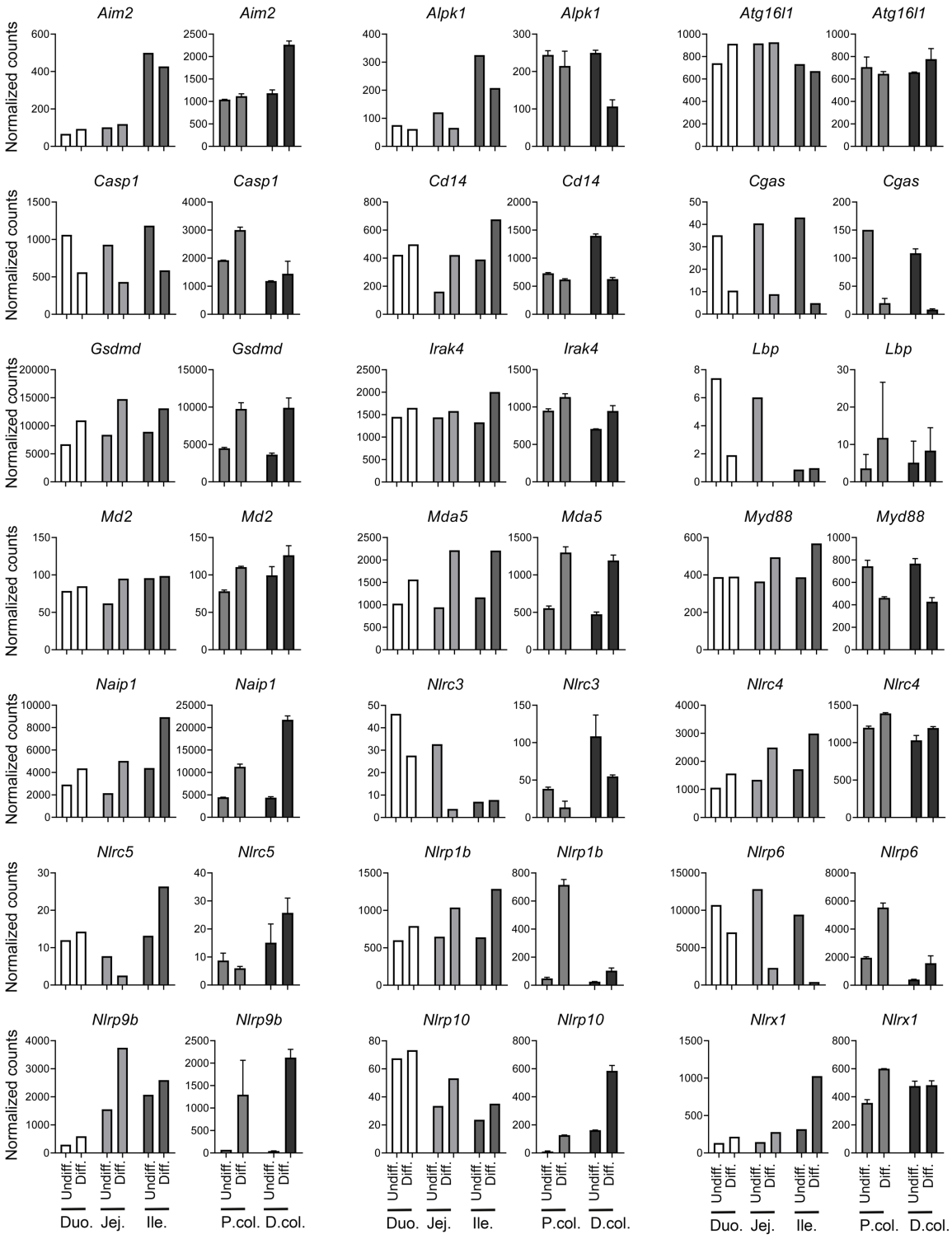


Figure S 7: Expression levels of selected PRRs according to the differentiation state of the small intestinal and colonic 2D monolayers. All experiments were performed by Kozuka *et al*. [188]. Small intestinal and colonic cells were seeded on transwells and either differentiated or not differentiated by using different medium conditions. Detailed culture conditions and durations can be found in **supplementary Figure S 6**. The data were generated from RNA-seq counts under the accession number GSE104803 in GEO database. Normalized gene counts for the genes related the innate immunity (same genes depicted in **Figure 19**). Bars represent mean ± SD, n = 2 biological samples for each group except small intestinal samples which are n=1.

