From Department of Medicine, Solna
Karolinska Institutet, Stockholm, Sweden

Genetic variation of tight junction structures in intestinal inflammation

Elisabeth Norén

Stockholm 2016

Published by Karolinska Institutet
Printed by E-Print AB 2016

ISBN 978-91-7676-465-7

Karolinska Institutet

Department of Medicine, Solna

Genetic variation of tight junction structures in intestinal inflammation

THESIS FOR DOCTORAL DEGREE (Ph. D.)

Public defense on Friday the 16th of December 2016, 09.00 at Karolinska University Hospital, Solna, S2:01, Rehabsalen.

By

Elisabeth Norén

Principal Supervisor:
Sven Almer, professor.
Karolinska Institutet
Department of Medicine, Solna

Co-supervisor:
Jan Söderman, associate professor.
Region Jönköping County
Division of Medical Diagnostics

Faculty of opponent:
Alkwin Wanders, associate professor.
Umeå University
Department of Medical Biosciences

Examination board:
Jonas Mattsson, professor.
Karolinska Institutet
Department of Oncology-Pathology

Marie Carlson, professor.
Uppsala University
Department of Medical Sciences

Åsa Torinsson Naluai, associate professor.
The Sahlgrenska Academy
University of Gothenburg
Institute of Biomedicine

Stockholm 2016

”Jag har inga speciella talanger. Jag är bara passionerat nyfiken.”

Albert Einstein

Abstract

The epithelial barrier facing the external environment in the gastrointestinal (GI) tract is comprised of several components, including the tight junction (TJ) structures observed for the first time in 1963. TJ structures, which are multiprotein complexes composed of transmembrane proteins and a diverse spectrum of intracellular components, create a primary barrier to diffusion of ions, solutes, and water and they concomitantly prevent permeation of pro-inflammatory factors, such as pathogens, toxins, and antigens. Recent studies suggested that disturbance of epithelial integrity is associated with intestinal inflammatory conditions, such as inflammatory bowel disease (IBD), graft-versus-host disease (GVHD), celiac disease, but also diabetes.

The general aim of this thesis was to identify novel genetic variants related to the development of intestinal inflammation with a specific focus on the TJ structures, yielding implications for epithelial integrity and paracellular permeability.

Using a case-control study approach (Swedish cases and controls) in paper I, potential associations were investigated between IBD and three selected genetic markers in the genetic region of *CLDN1, CLDN2,* and *CLDN4* (one marker per gene). The strongest association was observed between Crohn's disease (CD) and the single nucleotide polymorphism (SNP) marker in the genetic region of *CLDN2*. The same SNP markers were further investigated using a family-based approach in non-Swedish families, but none of the identified associations from the Swedish case-control approach were confirmed. *MORC4* which is located in the same genetic region as *CLDN2* was also included in the investigation. A significant association was observed between a nonsynonymous SNP in *MORC4* and CD in the Swedish case-control cohort.

Similarities between IBD and GVHD include intestinal barrier defects and genetic contributions. GVHD is considered to be multifactorial, where the human leucocyte antigen (HLA) acts as a cornerstone; however, non-HLA genes have been identified in association with the outcome after stem cell transplantation (SCT). By using a case-control approach the relationship between non-HLA polymorphisms and emergence of GVHD as well as overall mortality after SCT was analyzed in paper II. The markers of *MORC4, CD14, TLR4,* and *NOD2* were found to associate with the outcome (overall mortality) after SCT. The SNP marker of *CD14* was the only analyzed marker that associated with acute GVHD.

In paper III, the associations between IBD and several TJ genes that encode proteins reported to interact with each other were analyzed in a Swedish population. The strongest associations were observed between IBD and SNP markers in the membrane-associated guanylate kinase inverted genes *MAGI2* and *MAGI3*. The *MAGI3* SNP was also associated with ileal *MAGI3* expression level in the non-inflamed non-IBD subgroup. Furthermore, no overlap between the expression levels of *PTEN* in inflamed colonic mucosa from patients with CD and those in the non-inflamed mucosa was detected, suggesting that *PTEN* is an inflammatory marker in CD.

In paper IV, the genetic associations between microscopic colitis (MC) and TJ genes were analyzed in a Swedish population. The strongest association was identified between a SNP marker in *PTEN* and MC and also the sub-phenotype collagenous colitis (CC). Furthermore, significant associations were observed between genetic variations of *MAGI1* and MC and also between a SNP marker in *F11R* and CC. Moreover, decreased expression levels of *PTEN* and *MAGI1* were primarily associated to CC and the MC subtype lymphocytic colitis (LC), respectively, in comparison with controls.

In conclusion, genes encoding proteins involved in the regulation of the intestinal epithelial integrity, including those contributing to TJ structures, may predispose individuals to intestinal inflammation, such as IBD and MC. Furthermore, *MORC4* may be a predisposing factor for CD and one-year mortality after SCT for hematological malignancies.

Sammanfattning på svenska

Barriären mot lumen i mag-tarmkanalen innefattar många komponenter, bland andra täta fogar (*tight junctions,* TJ). Dessa är multiproteinkomplex uppbyggda av transmembranproteiner och ett brett spektrum av intracellulära proteiner. Dessa strukturer reglerar passage av vatten och joner, men även genomsläppligheten av proinflammatoriska faktorer, såsom patogener, toxiner och antigener. Flertalet studier har identifierat en störd paracellular permeabilitet vid inflammation i tarmen, exempelvis vid inflammatorisk tarmsjukdom (IBD), transplantat-mot-värdsjukdom (GVHD), celiaki, men även vid diabetes.

Det övergripande syftet med avhandlingen var att identifiera genetisk variation i relation till inflammation i tarmen, med fokus på täta fogar.

Genetisk variation av *CLDN1, CLDN2* och *CLDN4,* en markör per gen, studerades i relation till IBD i arbete I genom applicering av ett fall-kontroll upplägg (svenska fall och kontroller). Den starkaste associationen observerades mellan en enbaspolymorfi (*single nucleotide polymorphism*, SNP) i *CLDN2* och Crohns sjukdom (CD). Samma genetiska variationer studerades vidare i ett familjebaserat upplägg, bland icke-svenska familjer, där inga signifikanta associationer noterades. *MORC4*-genen är belägen i samma genetiska region som *CLDN2* och inkluderades därför i studien. En icke-synonym SNP i *MORC4* analyserades bland svenska fall och kontroller (fall-kontroll upplägg), varpå signifikant association identifierades mellan markören och CD.

GVHD anses vara multifaktoriell, där *human leucocyte antigen* (HLA) spelar en viktig roll. Tidigare studier har visat att icke-HLA relaterade gener kan påverka utfallet efter stamcellstransplantation. I ett fall-kontroll upplägg, undersöktes genetisk variation (icke-HLA relaterade) i relation till utfallet efter stamcellstransplantation (arbete II).

Genetisk variation inom *MORC4, CD14, TLR4* och *NOD2* uppvisade association till generell mortalitet, och enbaspolymorfin inom *CD14* var dessutom associerad till uppkomst av akut GVHD.

I arbete III analyserades en svensk population av fall och kontroller avseende genetisk association mellan IBD och flertalet TJ-relaterade gener kodandes för interagerande proteiner. Den starkaste associationen noterades mellan IBD och genetisk variation i *MAGI2* och *MAGI3*. *MAGI3*-markören påverkade även genuttrycksnivån av *MAGI3* i ileum bland kontrollerna. Utöver detta noterades ett nedreglerat uttryck av *PTEN* i inflammerad kolonmukosa hos CD patienter, utan överlappning med uttrycksnivån i icke-inflammerad mukosa, bland CD-patienterna. Detta tyder på att *PTEN* kan utgöra en inflammationsmarkör för CD.

I det fjärde arbetet studerades genetisk association mellan gener som kodar för interagerande TJ-proteiner och mikroskopisk kolit (MC). Den starkaste associationen noterades mellan en SNP inom *PTEN* och MC samt undergruppen kollagen kolit (CC), men även association mellan MC och *MAGI1* samt mellan CC och *F11R* identifierades. Vidare noterades att en sänkt genuttrycksnivå av *MAGI1* och *PTEN* var associerade med undergrupperna lymfocytär kolit respektive CC, jämfört med kontroller.

Avhandlingen visar att genetisk variation av gener som kodar för TJ-relaterade proteiner kan predisponera för kronisk inflammation i tarmen (IBD och MC). Avhandlingen visar ytterligare att *MORC4* möjligen predisponerar för CD. Genetisk variation inom *MORC4* kan även påverka generell dödlighet efter allogen stamcellstransplantation.

List of publications

This thesis is based on the following papers, which will be referred to in the text by the corresponding Roman numerals:

I. *Söderman J, ***Norén E**, Christiansson M, Bragde B, Thiébaut R, Hugot J-P, Tysk C, O'Morain CA, Gassull M, Finkel Y, Colombel J-F, Lémann M, Almer S. Analysis of single nucleotide polymorphisms in the region of *CLDN2-MORC4* in relation to inflammatory bowel disease. *World Journal of Gastroenterology* 2013; 19: 4935-4943.

*Söderman J and **Norén E** contributed equally to this work.

II. **Norén E**, Verma D, Söderkvist P, Weisselberg T, Söderman J, Lotfi K, Almer S. Single nucleotide polymorphisms in *MORC4*, *CD14*, and *TLR4* are related to the outcome of allogeneic stem cell transplantation. *Annals of Transplantation* 2016; 21: 56-67.

III. **Norén E**, Almer S, Söderman J. Genetic variation and expression levels of tight junction genes identifies association between *MAGI3* and inflammatory bowel disease. *Manuscript submitted.*

IV. **Norén E**, Mellander MR, Almer S, Söderman J. Variations of the tight junction genes *F11R, MAGI1, MAGI2, MAGI3, PARD3, PTEN,* and *TJP1* in relation to microscopic colitis. *Manuscript*.

List of abbreviations

AJ	Adherens junctions
APC	Antigen presenting cell
CC	Collagenous colitis
CD	Crohn's disease
cDNA	Complementary DNA
Ct	Threshold cycle
CI	Confidence interval
DC	Dendritic cells
dNTP	Deoxynucleoside triphosphate
ddNTP	Dideoxynucleoside triphosphate
dHPLC	Denaturing high-performance liquid chromatography
ECL	Extracellular loop
Fc	Fold change
GI	Gastrointestinal
GVHD	Graft-versus-host disease
GWAS	Genome-wide association study
HD	Human defensin
HLA	Human leucocyte antigen
IBD	Inflammatory bowel disease
IEC	Intestinal epithelial cell
IFN	Interferon
IL	Interleukin
JAM	Junctional adhesion molecule
LC	Lymphocytic colitis
LD	Linkage disequilibrium
MC	Microscopic colitis
M cell	Microfold cell
mRNA	Messenger ribonucleic acid
miRNA	Micro ribonucleic acid

MLC	Myosin light chain
NFκB	Nuclear factor-κB
OR	Odds ratio
PDZ	PSD-95/DlgA/ZO-1
PIP2	Phosphatidylinositol (4,5)-bisphosphate
PIP3	Phosphatidylinositol (3,4,5)-trisphosphate
PMN	Polymorphonuclear leukocytes
RQ	Relative quantification
PRR	Pattern recognition receptor
RFLP	Restriction fragment length polymorphism
RT-qPCR	Reverse transcription quantitative PCR
SCT	Stem cell transplantation
SNP	Single nucleotide polymorphism
TER	Transepithelial electrical resistance
T_H cell	T helper cell
TJ	Tight junction
TLR	Toll-like receptor
TNF	Tumor necrosis factor
T_{reg} cell	Regulatory T cell
UC	Ulcerative colitis
ZO	Zonula occludens

Table of contents

Introduction

The human intestine

The gastrointestinal (GI) tract provides the body's largest barrier to the external environment. In humans the small and large intestines differ prominently in size, as the small intestine of an adult is approximately 6-7 m, whereas the colon is wider in diameter and considerable shorter (1.5 m) (Crawford, 1999; Mowat and Agace, 2014). The surface of the intestine is equivalent to half a badminton court (approximately 32 m^2) (Helander and Fandriks, 2014). Since the intestine every day is exposed to 1-2 kg of food, the intestine processes approximately 45 tons of food during a lifetime. The passage time is approximately 5 h in the small intestine and 20 h in the colon. The epithelium in the intestine is completely renewed every 3 to 4 days.