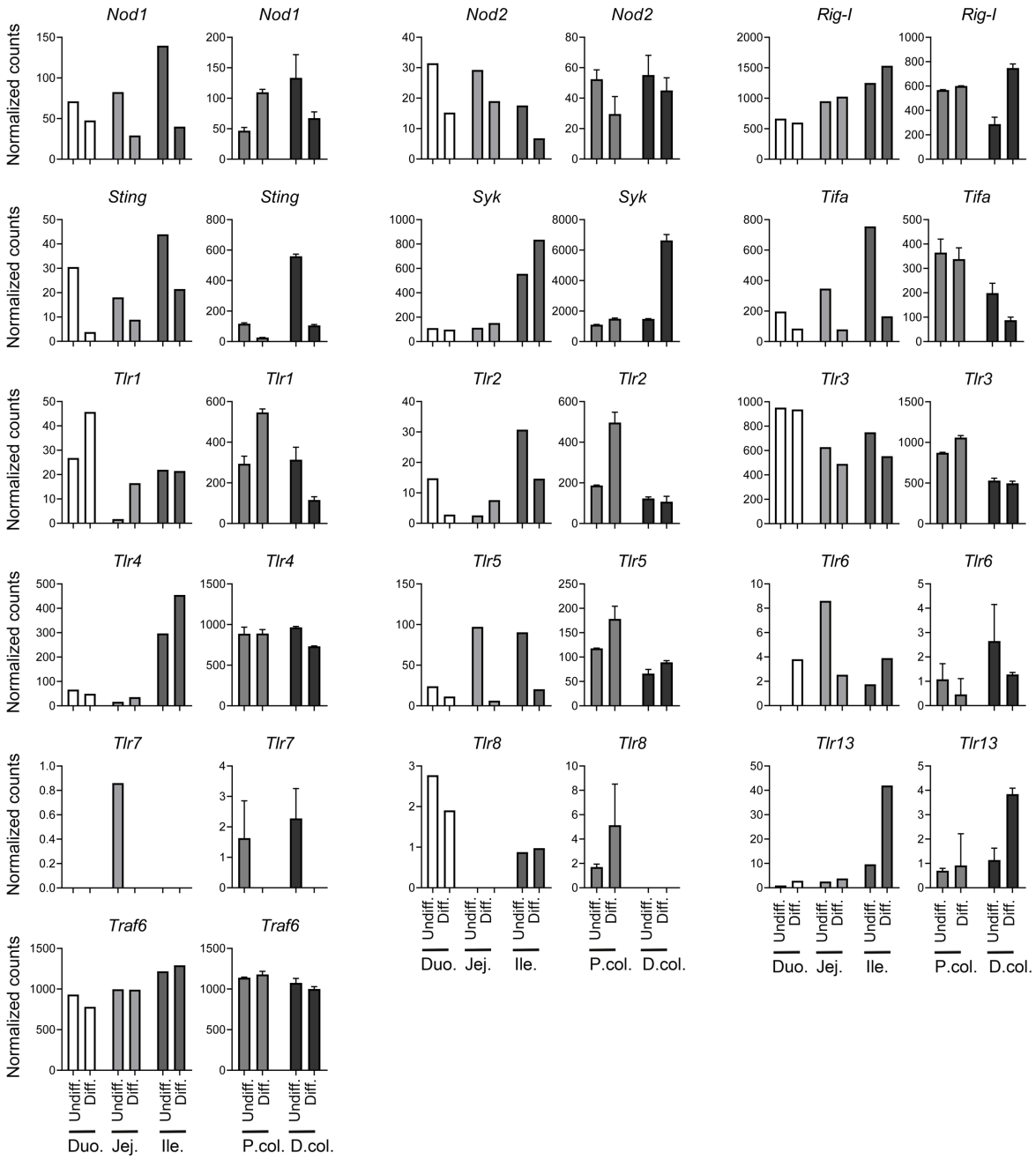


Figure S 7: continuation

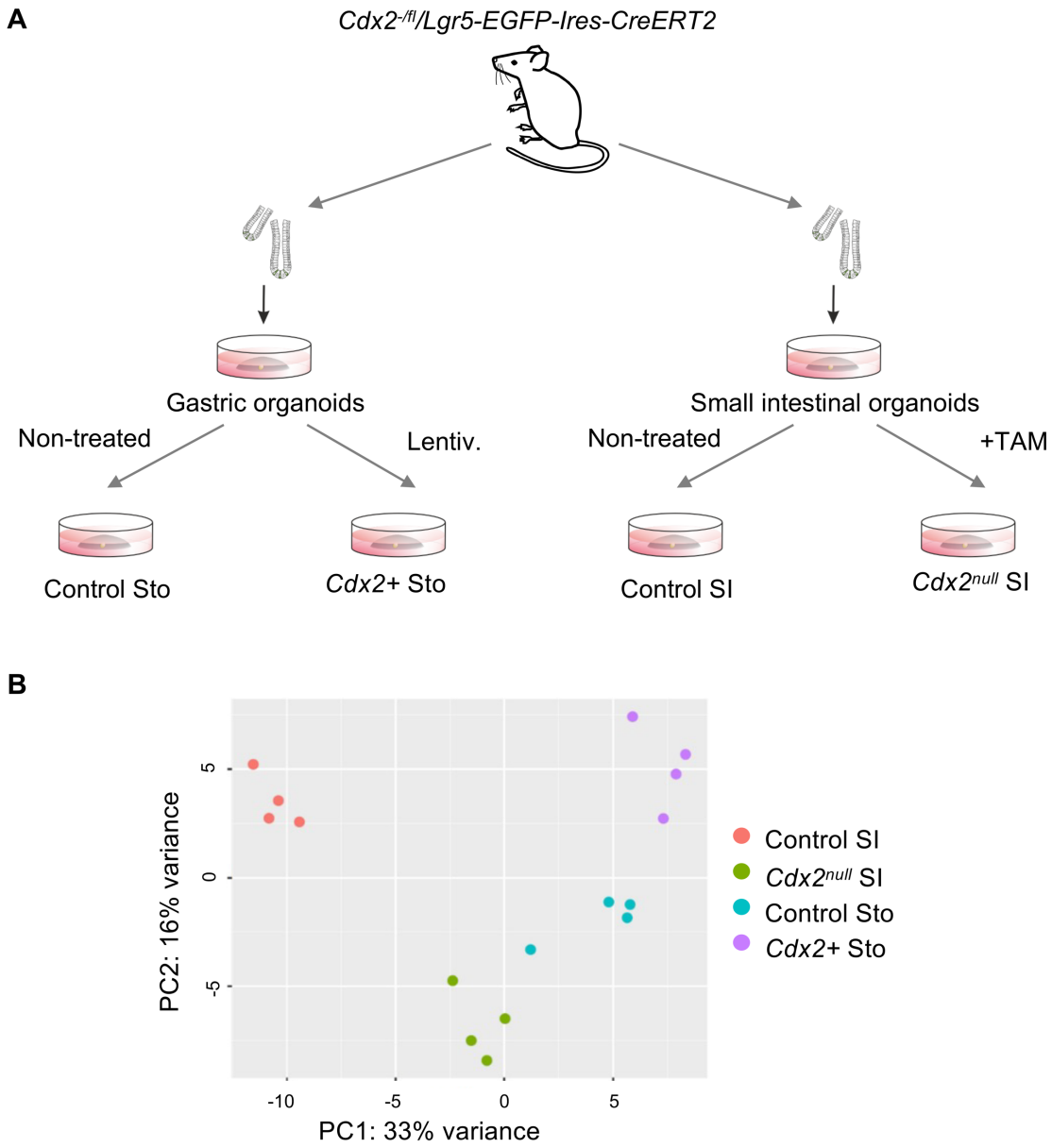


Figure S 8: RNA-seq data from murine small intestinal organoids with genetically ablated *Cdx2* and gastric organoids with overexpressed *Cdx2* were analyzed. (A) All experiments were performed by Simmini *et al*. [73]. Briefly, small intestinal and gastric organoids were generated from isolated crypts and glands of *Cdx2-/fl/Lgr5-EGFP-Ires-CreERT2* mice which express conditional *Cdx2* gene and *Lgr5-EGFP* stem cells. To obtain $Cdx2^{null}$ intestinal stem cells carrying the Lgr5-EGFP marker ($Cdx2^{null}$ SI), 5-6 days old small intestinal organoids were incubated with 1 µM of 4-hydroxytamoxifen (TAM) in intestinal culture medium for 16 hours to activate the Cre-recombinase. TAM-untreated small intestinal (Control SI) and gastric (Control Sto) organoids served as controls. Organoids were dissociated and sorted for high EGFP signal. Subsequently, *Cdx2*+ gastric organoids (*Cdx2*+ Sto) were generated by infection of the wild-type gastric organoids with lentiviral stock expressing the *Cdx2* gene. After all the groups were cultured in gastric culture medium (ENRWfg), total RNA was extracted from all samples and sequenced. The data were generated from RNA-seq counts under the accession number GSE51751 in GEO database. **(B)** PCA plot of the RNA-seq data.

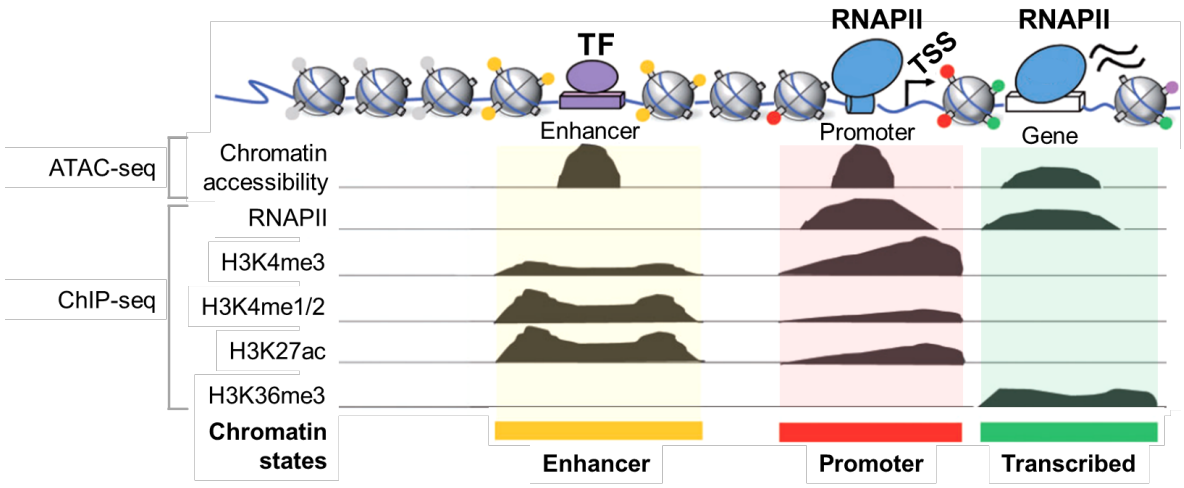


Figure S 9: Chromatin states defined by chromatin accessibility, different histone modifications and RNAPII binding. As the gene expression itself is oftentimes regulated by epigenetic mechanisms, the factors affecting the expression of a gene can be analyzed by investigating epigenetic factors regulating the gene expression. Transcription of a gene begins at the transcription start sites (TSS) which are a part of each gene's promoter. Promoters in turn are short sequences that are able to specify the precise location and direction of the transcription. Additionally, enhancers (>2 Kb upstream or >1 Kb downstream of the TSSs) control the access to the DNA region and much of the regulatory code driving gene expression is found in these distal elements. The transcriptional machinery including RNA polymerase II (RNAPII) and transcription factors (TF) are assembled at the promoter region of a gene. The epigenetic regulations include both large changes in the chromatin structure or progressive transition in the chromatin from a more open to a more packed chromatin state, or the other way round. Chromatin is structurally supported by DNA-binding proteins called histones. The open and closed state of the chromatin is largely determined by modifications on amino acid residues of the histones, which can alter their interaction with the DNA. Particularly, acetylation (ac) and methylation (me) of the histones are critical for epigenetic gene regulation. Histone modifications are strongly related to gene activation or repression and are therefore used as epigenetic indicators of the chromatin state. For example, trimethylation of the fourth lysine residue on histone 3 (H3K4me3), acetylation of