The small intestine is divided into three main parts: duodenum, jejunum, and ileum (Figure 1) (Mowat and Agace, 2014). The luminal surface of the GI tract plays several functions, such as absorption of essential nutrients and prevention of entry of harmful contents (Salim and Soderholm, 2011). The mucosal surface of the small intestine is characterized by finger-like projections, termed *villi*, which extend to the lumen and increase the area of the active epithelium. The villi are, however, absent in the cecum and the colon (Figure 2). The colon begins right after the small intestine at the cecum, followed by the ascending colon, the transverse colon, the descending colon, and lastly the rectum that terminates at the anus (Mowat and Agace, 2014).

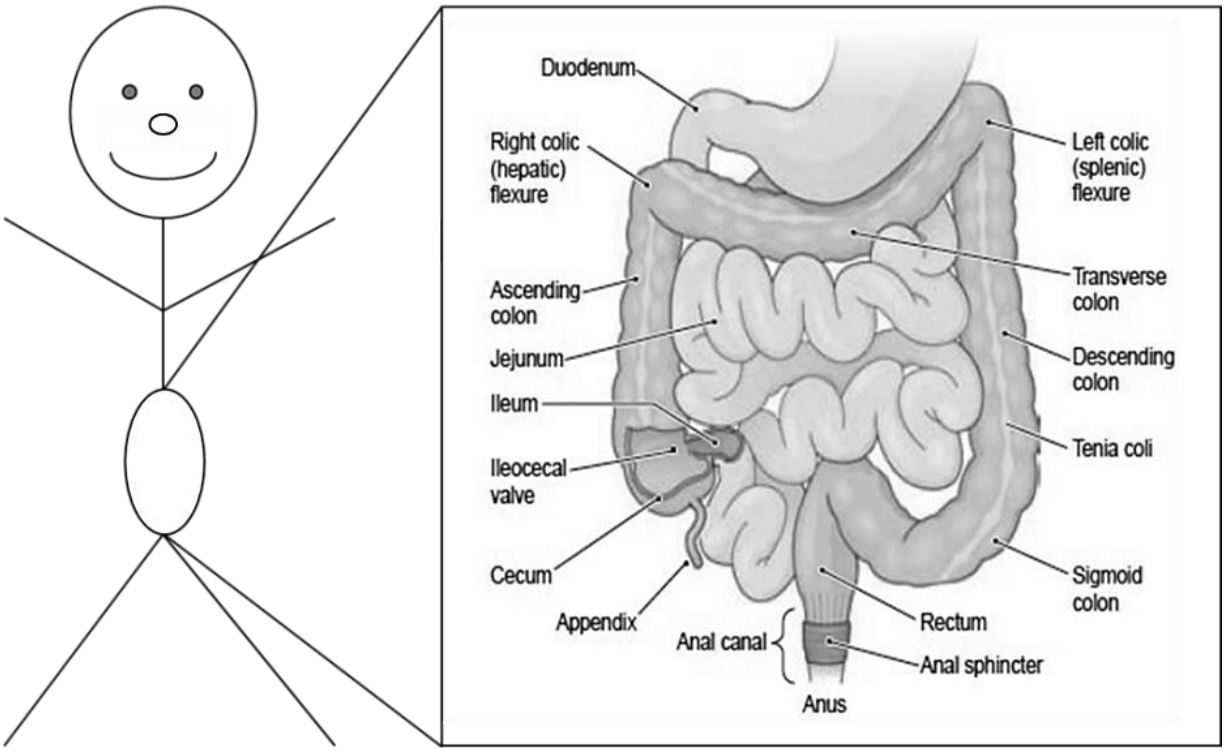


Figure 1. The human intestine. The framed figure is reprinted with permission (encyclopedia.lubopitko-bg.com).

Mucosal barrier components in the intestine

Constitution of a distinct internal environment is essential for multicellular organisms to maintain life. All surfaces (*e.g.* GI tract, skin, eyes, and respiratory tract) are therefore covered by epithelial cells in order to form a boundary between the external environment and the internal environment (Sawada, 2013). This barrier function is maintained by a complex multilayer system that comprises an outer physical barrier as well as an inner functional immunological barrier.

Outer mucosal components

The mucus layer (protective gel-like layer) is produced by *goblet cells* and overlays the epithelium, protecting the body against luminal bacteria (Figure 2). The goblet cells comprise 25% of the epithelial cells in the distal colon compared to only 10% or less of the epithelial cells in the small intestine. The mucus layer consists of two layers in the colon, the outer and inner layers. Bacteria can reside in the outer mucus layer, but these organisms rarely penetrate into the inner mucus layer. The purity of the inner mucus layer is dependent on the presence of antimicrobial peptides, such as defensins

produced by *Paneth cells*; however, increased penetration of bacteria into the epithelial surface of the colon may result in increased susceptibility to colitis and colorectal cancer (Mowat and Agace, 2014; Salim and Soderholm, 2011).

Intestinal epithelial cells

Pluripotent stem cells located in the crypt of the epithelium give rise to several mature intestinal epithelial cells (IEC), such as enterocytes, goblet cells, Paneth cells, microfold cells (M cells), and enteroendocrine cells (Figure 2) (Menard *et al.*, 2010; Wehkamp and Stange, 2010).

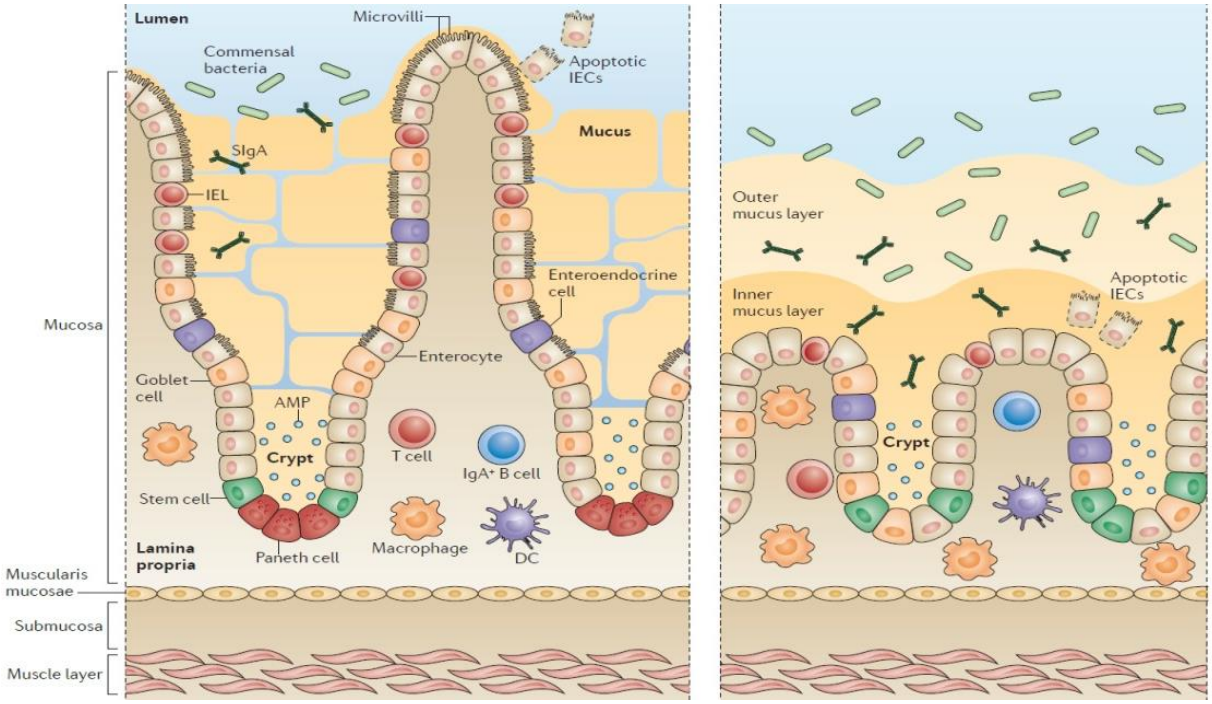


Figure 2. Anatomy of the intestinal mucosa (left, small intestine; right, colon). Figure reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Immunology, copyright 2014 (Mowat and Agace, 2014).

The main task of the intestinal *enterocytes* is regulation of nutrient absorption into the circulation. The *goblet cells*, which are spread out on the continuous epithelial layer, produce the mucus layer that protects against infiltration of bacteria into the epithelial cells. The *Paneth cells* are particularly concentrated in the distal part of the ileum. After differentiation from stem cells, Paneth cells migrate to the very bottom of the crypts where they exert their function (Mowat and Agace, 2014; Salim and Soderholm, 2011). Dysregulated Paneth cells have been observed in Crohn's disease (CD), concomitant with decreased levels of ileal defensins (Wehkamp *et al.*, 2005).

Finally, the *M cells* are IECs that occur over organized gut-associated lymphoid tissue, including Peyer's patches in the small intestine. These cells are specialized for transport of antigens from the lumen into a region enriched in *dendritic cells* (DC) and *lymphocytes*. Here, the antigens can thereby be presented to the adaptive immune system. The size of the Peyer's patches increases from the jejunum to the ileum, and these patches are primary concentrated in the distal part of the small intestine. The frequency of colonic lymphoid follicles increases from the ascending colon and reaches its maximum in the sigmoid colon (Gullberg and Soderholm, 2006; Menard *et al.*, 2010; Mowat and Agace, 2014).

Pattern recognition receptors

The epithelial cells express a wide range of pattern recognition receptors (PRR), such as Toll-like receptors (TLR) and nucleotide-binding oligomerization domain 2 (NOD2). NOD2, an intracellular muramyl dipeptide receptor, is mainly expressed by epithelial cells in the ileum; the Paneth cells (Mowat and Agace, 2014; Salim and Soderholm, 2011). This is consistent with the identified association between genetic variation of *NOD2* and *ileal* CD, but not colonic type of CD (Cuthbert *et al.*, 2002; Wehkamp and Stange, 2010).

Furthermore, the pattern of TLR expression is heterogeneous along the intestine, as TLR2 is primarily expressed in the proximal part of the colon and decreases along the colon distally. TLR4 and CD14 are generally expressed at higher levels by epithelial cells in the colon than by those in the small intestine. Nevertheless, dysregulated intestinal TLR4 and CD14 expression may predispose activation of the mucosal innate immune system, and thus also acting as an important factor in the pathogenesis of inflammatory bowel disease (IBD) (Frolova *et al.*, 2008). Subsequently, nuclear factor-κB (NFκB) is activated in response to PRR binding, contributing to an increased level of pro-inflammatory cytokines (interleukin (IL)-1, IL-6, tumor necrosis factor (TNF), interferon (IFN) gamma), chemokines, and antibacterial peptides, *e.g.*, human defensins (HD)-5 and HD-6 secreted by Paneth cells (Mowat and Agace, 2014; Salim and Soderholm, 2011; Wehkamp and Stange, 2010).