H3K27 (H3K27ac) and high occupancy of RNAPII mark the TSS and promoter region of the actively transcribed genes (red). The active gene body itself carry trimethylated H3K36 (H3K36me3) as well as RNAPII (green). Active enhancers on the other hand are characterized by the high occupancy of both mono- and dimethylated H3K4 (H3K4me1, H3K4me2) and H3K27ac (yellow) [225–227]. Chromatin state and TF-DNA interactions can be analyzed by assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) and chromatin immunoprecipitation followed by sequencing of associated DNA fragments (ChIP-seq). Of the two, ATAC-seq identifies the accessible regions of the chromatin [184], whereas ChIP-seq helps define the histone modifications or binding sites of the corresponding transcription factors, to reveal the mechanisms involved in gene regulation. Following the sequencing, genome browsers such as the Integrative Genome Browser (IGV) can be used to visualize the ATAC- or ChIP-seq data. In ATAC-seq tracks, the sequencing reads enriched for accessible chromatin are generally distinct with sharp peaks along the genome. In ChIP-seq tracks, the peaks represent the DNA regions where the histone modification or transcription factor of interest is enriched. The sequencing regions enriched for transcription factors are sharp, similar to the ATAC-seq peaks. On the other hand, the distribution of reads for the histone modifications or RNAPII can be more continuous, resulting in broader peaks (adapted with permission from reference [228]; copyright 2018 Oxford University Press).