Epithelial integrity

The epithelium is a selective barrier that limits the permeation of pro-inflammatory molecules from the lumen into the mucosal tissue, while simultaneously allowing absorption of water and nutrients (Groschwitz and Hogan, 2009; Salim and Soderholm, 2011). In addition to the paracellular passage of molecules, a transcellular route has been described (Salim and Soderholm, 2011). The epithelial cells are further responsible for the immune interface, *e.g.* PRR and HD (Menard *et al.*, 2010; Stappenbeck and McGovern, 2016; Wehkamp and Stange, 2010).

The barrier is rigorously formed by cell-cell junctions, including tight junctions (TJs), adherens junctions (AJs), and desmosomes (Figure 3) that consist of transmembrane interacting proteins to link the epithelial cells together. The paracellular permeability is primarily regulated by the TJ complexes, which form a continuous band that encircles the epithelial cells and forms intercellular interactions, which are termed "*kissing points*" (Groschwitz and Hogan, 2009; Suzuki, 2013).

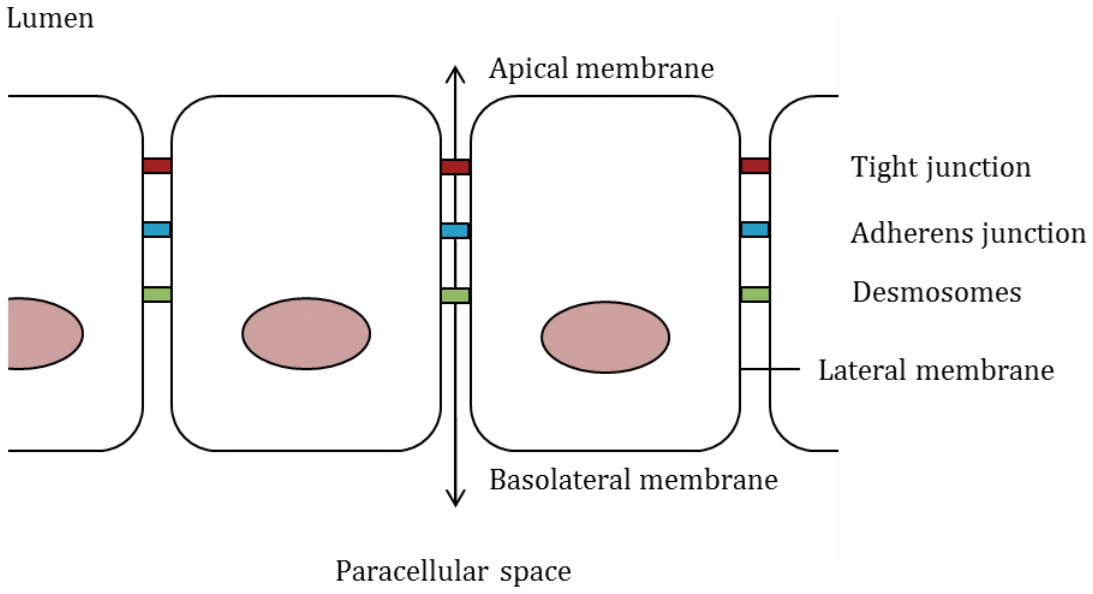


Figure 3. Illustration of IEC and TJs, AJs, and desmosomes in the intercellular space between the epithelial cells.

AJs link cells by bridging the cytoskeleton of adjacent cells via interactions between transmembrane proteins (cadherin), intracellular structures (catenin), and the cytoskeleton. Cadherin-catenin interactions do not only contribute to a strong adhesive effect between adjacent cells but also to cell polarity, regulation of proliferation, and migration (Groschwitz and Hogan, 2009).

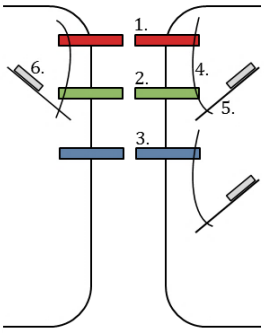
Tight junctions and their molecular composition

TJ structures, which were first described 1963 (Farquhar and Palade, 1963), are dynamic complexes composed of transmembrane proteins as well as a spectrum of intracellular proteins. These multiprotein complexes create a primary barrier to the diffusion of ions, solutes, and water due to the selective/semipermeable barrier and also prevent the permeation of pro-inflammatory agents, such as pathogens, toxins, and antigens (Groschwitz and Hogan, 2009; Salim and Soderholm, 2011). Moreover, the sealing function is tissue-specific as well as both size- and charge-selective (Van Itallie *et al.*, 2003). Studies have identified the presence of two separate TJ paracellular pathways; a CLDN-dependent size restrictive pore pathway responsible for flux of small and ionic molecules (<4 Å) and an occludin-dependent large-channel pathway responsible for flux of macromolecules (Al-Sadi *et al.*, 2011; Van Itallie *et al.*, 2008).

The transmembrane proteins CLDN, occludin, and junctional adhesion molecule (JAM) form the barrier via both homophilic and heterophilic interactions to yield the kissing points in the intercellular space (Figure 4). The intracellular domains of these transmembrane structures interact with intracellular proteins, such as PSD-95/DlgA/ZO-1 homology (PDZ) containing proteins such as zonula occludens (ZO) proteins (Luissint *et al.*, 2016), thus anchoring the transmembrane protein to the actomyosin ring (protein complex composed of actin and myosin molecules). Paracellular permeability is regulated by several physiological functions, such as myosin light chain (MLC) phosphorylation/dephosphorylation, as well as the intrinsic CLDN composition (Turner *et al.*, 1997). Furthermore, the amount of occludin contributes to the quantity of strands in the paracellular space (McCarthy *et al.*, 1996).

Contraction of the actomyosin ring is regulated by MLC activity. Indeed phosphorylation of MLC by MLC kinase induces contraction of the actomyosin ring, thereby opening the paracellular pathway and further increasing the permeability (Suzuki, 2013).

The cell-membrane of the IEC has two domains; an apical domain and a basolateral domain. Under physiological conditions, TJs act as a fence to prevent intermixing of the proteins from the apical part of the cell to the basolateral part, thereby maintaining the polarity of the cells (Sawada, 2013).



1. Claudin
2. Occludin
3. JAM
4. ZO-1
5. Actin
6. Myosin

Figure 4. Schematic illustration of the TJ structure.

The claudin family

The first members of the CLDN-family, CLDN1 and CLDN2, were discovered by Furuse *et al.* (1998), and today, at least 27 members in mice and humans have been identified with unique expression patterns in different tissues (Mineta *et al.*, 2011). The CLDNs are transmembrane structures with two extracellular loops (ECL; Figure 5), creating homophilic and heterophilic interactions with adjacent epithelial cells (Krug *et al.*, 2014).

The CLDNs may be divided into two groups: *barrier-forming* (tight epithelia) and *channel-forming* (leaky epithelia) CLDNs; in leaky epithelium the paracellular pathway is more ion-conducive than the transcellular pathway. The channel-forming CLDNs exhibit different types of selectivity, *i.e.* for cations, for anions, or for water (Krug *et al.*, 2014).

Moreover, overexpression of CLDN2 causes increased paracellular permeability due to an increase in the number of pores without effects on the charge selectivity (Furuse *et al.*, 2001; Van Itallie *et al.*, 2008), whereas other CLDNs, such as CLDN4, affect the charge

selectivity without affecting the number of pores (Inai *et al.*, 1999; Van Itallie *et al.*, 2003). A tightness function has been observed in various epithelia for several of the CLDNs (*e.g.* CLDN1 and CLDN4) when using knockout mouse models (Menard *et al.*, 2010).

The combination and mixing ratios of the CLDN structures incorporated in the same strand have been found to determine the tightness of the IECs (Luissint *et al.*, 2016). Moreover, the number of TJ strands correlates with an increased transepithelial electrical resistance (TER; measure of ionic flux) value and decreased permeability; however, the number of TJ strands does not seem to be simply determined by the total amounts of CLDNs expressed by the cells (Furuse *et al.*, 2001).

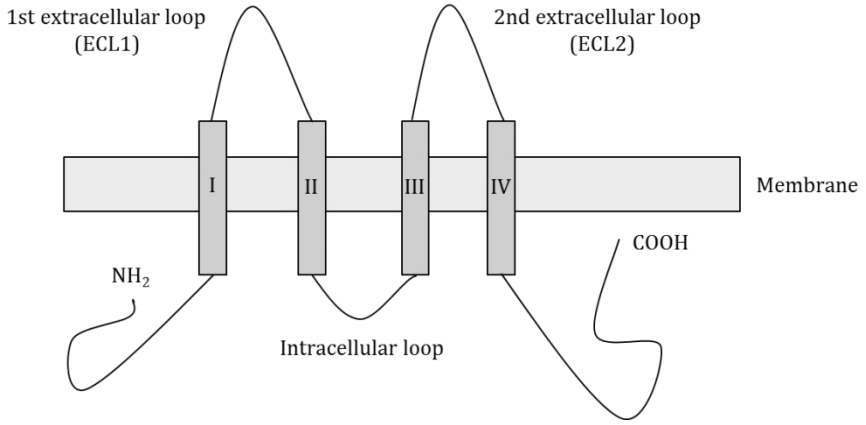


Figure 5. Schematic illustration of the CLDN protein. COOH, carboxyl group; NH_2, amino group.

Expression of CLDN proteins may be affected in intestinal inflammation, such as IBD (Krug *et al.*, 2014; Kucharzik *et al.*, 2001; Prasad *et al.*, 2005; Weber *et al.*, 2008; Zeissig *et al.*, 2007) and collagenous colitis (CC) (Burgel *et al.*, 2002). Increased levels of CLDN1 and CLDN2 proteins have been observed in colonic tissue from IBD patients with active disease, but not in those with inactive disease (Poritz *et al.*, 2011; Weber *et al.*, 2008). In contrast, CLDN4 expression does not correlate with disease activity (Prasad *et al.*, 2005; Weber *et al.*, 2008). If the CLDN expression is affected in intestinal graft-versus-host disease (GVHD) is not known.

Occludin

Occludin was the first transmembrane TJ protein discovered (Furuse *et al.*, 1993), but its contribution to TJ regulation remains incompletely understood (Luissint *et al.*, 2016). Similar to CLDN, occludin is a transmembrane protein with two extracellular domains and intracellular domains. The intracellular domains interact with ZO-1 and thus link occludin to the cytoskeleton (Fanning *et al.*, 1998; Itoh *et al.*, 1997). Induced expression of occludin has been shown to associate with an increased TER value and also with an increased number of TJ strands in the paracellular space (McCarthy *et al.*, 1996).

Deletion of occludin contributes to increased flux of macromolecules in the paracellular pathway and this influx may further induce an inflammatory response, without affecting the TER. However, the deletion of occludin only had a modest effect on the flux of smaller molecules (Al-Sadi *et al.*, 2011). Some studies have shown decreased intestinal occludin expression in several pathological conditions, including IBD (Poritz *et al.*, 2011; Yamamoto-Furusho *et al.*, 2012), CC (Burgel *et al.*, 2002), and GVHD (Noth *et al.*, 2011).

Junctional adhesion molecule

JAMs are included in a family composed of five members (JAM-A, JAM-B, JAM-C, JAM-4, and JAM-like) of which JAM-A was the first to be identified (Martin-Padura *et al.*, 1998; Vetrano *et al.*, 2008). These molecules are transmembrane proteins that span once across the cell membrane and contribute to epithelial integrity as gate-keeper via intercellular interactions (Vetrano and Danese, 2009). Knockout mice as well as JAM-A-deficient epithelial cells exhibit increased permeability, assessed by a decrease in the TER (Luissint *et al.*, 2014; Luissint *et al.*, 2016). Furthermore, JAM-A seems to play a role in the regulation of intestinal permeability to both smaller and larger molecules (Luissint *et al.*, 2014). Significantly increased levels of lymphoid aggregates, which consist of B and T lymphocytes as well as polymorphonuclear leukocytes (PMN), are observed in JAM-A$^{-/-}$ mice (Laukoetter *et al.*, 2007). Moreover, decreased colonic expression of JAM-A has been described in IBD (Laukoetter *et al.*, 2007; Vetrano *et al.*, 2008).