Table S 1: Expression of several pattern recognition receptors is affected by the expression of *Cdx2*. Experiments were performed by Simmini *et al.* [73] as explained in **supplementary Figure S 8**. The data were curated from normalized RNA-seq counts provided to the GEO with accession number GSE51751. Table reveals the differentially expressed genes (DEGs) found in different sections of the Venn diagram found in **Figure 42**. Comparing 4 Control Sto and 4 *Cdx2*+ Sto organoids (Stomach), 4 Control SI to 4 *Cdx2*^{null} SI organoids (Small intestine), DEGs in both groups (Intersection) ($p<0.05$ and log2Fold Change > ±1 (2-fold change)).

Stomach			Intersection	Small Intestine			
2310002J15Rik	*Gabrp*	*Serpinb11*	*Aim2*	*0610009B14Rik*	*Cdh17*	*Myo7b*	*Slc7a11*
5830428M24Rik	*Gatm*	*Slc26a2*	*Bhlha15*	*1700011H14Rik*	*Cdhr5*	*Myoz3*	*Slc9a5*
A230050P20Rik	*Gbp9*	*Slc35g1*	*C4b*	*1700016K19Rik*	*Cds1*	*Naip1*	*Slc9b2*
AA467197	*Gm11545*	*Socs1*	*Ces1d*	*2010001M06Rik*	*Cdx1*	*Naip6*	*Slco3a1*
Acox2	*Grin1*	*Sorcs3*	*Cpq*	*2200002J24Rik*	*Cdx2*	*Nbl1*	*Smarcd3*
Adrbk2	*Gsta3*	*Spns3*	*Cyp2s1*	*2210039B01Rik*	*Celf2*	*Nr2e3*	*Smoc2*
Ak4	*Habp2*	*Spp1*	*Dhh*	*2610528A11Rik*	*Cers4*	*Nrarp*	*Snx22*
Akr1e1	*Heatr5a*	*Sprr2a1*	*Eid2*	*3930402G23Rik*	*Ces2a*	*Nrp2*	*Sox12*
Aldh3a1	*Hkdc1*	*St8sia6*	*Fam189b*	*A4gnt*	*Ces2e*	*Nt5dc2*	*Spink4*
Alox12e	*Hmox1*	*Tceal8*	*H2-Q10*	*A630076J17Rik*	*Cfi*	*Oat*	*St6galnac4*
Atp2c2	*Hspb8*	*Tff1*	*Kcne3*	*Aadac*	*Chgb*	*Osmr*	*Stk17b*
Atp6v0e2	*Htra1*	*Thnsl1*	*Mgat4c*	*Abcg5*	*Chst4*	*Ovol1*	*Stxbp1*
B930041F14Rik	*Igsf23*	*Thrb*	*Muc20*	*Abi3*	*Clca4*	*Pamr1*	*Sulf1*
Bmp2	*Ildr1*	*Tmem14a*	*Nlrp6*	*Acta1*	*Cldn15*	*Pcsk7*	*Sult1a1*
Bnip3	*Inha*	*Tnfaip2*	*Otop3*	*Adam22*	*Clec2h*	*Pcsk9*	*Sult1d1*
Bpifb1	*Inhbb*	*Trp53i11*	*Rasa4*	*Adh6a*	*Clrn3*	*Pde2a*	*Syt7*
Bsg	*Iyd*	*Tsku*	*Rhoj*	*Adora1*	*Cmtm3*	*Pde5a*	*Tac1*
Btn1a1	*Kbtbd11*	*Tspan33*	*Rnf186*	*Agmat*	*Col16a1*	*Pde9a*	*Tesc*
Camk2b	*Khdc1a*	*Tyro3*		*Agpat4*	*Col9a3*	*Pea15a*	*Tff3*
Car9	*Kremen2*	*Ugt2b34*		*Akap5*	*Cox7a1*	*Pgc*	*Tgfb3*
Cbr2	*Lcn2*	*Wnt7b*		*Akap7*	*Cpe*	*Phgr1*	*Tm4sf20*
Cetn4	*Lgals1*	*Zfp287*		*Aldh1b1*	*Cpt1c*	*Pkdcc*	*Tm4sf5*
Cfh	*Lgals3*	*Zfp488*		*Aldob*	*Creb3l4*	*Pkp1*	*Tmem237*
Chad	*Ltf*	*Zik1*		*Alpi*	*Cth*	*Plgrkt*	*Tmem45a*
Chic1	*Ly6d*	*Zzef1*		*Amica1*	*Ctse*	*Pm20d1*	*Tmem64*
Chst8	*Mcmdc2*			*Amn*	*Ctss*	*Pnliprp2*	*Trim31*
Cish	*Mdfic*			*Ang4*	*Cyp2c55*	*Podxl*	*Trit1*
Clip4	*Mgat3*			*Ankrd37*	*Cyp2c65*	*Porcn*	*Trpv4*
Clps	*Mkx*			*Anxa13*	*Cytip*	*Pparg*	*Tspan9*
Cstf2t	*Mmp11*			*Apobec2*	*Dach1*	*Ppp1r14d*	*Tstd1*
Cxcl5	*Ndn*			*Apoc2*	*Dact3*	*Prap1*	*Uevld*
Cym	*Ndrg1*			*Arhgap22*	*Dbn1*	*Prkg2*	*Ush1c*
Cyp2b10	*Nenf*			*Atoh8*	*Ddc*	*Procr*	*Vdr*
Cyp2c44	*Nupr1l*			*Atp6v0a4*	*Defa17*	*Ptgs1*	*Vipr1*
Cyp4b1	*Olfr856-ps1*			*Avil*	*Defa24*	*Qsox2*	*Wdr72*
Dmrta2	*P4ha1*			*Axl*	*Depdc7*	*Rab11fip5*	*Wdr92*
Dnase1l3	*P4ha2*			*B4galnt2*	*Egfl7*	*Rabl2*	*Wfdc18*
Duoxa2	*Parm1*			*B4galnt4*	*Enho*	*Rdx*	*Wnt10a*
Dynap	*Pdp2*			*B4galt4*	*Entpd8*	*Reg4*	*Xlr3a*
E330020D12Rik	*Pdpn*			*B4galt5*	*Epas1*	*Rfxank*	*Zg16*
ERCC-00053	*Pfkfb3*			*Baiap2l2*	*Epdr1*	*Rgcc*	
Efemp1	*Pfn2*			*Blm*	*Ephx4*	*Rgs17*	
Efemp2	*Pgm2l1*			*Bmp8b*	*Etv4*	*Rnf214*	
Egln3	*Pik3ip1*			*C1ra*	*Evx1*	*S100g*	
Epha4	*Pla2g10*			*C2cd4b*	*F5*	*Selm*	
Ero1l	*Pla2g4c*			*C3*	*F830016B08Rik*	*Sema4g*	
Eva1b	*Plekhf1*			*Cald1*	*Fabp2*	*Serpine2*	
Evpl	*Prnp*			*Calml4*	*Fam101b*	*Slc16a10*	
Exoc3l4	*Ptges*			*Car8*	*Fbln2*	*Slc16a2*	
Fam124a	*Rab15*			*Casp1*	*Fbp1*	*Slc17a4*	
Fam162a	*Rasgef1a*			*Cbfa2t3*	*Fcgbp*	*Slc1a2*	
Fam46c	*Ripk3*			*Ccdc80*	*Frat2*	*Slc1a4*	
Fam71e1	*Rps4y2*			*Ccl25*	*Fscn1*	*Slc25a45*	
Fbxo44	*S100a4*			*Ccl6*	*Gdf15*	*Slc30a10*	
Fggy	*Selenbp1*			*Cd177*	*Ggt6*	*Slc39a14*	
Fhit				*Cd38*	*Ggta1*	*Slc5a1*	