Intracellular molecules

Intracellular domains of the transmembrane proteins interact with intracellular scaffolding proteins, such as ZO, which in turn anchor the transmembrane proteins to the actomyosin ring. These interactions between transmembrane protein and actomyosin ring provide cytoskeletal regulation of epithelial integrity. Thus, contraction and tension of the actomyosin ring regulated by MLC activity cause increased paracellular permeability (Suzuki, 2013).

Intracellular TJ proteins are mainly divided into two groups: PDZ domain-containing proteins (interacting with proteins with a PDZ motif) and non-PDZ domain-containing proteins. These non-PDZ proteins are primarily involved in signaling regulation rather than direct contribution to the basic structural elements. A large group of intracellular TJ-related proteins do, however, contain one or multiple PDZ domains (Guillemot *et al.*, 2008).

The first identified intracellular TJ specific protein ZO-1 was described by Stevenson *et al.* (1986). ZO-1 is a scaffolding protein of the membrane-associated guanylate kinase family and contains three PDZ domains, a SH3 domain, and a guanylate kinase domain in the amino-terminal region. ZO-1 interacts with TJ-related transmembrane proteins, such as CLDNs and occludin. The carboxyl-terminal region of ZO-1 directly interacts with the actomyosin ring (Hu *et al.*, 2013). Furthermore, ZO-1 is at the center of a protein interaction network, and the interaction between the ZO-1 and CLDNs may play an important role in the assembly of the TJ strands (Umeda *et al.*, 2006).

Other examples of intracellular scaffolding proteins include PARD3 and the membrane-associated guanylate kinase inverted (MAGI) proteins. These scaffolding proteins interact with intracellular signaling proteins (*e.g.*, protein kinases and PTEN) and are thereby involved in the regulation of different pathways, including cell polarization and Akt signaling (Zihni *et al.*, 2016).

Disrupted barrier function and disease development

Studies in recent years suggest that disruption of essential elements of the intestinal barrier is associated with permeation of molecules, followed by disruption in the immune system, intestinal inflammation, and tissue damage. This perturbed barrier is intimately associated with several intestinal inflammatory conditions, such as IBD (Almer *et al.*, 1993; Bjarnason, 1994; Clayburgh *et al.*, 2004; Hollander *et al.*, 1986; Zundler and Neurath, 2015), microscopic colitis (MC) (Barmeyer *et al.*, 2012; Munch *et al.*, 2005), GVHD (Noth *et al.*, 2011), and celiac disease (van Elburg *et al.*, 1993), as well as with type 1 diabetes (Bosi *et al.*, 2006; Carratu *et al.*, 1999). Whether the increased permeability in IBD is a consequence of the inflammatory process or whether the inflammatory process is a consequence of the enhanced permeability remains unclear. *Which came first, the chicken or the egg?*

Inflammatory cytokines, such as IL-1β, IL-4, INF-γ, and TNF-α, have been shown to affect the expression of TJ-related proteins (*e.g.* CLDN, occludin, and JAM-A) and also the epithelial integrity (Al-Sadi and Ma, 2007; Cui *et al.*, 2010; Luissint *et al.*, 2014; Poritz *et al.*, 2011; Wisner *et al.*, 2008; Zeissig *et al.*, 2007). Studies have further identified novel roles of non-coding micro ribonucleic acid (miRNAs) in the regulation and maintenance of TJ structures, and some of these identified miRNAs directly regulate the expression level of TJ proteins, such as occludin (Liang and Weber, 2014).

Inflammatory bowel disease

IBD is the term used to describe some, but not all, types of chronical inflammation in the GI tract. Inflammation caused by known microorganisms is not included in this category nor is inflammation caused by toxic substances (Lindgren and Löfberg, 2011). The diagnosis of IBD is, however, based on a combined assessment of the symptoms of the patient, findings of endoscopy, and histopathology findings.

The two major forms of IBD have traditionally been classified as CD and ulcerative colitis (UC). Cleynen *et al.* (2016) has suggested that IBD is better described by three groups (ileal CD, colonic CD, and UC), instead of the traditional separation. Overall, IBD is characterized by chronic inflammation in the GI tract in genetically predisposed individuals upon exposure to certain environmental factors (Molodecky *et al.*, 2012). IBD is associated with abdominal pain and diarrhea, which span a wide spectrum of severity.

Clinical and pathological features of CD and UC may substantially overlap. The major hallmarks that distinguish CD from UC are the *skip-lesions*, which in CD occur anywhere in the GI tract, deep-penetrating ulcers, as well as fistulas. UC lesions are homogenous, and the inflammation area is mainly restricted to the mucosa. Phenotypic overlap and inflammation limited to the colon may make diagnosis of CD or UC difficult or impossible. These patients are classified as IBD-undetermined or indeterminate colitis and may in time develop the clinical presentation of either CD or UC (Henriksen *et al.*, 2006; Liu and Stappenbeck, 2016). Furthermore, a change in diagnosis over time (5 years) from CD to UC and vice versa has been reported to occur in a few cases (3%). A possible explanation for this change in diagnosis over time is the initial coexistence of both phenotypes (CD and UC). A small fraction of patients (6%) with an initial diagnosis of CD or UC are also found to not have IBD after 5 years, and this change may result due to difficulties in the discrimination between a patient with IBD in remission over a long time with an initial relapse and self-limiting colitis (Henriksen *et al.*, 2006).

Extra-intestinal manifestations of the disease involving joints, skin, eyes, liver, and kidney are often present. Furthermore, IBD patients often experience reduced quality of life and capacity for work (Cosnes *et al.*, 2011; Foersch *et al.*, 2013; Geremia *et al.*, 2014).

Epidemiology of inflammatory bowel disease

Extensive variation in the epidemiological data of IBD within as well as between geographic regions has been reported all around the world but estimates of prevalence reach the highest rates in Canada and Europe. Moreover, industrialization is believed to be associated with IBD, and in line with this notion, additional increases in disease prevalence have been observed as developing countries have become more industrialized. IBD occurs more frequently in urban regions than in rural areas. People raised in urban regions of industrialized nations are exposed to considerably different environmental factors than people raised outside these regions. Moreover, urbanization and industrialization of societies have been associated with modifications in microbial exposure, lifestyle behavior, medication, and sanitation, all of which have been suggested to be potential environmental risk factors for IBD. In addition, the increased incidence of IBD during recent years may also be a consequence of increased awareness of IBD by healthcare professionals as well as advancements in diagnostic methods. Increased access to medical services, such as colonoscopy, may also contribute (Molodecky *et al.*, 2012).

Myren *et al.* (1971) reported Norwegian epidemiological IBD data from the late 1960s and documented an incidence of 3 per 10^5 inhabitants and 1 per 10^5 for UC and CD, respectively. These data were in line with a Swedish study that reported UC incidence of 2 per 10^5 during the late 1950s and also an increased incidence rate to a maximum of 5 per 10^5 during the early 1970s (Nordenvall *et al.*, 1985). In the 1990's the Inflammatory Bowel South-Eastern Norway (IBSEN) group reported an incidence of 14 per 10^5 and 6 per 10^5 for UC and CD, respectively (Moum *et al.*, 1996a; Moum *et al.*, 1996b), indicating increased numbers of IBD patients. Moreover, today, both UC and CD are worldwide health problems with incidence rates of 24.3 per 10^5 and 10.6 per 10^5, respectively, in Northern Europe. Based on observed estimates, approximately 0.8% of the population in Europe has IBD, and furthermore, areas with traditionally low occurrences of IBD (Asia

and Africa) have witnessed increasing numbers of IBD patients in recent years (Ek *et al.*, 2014; Loddo and Romano, 2015; Molodecky *et al.*, 2012). The North-South gradient that was reported for IBD in the European population is becoming less prominent with time (Molodecky *et al.*, 2012; Shivananda *et al.*, 1996), whereas an East-West gradient is appearing (Figure 6) (Burisch *et al.*, 2014a; Cosnes *et al.*, 2011; Vegh *et al.*, 2014) with the highest incidence rate on the Faroe Islands (82 per 10^5) (Burisch *et al.*, 2014a).

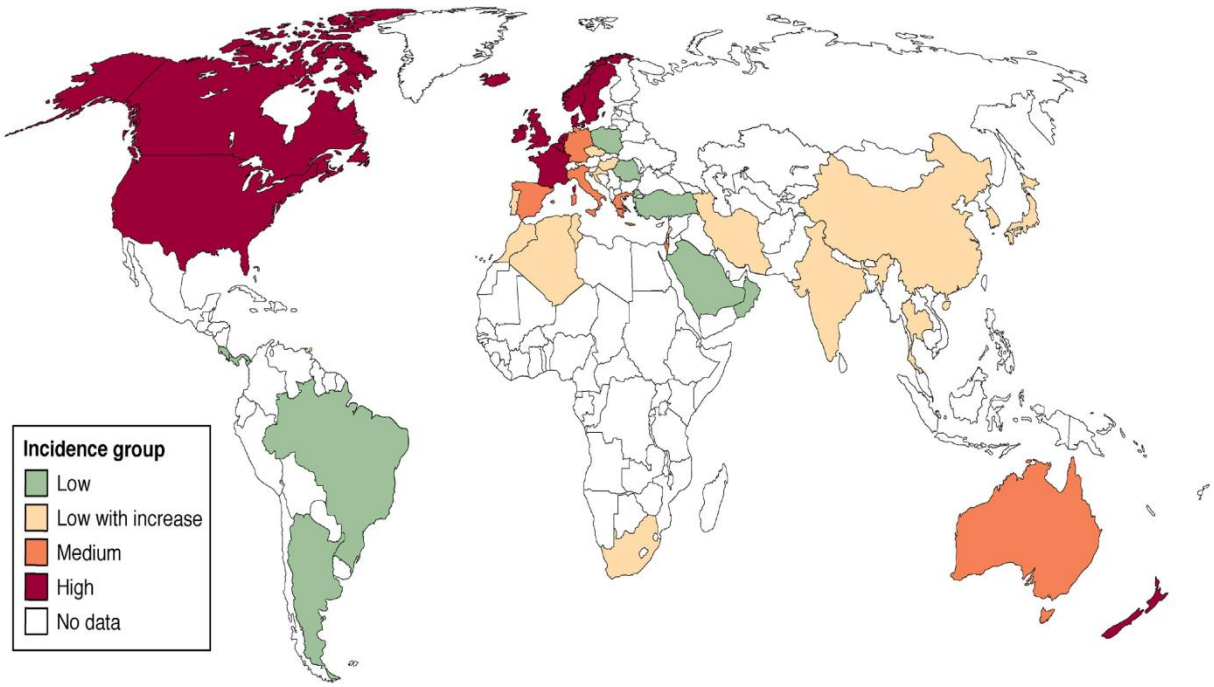


Figure 6. Global map depicting IBD. Red indicates annual incidence >10/10^5, orange indicates incidence 5-10/10^5, green indicates incidence <4/10^5, and yellow indicates countries with a low but increased incidence. Figure reprinted from Cosnes et al. (2011), Copyright 2011, with permission by Elsevier.

The onset of CD and UC peaks at a young age, between 15-25 years of age and 25-35 years of age, respectively (Cosnes *et al.*, 2011; Moum *et al.*, 1996a; Moum *et al.*, 1996b; Vatn and Sandvik, 2015). Approximately 20% of IBD patients are diagnosed in childhood or adolescence (Ek *et al.*, 2014; Loddo and Romano, 2015). UC tends to be overrepresented in men, whereas CD occurs more often in women (Cosnes *et al.*, 2011; Moum *et al.*, 1996a; Vatn and Sandvik, 2015). The higher incidence of CD among women, however, is not observed worldwide, as in Europe and North America the incidence of CD among men has increased to that among women (Cosnes *et al.*, 2011). Although advances in clinical care have been made for IBD, the disease constitutes a huge clinical

problem with approximately 2.2 million affected Europeans (Vatn and Sandvik, 2015) and 1.3 million affected Americans in the United States (Cosnes *et al.*, 2011).

Etiology of inflammatory bowel disease

Although the etiology of IBD has been investigated during the past decades, the pathogenesis of IBD is still not completely understood. The strongest risk factor identified to date is family history, and furthermore, the adaptive immune system has been suggested to play a major role in IBD pathogenesis. Recent research in genetics and immunology has documented the involvement of the innate immune system. The development of IBD may be a consequence of increased permeability, inadequately functioning innate immune system, and dysregulated adaptive immune system; however, which factors are responsible for the initiation of the disease remain unknown.

Environment factors in relation to IBD

Numerous environmental factors have been found to either protect against or act as a risk factor for CD and UC, and these risk factors differ among geographic regions, such as between Western and Eastern Europe (Burisch *et al.*, 2014b).

Several putative risk factors have been investigated for their relation to IBD. For example, smoking (Burisch *et al.*, 2014b; Mahid *et al.*, 2006) and appendectomy (Ananthakrishnan, 2015; Baron *et al.*, 2005; Burisch *et al.*, 2014b) have been shown to be associated with CD and protective against UC. Furthermore, breastfeeding may reduce the risk of IBD development (Frolkis *et al.*, 2013; Klement *et al.*, 2004), although this association was not found by Baron *et al.* (2005). Whether delivery by cesarean section is a risk factor for IBD (Bager *et al.*, 2012) or not (Bernstein *et al.*, 2016) has been unambiguously determined.

The hygiene hypothesis suggests that improved sanitation and reduced exposure to microorganisms during childhood contribute to an improper immunological response later in life. Living on a farm during childhood, living with multiple siblings, and consuming unpasteurized milk may be associated with reduced risk of IBD (Frolkis *et al.*, 2013). The gut microbiota is the largest reservoir of microbes in the body. Several

factors, such as IBD genetic risk variants in the host, diet, genetics, hygiene, geography, drugs, and smoking, may affect the stability and complexity of the gut microbiota (Imhann *et al.*, 2016; Kostic *et al.*, 2014), which, in turn, possibly predispose individuals to IBD (Peterson *et al.*, 2015).

Epithelial integrity

IBD is characterized by intestinal inflammation, increased paracellular and transcellular permeability, and ensuing infiltration of pathogens. An impaired epithelial barrier contributes to increased exposure of the mucosal immune system to luminal antigens, resulting in an inflammatory process. Loss of the intestinal barrier function may also enable passive passage of water and ions into the lumen, resulting in *leak flux-induced* diarrhea (Keita *et al.*, 2008; Krug *et al.*, 2014; Lee, 2015; Prasad *et al.*, 2005; Schmitz *et al.*, 1999; Zundler and Neurath, 2015). Increased intestinal permeability has been observed in IBD patients and also among first-degree relatives and spouses of IBD patients (Hilsden *et al.*, 1996; Hollander *et al.*, 1986; Landy *et al.*, 2016; Peeters *et al.*, 1997; Schmitz *et al.*, 1999; Secondulfo *et al.*, 2001; Soderholm *et al.*, 1999).

The intestinal permeability seems to be dependent on not only hereditary factors but also environmental factors (Peeters *et al.*, 1997; Soderholm *et al.*, 1999). Inflammatory cytokines associated with intestinal inflammation may alter the epithelial integrity due to the effects of these cytokines on TJ structures (Huang and Chen, 2016; Landy *et al.*, 2016).

Immune response

The immune system is crucial for our survival and protects us from different foreign invaders, such as bacteria, viruses, and parasites. In addition, the immune system develops tolerance to self-antigens. The immune response must be carefully regulated to prevent immunological *hyperactivity* as evident in IBD patients (Dave *et al.*, 2014; Foersch *et al.*, 2013; Geremia *et al.*, 2014; Loddo and Romano, 2015), where infiltrated mucosal immune cells and increased production of pro-inflammatory cytokines are characteristic (Zundler and Neurath, 2015). The immune system can be divided into two different components: the *innate immune system* and the *adaptive immune system* (Huang and Chen, 2016).

The non-specific innate immune system constitutes the first line of active defense against pathogens (Dave *et al.*, 2014; Foersch *et al.*, 2013), whereas cells of the adaptive immune system recognize conserved structural motifs (pathogen-associated molecular patterns) on the microorganism via PRR, such as TLRs and NOD2 (Corridoni *et al.*, 2014; Foersch *et al.*, 2013). Defensins, which are secreted by Paneth cells, are produced in response to recognition of bacterial components, and decreased expression of α-defensins has been identified in CD patients, especially in the case of *NOD2* variants (Wehkamp *et al.*, 2004). DCs, which are phagocytic antigen presenting cells (APC), further interact with B and T cells and play a key role in the crosstalk between the innate and the adaptive immune systems (Foersch *et al.*, 2013; Geremia *et al.*, 2014).

Upon activation, T helper 0 cells (T_H0 cells) differentiate into T_H1, T_H2, T_H17, or regulatory T cells (T_{reg} cells). The differentiation of T_H0 cells is driven by cytokines that are primarily released from cells of the innate immune system (Foersch *et al.*, 2013; Huang and Chen, 2016). The T_H1 (IL-1, IL-2, IL-6, and IL-8) and T_H2 (IL-4, IL-10, and IL-13) cells are in balance under normal conditions, but an imbalance of the systems determines the form of immunological disorder that will occur (Huang and Chen, 2016). Furthermore, increased T_H1 and T_H2 cell activity has been described in CD and UC, respectively (Zanello *et al.*, 2014; Zundler and Neurath, 2015).

T_H1 cells are essential for the elimination of intracellular pathogens, whereas T_H2 cells are involved in the protection against parasites and the mediating of allergic reactions. T_H17 cells contribute to the host defense against extracellular fungi and bacteria (Geremia *et al.*, 2014). T_{reg} cells inhibit T_H cells (*i.e.*, T_H1, T_H2, and T_H17) via secretion of several cytokines (IL-10 and transforming growth factor (TGF)-β). Reduced level of T_{reg} cells in colonic tissue is described to associate with IBD pathogenesis (Huang and Chen, 2016).

Genetic contribution to IBD

CD and UC are multifactorial diseases with significant genetic influence. Genetic studies have demonstrated that a large number of genes with a small effect are associated with IBD susceptibility (Ek *et al.*, 2014).

Since the 1960s, familial aspects of IBD have been of interest (Kirsner and Spencer, 1963), and further attempts to improve the understanding of genetic contribution to the IBD pathophysiology have greatly increased in recent years. Data from epidemiological studies indicate that the first-degree relatives of patients with CD or UC have a significantly increased lifetime risk of developing IBD compared to general population (Orholm *et al.*, 1991; Probert *et al.*, 1993). The highest risk of developing IBD has however been observed among the offspring of two affected parents, as approximately 30% of these offspring develop disease before 30 years of age (Halme *et al.*, 2006). The Swedish twin study performed by Tysk *et al*. (1988) showed a stronger genetic contribution to disease for the CD phenotype compared to the UC phenotype (Tysk *et al.*, 1988), and these findings have subsequently been confirmed by other studies (Halfvarson *et al.*, 2003; Orholm *et al.*, 2000). Moreover, Halfvarson *et al.* (2003) observed a higher degree of concordance among monozygotic twin pairs (UC 6-20% and CD 30-50%) than among dizygotic twin pairs (UC 4-5% and CD 4-10%).

Genetic variation of *NOD2* was the first polymorphisms identified to associate with CD (Hugot *et al.*, 2001; Ogura *et al.*, 2001). However, significant geographic variations regarding allele frequencies of *NOD2* have been identified, highlighting the existence of other contributing genetic loci in IBD (Cavanaugh, 2006). Presently, 206 susceptibility loci have been described in association with IBD (Ellinghaus *et al.*, 2016; Jostins *et al.*, 2012; Liu *et al.*, 2015). Only few of these associated signals correspond however to nonsynonymous coding single nucleotide polymorphisms (SNPs), with a function on the protein function (Cleynen and Vermeire, 2015). Yet, known loci explain a mere 10.9% and 7.7% of the heritability of CD and UC, respectively (Ellinghaus *et al.*, 2016).

Four years after the identification of the relationship between genetic variation of *NOD2* and CD, the first genome-wide association study (GWAS) using SNP was performed in an attempt to identify other IBD susceptibility genes (Yamazaki *et al.*, 2005), and this study uncovered significant associations between genetic variation of *TNFSF15* and CD.

Assays for markers in genomic regions that are associated with different immune-mediated overlapping phenotypes have been developed as an *Immunochip*, which encompass approximately 200.000 SNP markers at 186 loci. Jostins *et al.* (2012)

combined GWAS with *Immunochip* data and identified 163 loci that were significantly associated with IBD, with a substantial degree of overlap between CD and UC. Of these 163 loci, 110 were significantly associated with both CD and UC, whereas 30 loci were associated only with CD and 23 were associated only with UC.

The GWAS studies have revolutionized our understanding of genetic susceptibility in IBD. Several specific loci (*e.g,. CDH1, GNA12, HNF4A, LAMB1, MUC1, OSMR,* and *PTPN22*) detected in GWAS studies suggest an involvement of *the epithelial barrier* in the pathogenesis of IBD (Anderson *et al.*, 2011; Barrett *et al.*, 2008; Ellinghaus *et al.*, 2016; Franke *et al.*, 2010; Jostins *et al.*, 2012; Liu *et al.*, 2015; McGovern *et al.*, 2010). Also, a number of genes related to TJ assembly (*e.g. MAGI2, GNA12,* and *MYO9B*) have been identified in association studies using a candidate gene approach (Anderson *et al.*, 2011; Cooney *et al.*, 2009; Li *et al.*, 2016; McGovern *et al.*, 2009; Wapenaar *et al.*, 2008).

Microscopic colitis

Inflammatory conditions of the colon, such as lymphocytic colitis (LC) and CC, are subgroups of MC, which is characterized by watery, non-bloody diarrhea and less frequently abdominal pain and weight loss. Approximately 10% of patients with chronic diarrhea are diagnosed with MC (Munch and Langner, 2015). In several epidemiological investigations, female predominance was found for both LC and CC, although this predominance is more pronounced for CC (Bonderup *et al.*, 2015; Mellander *et al.*, 2016; Munch and Langner, 2015; Olesen *et al.*, 2004). MC was initially considered to be a rare disease, but an increase in annual incidence has been reported (from approximately 5 per 10^5 inhabitants in 2002 to 25 per 10^5 in 2011) perhaps due to increased endoscopic activity and awareness of the disease (Bonderup *et al.*, 2015).

The majority of patients with MC have macroscopic normal colonic mucosa, and therefore, colonoscopy with multiple biopsies is necessary to obtain the diagnosis. The histopathological hallmark for MC (for LC as well as CC) is increased infiltration of inflammatory cells in the colonic mucosa at both the epithelial surface and the lamina propria. An increased number of intraepithelial lymphocytes (>20 per 100 epithelial cells) is observed in LC, whereas the key histological feature in CC is a thickened collagen band (>10 μm) in addition to the infiltration of IELs. Damage to the surface epithelium is usually more pronounced in CC compared to LC. Whether LC and CC are two histological manifestations of the same disease entity remains controversial (Langner *et al.*, 2015; Magro *et al.*, 2013; Mellander *et al.*, 2016; Munch and Langner, 2015; Pisani *et al.*, 2016; Storr, 2013).