10 Tables

Supplementary tables

11 Figure legends

Supplementary figures

12 Curriculum Vitae

13 Publications

Kayisoglu, O., Weiss, F., Niklas, C., Pierotti, I., Pompaiah, M., Wallaschek, N., Germer, C. T., Wiegering, A., & Bartfeld, S. (2021). Location-specific cell identity rather than exposure to GI microbiota defines many innate immune signaling cascades in the gut epithelium. Gut, 70(4), 687–697. https://doi.org/10.1136/gutjnl-2019-319919

Kayisoglu, Ö., Schlegel, N., & Bartfeld, S. (2021). Gastrointestinal epithelial innate immunity—regionalization and organoids as new model. Journal of Molecular Medicine, 99(4), 517–530. https://doi.org/10.1007/s00109-021-02043-9

Kayisoglu, Ö., Schlegel, N., & Bartfeld, S. (2020). 3.8 Die zelluläre Grenzschicht im Magen-Darm-Trakt und ihre Funktion in der Immunabwehr: Organoide als Modell des gastrointestinalen Epithels. In S. Bartfeld, H. Schickl, C. Alev M.D. Ph.D., B.-K. Koo Ph.D., A. Pichl, A. Osterheider, & L. Marx-Stölting (Eds.), Organoide: Ihre Bedeutung für Forschung, Medizin und Gesellschaft (1st ed., pp. 138–148). Nomos Verlagsgesellschaft mbH & Co. KG. https://doi.org/10.5771/9783748908326-138

Wallaschek, N., Reuter, S., Silkenat, S., Wolf, K., Niklas, C., Kayisoglu, Ö., Aguilar, C., Wiegering, A., Germer, C. T., Kircher, S., Rosenwald, A., Shannon-Lowe, C., & Bartfeld, S. (2021). Ephrin receptor A2, the epithelial receptor for Epstein-Barr virus entry, is not available for efficient infection in human gastric organoids. PLoS Pathogens, 17(2), 1–18. https://doi.org/10.1371/JOURNAL.PPAT.1009210

14 Acknowledgements

First of all, I would like to thank Dr. Sina Bartfeld for providing me the opportunity to conduct my PhD project in her group and work on this fascinating topic. It is amazing to realize how the project evolved and how valuable it became upon tackling each difficulty we encountered together. I will never forget how patient you were with me, that you never gave up on me and tried to see the best in me at my worst times. Thank you so much for everything.

To my co-supervisors Prof. Dr. Thomas Rudel and Dr. Bon-Kyoung Koo for kindly agreeing to evaluate this project and all their advice during the annual meetings.

To Nina, Sylvie, Carmen, Caro, Mastura and Mindaugas for their invaluable support and immense help. Also, for their patience with me during my turtle days and understanding. I could not do this without each of their support.

To Franziska, Jaap, Alisha, Dominik and Melissa who all helped me and supported this project during their time in our lab. Especially Franni, without whom this thesis would not exist in this form. Also, for always being there.

To GRK 2157 for providing funding and everyone involved for the nice discussions during the retreats we had. I would like to thank especially Dr. Emmanuel Saliba and Ehsan for all their help with the bioinformatics while I was struggling.

To Sarah Duin for being my strongest support and never letting me drown in the depths of darkness. For all her help with the thesis and also for being in my life.

To my beloved husband Ömer Kaya for… far too many things to recount here. You had an honorary master's degree in biology, now will have a PhD as well. I could not manage without your presence and also our apple tree.

Finally, to my parents and brother, again first of all, for their patience with me and for listening to me. Thank you so much for being there whenever I need it the most.

Affidavit

I hereby confirm that my thesis entitled “Analysis of gastrointestinal epithelial innate immune barrier using human and murine organoids as a model” is the result of my own work. I did not receive any help or support from commercial consultants. All sources and/or materials applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Place, Date Signature

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation "Analyse der Gastrointestinalen angeborenen Immunbarriere durch Humane und Murine Organoide als Modell" eigenständig, d.h. insbesondere selbständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Ort, Datum Unterschrift