Pathogenesis of microscopic colitis

The exact mechanism underlying the diarrhea in MC is still unclear, although MC is an inflammatory condition of the intestine and is considered a new member of the IBD group (Storr, 2013).

The pathogenesis of MC is multifactorial, and both an abnormal immune response and increased intestinal paracellular and transcellular permeability have been described (Barmeyer *et al.*, 2012; Munch *et al.*, 2005). An established risk factor for MC is concomitant autoimmune disease, such as celiac disease (Stewart *et al.*, 2011). Moreover, female gender, increased age, smoking, and consumption of some drugs, such as non-steroidal anti-inflammatory drugs, have also been observed to associate with MC development (Langner *et al.*, 2015; Storr, 2013).

Jarnerot *et al.* (2001) reported familial clustering of MC, and some studies identified an association between genetic variation (*e.g.*, *HLA*, *IL-6,* and *MMP*) and MC (Koskela *et al.*, 2011; Koskela *et al.*, 2008; Madisch *et al.*, 2011; Westerlind *et al.*, 2015). Although MC is characterized by affected paracellular permeability (Munch *et al.*, 2005), no studies have reported an association between TJ-related genes and MC.

Graft-versus-host disease

Approximately 25.000 hematopoietic stem cell transplantations (SCT) are performed worldwide each year, and SCT has proven to be a successful cure for patients suffering from hematological diseases, immune deficiencies, and metabolic disorders (Hymes *et al.*, 2012). Despite advances in human leucocyte antigen (HLA)-typing, acute GVHD (aGVHD) remains a significant cause of morbidity and mortality related to the procedure. In addition to the resultant morbidity and mortality, GVHD results in impaired quality of life; however, the severity of GVHD differs between patients. Furthermore, GVHD is directly correlated to the degree of mismatch between HLA proteins. However, 40-60% of the recipients of matched HLA grafts develop aGVHD that is in need of treatment with high-dose steroids; the frequency is even higher among HLA-mismatched unrelated-donor grafts (Ferrara *et al.*, 2009; Weisdorf, 2007).

Acute GVHD occurs most frequently within 100 days after SCT, and this time interval has by tradition previously defined aGVHD versus chronic GVHD (cGVHD). Today this definition of aGVHD and cGVHD has changed to propose a better definition of the severity of the disease and also adequate treatment; the clinical manifestations together with timing offer a better definition of the outcome (aGVHD versus cGVHD) (Hymes *et al.*, 2012).

The clinical manifestations depend on the compatibility of the donor-recipient HLA and the reactivity of the graft to foreign antigens present in the host. Furthermore, the recipient's tissue, that stimulates the lymphocytes from the donor, has been damaged by underlying disease and transplant conditioning. As a consequence of this, the tissue produces pro-inflammatory cytokines, such as IL-1, IL-6, IL-8, and TNF-α, which increase the expression of APC receptors and enhance the presentation of proteins to the donor immune cells (Ball *et al.*, 2008; Dickinson and Charron, 2005; Ferrara *et al.*, 2009). The T_{reg} cells may prevent GVHD due to suppression of T cells in animal models (Cohen and Boyer, 2006).

The organs affected by aGVHD are primary the skin, liver, and the GI tract. Overall grades of aGVHD are classified as mild (I), moderate (II), severe (III), and very severe (IV), based on clinical features of skin involvement, degree of bilirubin elevation, and the volume of diarrhea (Hymes *et al.*, 2012). Biopsies of involved tissue may confirm the diagnosis, especially if the signs are non-specific. Symptoms may involve the upper GI tract (nausea, anorexia, and vomiting), but the most characteristic manifestation is watery, usually voluminous diarrhea (>2 L daily) (Ferrara *et al.*, 2009).

Pathogenesis of graft-versus-host disease

GVHD pathogenesis involves many factors and may arise when T cells from the donor respond to proteins on the host cells, where the most important is the HLA, encoded by *major histocompatibility complex (MHC)* (Ferrara *et al.*, 2009). Non-HLA variants of genes encoding proteins involved in immunity (*INF-γ, IL-10, TNF, TLR4, CD14, NLRP,* and *NOD2*) seem to play a role in the transplantation outcomes (Cavet *et al.*, 1999; Dickinson and Charron, 2005; Dickinson and Holler, 2008; Elmaagacli *et al.*, 2006; Ferrara *et al.*, 2009; Granell *et al.*, 2008; Lin *et al.*, 2003; Sivula *et al.*, 2012). On the other hand, contradictory results have been published. For example, a Swedish study by Sairafi *et al.* (2008) did not replicate the identified association between genetic variation of *NOD2* and the outcome after SCT (Elmaagacli *et al.*, 2006). Inconsistent results may be a consequence of genetic heterogeneity among the patients and further relative small numbers of patients have been included in the studies.

Human genetics

Genetic variation

The most common variations within the human genome, which consists of ~3 billion base pairs, are SNPs. To date, more than 19 million validated SNPs have been deposited in the National Center for Biotechnology Information (NCBI) SNP database (www.ncbi.nlm.nih.gov/projects/SNP).

SNPs are defined as single nucleotide variants at a specific locus with a population frequency of 1% or more (Crawford and Nickerson, 2005; Pettersson *et al.*, 2009). Because autosomal regions carry one allele from the maternal chromosome and one from the paternal chromosome, an individual can exhibit one of three genotypes: homozygous for the major allele, heterozygous, or homozygous for the minor allele. The definitions of major and the minor allele refer to the observed frequencies in a specific population (Crawford and Nickerson, 2005). Some SNPs may affect gene expression levels (Kabakchiev and Silverberg, 2013), and SNPs that occur in coding regions may result in an amino acid substitution, affecting protein function (Sunyaev *et al.*, 2001). Most SNPs, however, are considered functionally neutral (Crawford and Nickerson, 2005).

Genetic association

Genetic association studies can be divided into *candidate gene studies* and *GWAS*, which both involve genotyping of SNPs in cohorts of cases and controls or in families with affected individuals. A candidate gene approach is based on prior hypotheses and known biological function, suggesting an involvement of a genetic region in a specific disease, whereas GWAS are primarily hypothesis-generating studies (Pettersson *et al.*, 2009).

The risk factor for cases in a genetic association study is the allele or the genotype for a specific SNP, where the penetrance reflects the risk of disease with respect to the susceptibility allele (Clarke *et al.*, 2011). The penetrance of genetic variations differs between rare (<0.5% frequency) and common (>5% frequency) polymorphisms, and GWAS studies are designed to detect effects of common variants. Consequently, most of the identified IBD susceptibility alleles are frequently found in the general population. Common genetic variants, located outside of coding regions, typically contribute to associations with a modest odds ratio (OR; OR<1.1), whereas rarer variants with higher effect size (OR >1.5) are more often located in coding regions (Liu and Stappenbeck, 2016).

Genetic association studies for complex diseases may be performed by either *population-based association studies* (unrelated cases and controls) or *family-based association studies*. Studies of unrelated cases and controls are the most commonly utilized approach. The primary advantage for the population-based association study is the enrollment of cases without the need for sampling specific family members. A major disadvantage to both approaches is *the population stratification* (Lewis, 2002; Mersha *et al.*, 2015); however, family-based association studies may be more resistant to this challenge than population-based studies (Cardon and Palmer, 2003; Zondervan and Cardon, 2007). To avoid *false-positive* results, *i.e. type I errors*, in population-based studies, the controls included in the study should be selected from the same population as the cases, to minimize the possibility that the associations may reflect population-specific differences in allele frequencies and not disease-specific associations (Zondervan and Cardon, 2007).

If the *causal* SNP is not genotyped, the effect of this SNP can be assessed through linkage disequilibrium (LD) with surrogate *marker* genotyping. Alleles at different loci that occur together more often than estimated by chance are said to be in LD in the population (Clarke *et al.*, 2011). Because of *e.g.* recombination, the LD is inversely associated with the genetic distance between the SNP markers (Khatkar *et al.*, 2006; Pritchard and Przeworski, 2001). Observed effect size of an association may be interpreted in terms of the effect size of the *causal* variant, the allele frequencies of the genotyped marker, the allele frequencies of the causal variant, and also the LD between

the genotyped marker and the causal variant. These parameters are of importance for the design of association studies (Wray, 2005). To reach (approximately!) the same power at the marker locus as at the susceptibility locus, the sample size would be increased by a factor of $1/r^2$ (Pritchard and Przeworski, 2001; Wray, 2005). Moreover, the statistical power is the probability to correctly reject the null hypothesis when the alternative hypothesis is true (Hong and Park, 2012).

Several measures have been evaluated for assessing the strength of the LD, and the most commonly used methods are r^2 and *D'*. Both of these measurements, which range from 0 (*no disequilibrium*) to 1 (*complete disequilibrium*), are dependent on recombination. Furthermore, the r^2-value only reaches a value of 1 if the allele frequencies of the markers are the same (Ardlie *et al.*, 2002; Wall and Pritchard, 2003). The r^2-value ≥0.8 (pairwise tagging) is the "gold standard" tagging set, *i.e.* the procedure for SNP selection in a manner to cover the genetic region of interest (Wilkening *et al.*, 2009).

Aims

The general aim of this thesis was to identify novel genetic variants in relation to development of intestinal inflammation, with a focus on genetic variation of the TJ structures with implications for epithelial integrity and paracellular permeability.

The specific aims were:

Paper I

To investigate a possible genetic influence of the TJ components *CLDN1, CLDN2*, and *CLDN4* on IBD susceptibility in a large European cohort.

Paper II

To investigate a possible relationship between genetic markers identified in paper I, together with other non-HLA polymorphisms implicated in the etiology of IBD, for the emergence of overall mortality and aGVHD following SCT for hematological malignancies.

Paper III

To investigate putative relationships between IBD and several SNP markers of TJ-related genes that encode interacting proteins, and their corresponding gene expression levels.

Paper IV

To investigate MC in relation to SNP markers of TJ-related genes that encode interacting proteins, and their corresponding gene expression levels.

Materials and methods

Study subjects and healthy controls

The number of patients and controls included in the separate papers are outlined in Table 1, and a detailed description of the study populations is provided below.

Table 1. Number of patients included in the allelic association studies.

	Paper I	*II*	*III*	*IV*
Non-Swedish individuals				
Families	463			
IBD total	715			
CD	528			
UC	151			
IBD-undetermined	36			
Swedish individuals				
IBD total	191[1]		295	
CD	103[1]		138	
UC	102[1]		157	
Controls	333		423[2]	
Patients who underwent SCT		127		
Donors		127		
MC total				104
CC				65
LC				39
Controls				423[2]

[1]In order to avoid bias due to genetic relatedness in the case-control study, one case per family was randomly selected from 191 Swedish families. [2]The same controls were used for paper III and paper IV.

Paper I

All of the IBD families included in paper I originated from the large European collaboration that contributed to the discovery of *NOD2* as a CD susceptibility gene (Hugot *et al.*, 2001). Conventional diagnostic criteria based on clinical, endoscopic, radiological and histological findings were used for the characterization of disease entities (Lennard-Jones, 1989). The case-control studies were based on 191 Swedish IBD patients (from 191 families) and 333 controls selected from an anonymized regional DNA bank, consisting of individuals living in the southeastern part of Sweden. The non-Swedish families (463 families, 715 patients) were included in the family-based genetic association studies.

Paper II

For these studies, 127 Swedish patients who underwent SCT for hematological malignancy at the Department of Hematology at Linköping University Hospital between May 1996 and April 2005 were included, together with their respective donors.

Paper III

Swedish IBD patients (n=295) sampled at Linköping University Hospital and Ryhov County Hospital and controls (n=423) selected from an anonymized regional DNA bank, consisting of individuals living in the southeastern part of Sweden, were used for investigation of genetic association (case-control study). A second Swedish cohort (IBD patients, n=52 and controls, n=33), which was sampled at Linköping University Hospital and had available RNA from endoscopic intestinal biopsies and DNA, was used for follow-up study of significant results of the initial case-control study. The same diagnostic criteria as in paper I were applied (see above).

Paper IV

Swedish MC patients (n=104) sampled at Linköping University Hospital and controls (n=423) selected from an anonymized regional DNA bank, consisting of individuals

living in the southeastern part of Sweden, were used for investigation of genetic associations (case-control study). A second Swedish MC cohort (MC patients, n=25 and controls, n=58) sampled at Linköping University Hospital and Karolinska University Hospital were used in cases where both RNA from endoscopic intestinal biopsies and DNA were available in order to follow up significant results of the initial case-control study.

Methods

The methods included in the separate papers are outlined in Table 2, and a more detailed description of each method is provided below.

Table 2. Methods used in the studies.

Methods	*Paper (I-IV)*
Sample preparation	
DNA preparation from blood	I-IV
DNA preparation from intestinal biopsies	III-IV
RNA preparation from intestinal biopsies	III-IV
cDNA synthesis	III-IV
DNA sequencing	
Sanger sequencing	I
Genotyping	
TaqMan Genotyping	I-IV
Restriction fragment length polymorphism	II
Denaturing high-performance liquid chromatography	II
SNuPe genotyping	II
TaqMan OpenArray	III-IV
mRNA quantification	
Relative RT-qPCR	III-IV

Endoscopic evaluation

In paper III, each biopsy was categorized as inflamed or non-inflamed based on a compound evaluation of endoscopic findings assessed by a single experienced endoscopist and routine histopathologic assessment for inflammation. Only biopsies with concordant results were included in the study. Biopsies were stratified based on sampling location (ileum versus colon) and inflammation for statistical analysis. Moreover, in paper IV, each biopsy was categorized as being diagnostic for MC or not based on the key histological features, described in Langner *et al.* (2015).

Sample preparation

Isolation of DNA from blood

DNA was isolated from blood samples with EDTA using QIAamp DNA minikit (Qiagen, Düsseldorf, Germany; paper II) or MagNA Pure LC DNA Isolation Kit and MagNA Pure extraction robot (Roche, Basel, Switzerland) (paper II-IV) according to the manufacturer's instructions.

Isolation of DNA from intestinal biopsies

Isolation of DNA from intestinal biopsies was performed using Allprep DNA/RNA Mini Kit (Qiagen; paper III-IV) according to the manufacturer's instructions.

Isolation of RNA from intestinal biopsies and conversion to cDNA

The biopsies were placed in RNA *later* (RNA stabilization reagent; Qiagen) and placed at +4°C overnight and thereafter at -20°C (paper III-IV) or at room temperature for one hour and thereafter at -80°C until extraction (paper IV). RNA was purified using the Allprep DNA/RNA Mini Kit (Qiagen) either manually or with the Qiacube extraction robot (Qiagen) according to the manufacturer's instructions. Furthermore, the RNA concentration and purity were determined using Nanodrop ND-1000 (Thermo Fisher Scientific Inc., Waltham, MA, USA), and the integrity of the RNA was assessed using Agilent RNA 6000 Nano kit on the Agilent 2100 bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Then, 2 µg RNA was converted to complementary DNA (cDNA) in a total

volume of 40 µL using the High-Capacity cDNA Reverse Transcription kit with RNase Inhibitor (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. The cDNA was diluted to 25 ng/µl using 0.1×TE buffer and then stored at -80°C until analysis.

Sanger DNA sequencing

Sanger DNA sequencing is based on the presence of free 3'-hydroxyl groups and the incorporation of nucleotides by DNA polymerase during replication and chain-termination (Figure 7). The chain-termination method requires a reaction mixture containing DNA, DNA primers, DNA polymerase, ordinary deoxynucleoside triphosphates (dNTP), and fluorescently labeled dideoxynucleoside triphosphates (ddNTP). The random incorporation of a ddNTP will stop the DNA synthesis, preventing further dNTP incorporation. The generated fragments are then separated by size using capillary gel-electrophoresis, allowing determination of the sequence based on the fluorescence of the ddNTPs.

The promoter region, exon-intron boundaries, and exons that harbor the 5'-untranslated region and protein-coding region of *CLDN2* were amplified by PCR and re-sequenced (paper I). New sequence variants were deposited in the NCBI SNP database (http://www.ncbi.nlm.nih.gov/projects/SNP/).

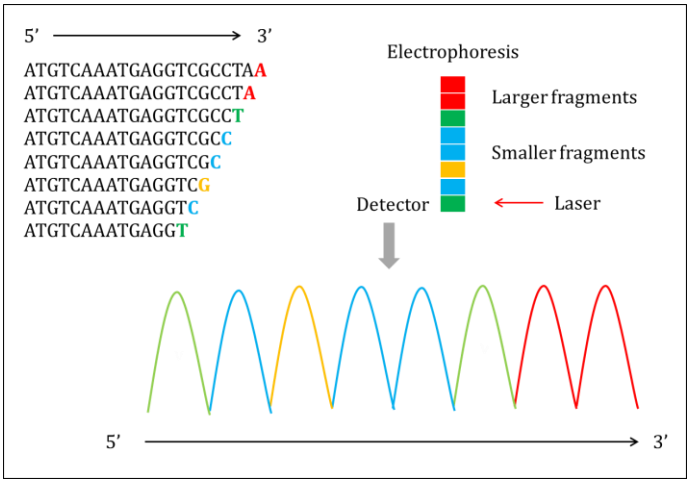


Figure 7. Illustration of Sanger sequencing based on free 3'-hydroxyl groups and the incorporation of nucleotides by DNA polymerase during replication and chain termination.

Genotype analysis

Genetic variants of several genes (Table 3) were investigated in relation to presence of disease, treatment outcome, or gene expression.

Table 3. Genes investigated in the studies.

Gene	*Genetic markers*	*Paper I*	*II*	*III*	*IV*
CLDN1	rs7620166	X			
CLDN2	rs12014762	X	X		
CLDN4	rs8629	X			
MORC4	rs6622126	X	X		
CD14	rs2569190		X		
TLR4	rs4986790		X		
NOD2	rs2066844		X		
NOD2	rs2066845		X		
NOD2	rs2066847		X		
NOD2	rs2066842		X		
SLC22A4	rs1050152		X		
SLC22A5	rs2631367		X		
CARD8	rs2043211		X		
NLRP3	rs35829419		X		
F11R	7 SNP markers			X	X
MAGI1	22 SNP markers			X	X
MAGI2	7 SNP markers			X	X
MAGI3	20 SNP markers			X	X
PARD3	1 SNP marker			X	X
PTEN	4 SNP markers			X	X
TJP1	2 SNP markers			X	X
PTPN22	1 SNP marker			X	

In paper I and II, one or a few SNPs per gene were investigated in relation to IBD and the outcome after SCT, whereas several polymorphisms in each genetic region were studied in relation to IBD and MC in paper III and paper IV.

TaqMan genotyping

TaqMan allelic discrimination assays employ two TaqMan probes that are complementary to the alleles at the polymorphic site (Figure 8), and this assay is based on a 5' reporter dye (FAM or VIC) that is linked to each probe and released in the presence of the probe's target.

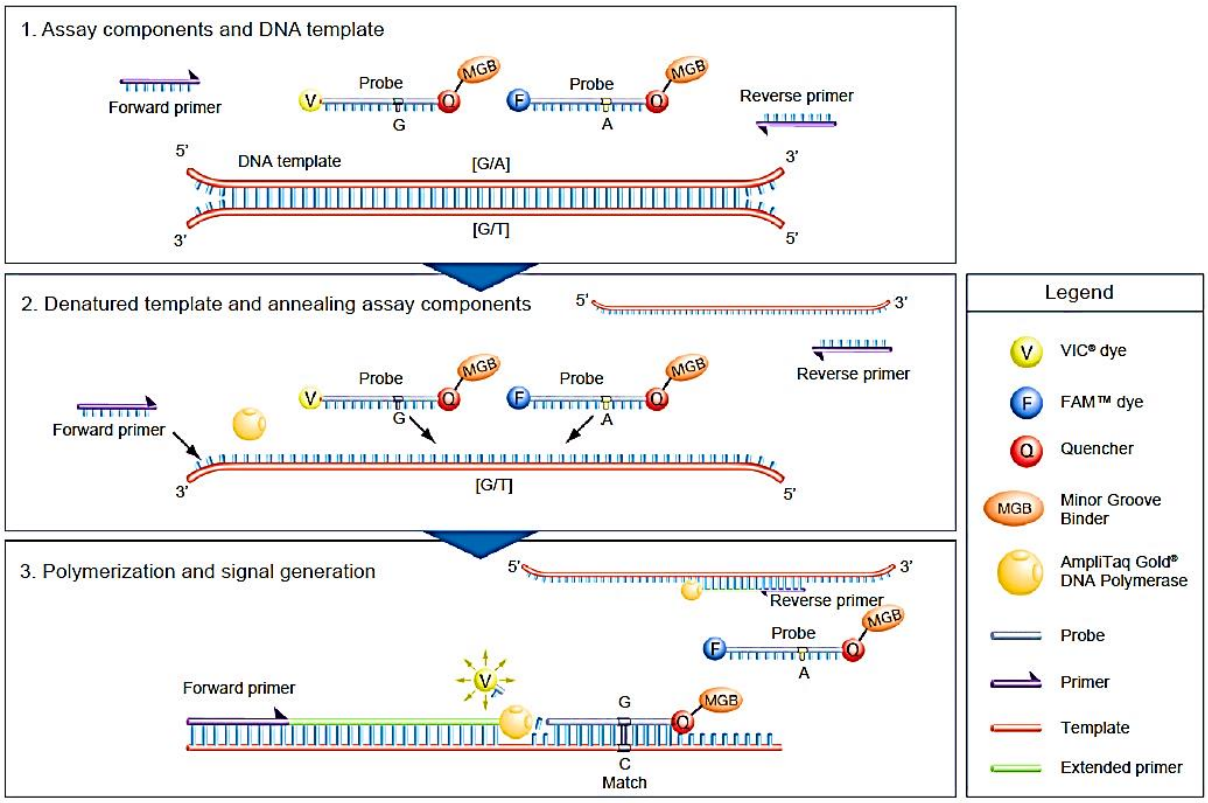


Figure 8. TaqMan chemistry illustration reprinted with permission from www.thermofisher.com. ©Thermo Fisher Scientific Inc.

Allelic discrimination was performed using either the TaqMan real-time PCR system (paper I-IV) or TaqMan OpenArray system (paper III-IV) with the TaqMan SNP genotyping assay (Applied Biosystems, Foster City, CA, USA) in accordance with the manufacturer's recommendations.

Restriction fragment length polymorphism

PCR restriction fragment length polymorphism (RFLP) analysis is a technique that assesses genotype of a specific SNP marker via analysis of the pattern derived from

restriction enzyme digestion of amplified DNA (Figure 9), as restriction enzymes cleave the amplified DNA at a specific recognition sequence (Chuang *et al.*, 2008).

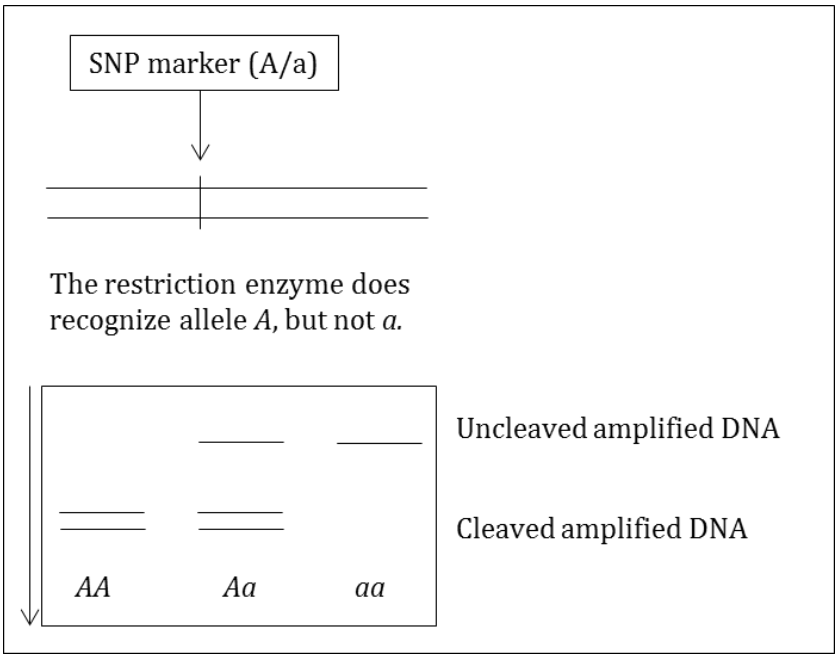


Figure 9. The principle of RFLP analysis, which involves assessment of the pattern following digestion of amplified DNA with a restriction enzyme that recognizes the (A) allele but not the (a) allele. The cleaved products are separated on agarose gels and visualized by DNA-specific staining.

Variants at three SNP positions (*NOD2*-R702W rs2066844, *NOD2*-P268S rs2066842, and *CD14* rs2569190) were determined using RFLP analysis (paper II). Restriction enzyme digestion was performed using *Hpa*II, *Bam*HI, or *Ava*II, and the fragments were further separated on an agarose gel and visualized with a DNA-specific stain.

Denaturing high-performance liquid chromatography

Denaturing high-performance liquid chromatography (dHPLC) is a technology that is used to detect the presence of heteroduplex DNA strands (Figure 10). Amplified products are denatured at a high temperature (95°C) and then re-annealed prior to analysis by dHPLC. During the re-annealing step, both homoduplexes (DNA strands that correspond to the same allele) and heteroduplexes (DNA strands that correspond to different alleles) are formed. Heteroduplexes, which are less stable than homoduplexes, have a shorter retention time compared to homoduplex DNA during dHPLC and therefore give rise to additional peaks. Thus, the first double peaks represent the less stable heteroduplex formations, while the second double peaks represent the

homoduplex formations (Frueh and Noyer-Weidner, 2003; Premstaller and Oefner, 2003).

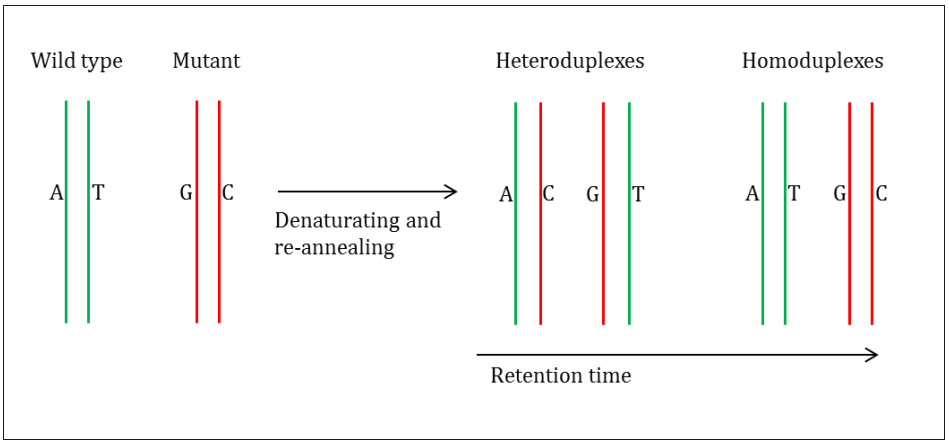


Figure 10. Illustration of the separation of heteroduplex formations from homoduplex formations via dHPLC.

A SNP marker in *TLR4* (rs4986790) was genotyped in paper II using dHPLC. A total of 5 µl of the PCR product was injected into an automated liquid chromatography system (Transgenomic, Dallas, TX, USA), according to the manufacturer's recommendations.

Single nucleotide primer extension

The single nucleotide primer extension (SNuPe) technique involves two distinct steps. In the first step, the genetic region spanning the polymorphic site is amplified and purified from the remaining PCR reagents, and in the second step, an extension primer is added to the reaction. This extension primer is extended with a single fluorescently labeled ddNTP at the polymorphic site (immediately at the 3' end of the primer). The determination of the ddNTP at the specific site, and thereby the genotype, is performed using a DNA sequencer (Syvanen, 1999).

Genetic variants of *NOD2*-G908R (rs2066845), *NOD2*-3020insC (rs2066847), *CARD8* (rs2043211), and *NLRP3* (rs35829419) were determined using the MegaBACE SNuPe genotyping kit (GE Healthcare, Bucks, UK) in paper II. Amplified PCR products were purified using exonuclease I and shrimp alkali phosphatase (Exo-SAP-IT; GE Healthcare) and further mixed with the SNuPe premix and SNP-specific primers. Finally, the genotype was determined using a MegaBACE 1000 DNA sequencing system (GE Healthcare).

TaqMan Gene Expression

Successful quantification of messenger RNA (mRNA) requires input of high-quality RNA. Because RNA molecules are susceptible to degradation by RNase proteins, the quantity and quality of RNA requires verification before reverse transcription (cDNA synthesis) is performed.

Target gene expression can be measured using TaqMan probes, which are short oligonucleotides, and reverse transcription quantitative PCR (RT-qPCR) (Figure 11). For these reactions, one TaqMan probe per transcript is used in combination with the specific unlabeled primers in the amplification. The probe is constructed with a reporter dye at the 5' end as well as a quencher dye at the 3' end. The probe binds to the target between the forward primer and the reverse primer. During the PCR reaction, the *Taq* polymerase cleaves the probe and promotes a separation of the reporter from the quencher, thereby generating a signal that is proportional to the amount of target molecule.

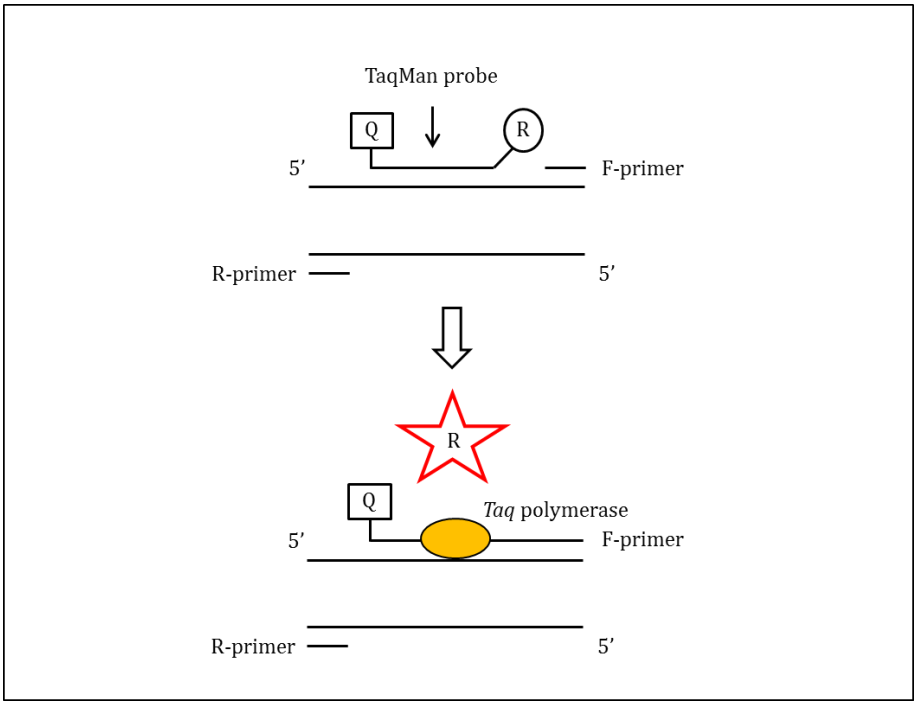


Figure 11. Illustration of the detection of target gene expression using sequence-specific unlabeled primers and TaqMan probes constructed with a reporter and a quencher at the 5' and 3' end of the probe, respectively. F-primer, forward primer; R-primer, reverse primer; R, reporter; and Q, quencher.

The threshold cycle (Ct) value is defined as the number of PCR cycles required to exceed the background level of the PCR reaction. The Ct values are inversely related to the amount of target molecules in the reaction; a lower Ct value corresponds to a higher amount of target (*i.e.*, proportional to the initial gene-specific cDNA). The amount of added cDNA to the PCR reaction is further normalized by endogenous controls, the reference genes.

The cDNA levels were quantified in paper III and paper IV using the TaqMan Gene Expression Assay, TaqMan Universal Mastermix, and a 7500 Fast Real-Time PCR System (Applied Biosystems). Each individual reaction contained 10 ng cDNA in a total volume of 20 μL.

Statistical analysis

In paper I, allelic OR with accompanying 95% confidence interval (CI) and *p*-values based on the χ^2 statistic were calculated using likelihood-based analysis for genetic association (Unphased software version 3.0.13; Dudbridge (2008)) for both the case-control study and the family-based study. Since *CLDN2* and *MORC4* are located on the X-chromosome, males contribute only one allele each, while females contribute two alleles.

Overall mortality and aGVHD (paper II) were analyzed in relation to SNP markers using χ^2 statistics (Microsoft Office Excel 2003; Microsoft Corporation, Redmond, WA, USA) or Fisher's exact test (Statistica version 10; StatSoft Inc., Tulsa, OK, USA) using a case-control approach. Significant findings observed in the genetic association studies were further explored using logistic regression (Statistica version 12). Additionally, to follow up the significant findings identified in the case-control study, Kaplan-Meier survival curves were generated to investigate one-year mortality after SCT (JMP Genomics 6.0; SAS Institute Inc., Cary, NC, USA).

In paper III and paper IV, the allelic OR and *p*-values, based on χ^2 tests were calculated using JMP Genomics 6.0 (JMP Genomics 6.0; SAS Institute Inc.). For these statistical tests, $p<0.05$ (paper III) and $p<0.007$ (paper IV; Bonferroni adjustment based on the number of analyzed genes) were considered significant. Group differences in gene expression were investigated using Kruskal-Wallis ANOVA or the Mann-Whitney U-test (Statistica 12). Logistic regression was used to investigate gene expression levels in relation to intestinal inflammation, phenotype, and gender.

Ct-values were established (ExpressionSuite Software Version 1.0.3; Applied Biosystems) and normalized to the average of selected reference genes (*CASC3* [Hs00201226_m1], *UBA52* [Hs03004332_g1], and *POP4* [Hs00198357_m1]) (Söderman *et al.*, 2015), generating the delta-threshold cycle (ΔCt) values. Relative quantification (RQ) values were further established via relating the ΔCt values to the sample with the

lowest gene expression for each gene using Microsoft Office Excel 2010 (Microsoft Corporation), according to the methods of Livak and Schmittgen (2001). Group differences in gene expression levels were calculated as a fold change (fc) based on RQ values. Bonferroni-adjusted (based on the number of analyzed genes) *p*-values <0.008 (paper III) and <0.017 (paper IV) were considered significant for statistical analysis based on gene expression levels.

Ethics

The four studies were conducted with approval by the ethics committees as follows:

- *Paper I*
  Ethics committees of Linköping University (Dnr 97271) and Karolinska Institutet (Dnr 97-327).

- *Paper II*
  Ethics committee of Linköping University (Dnr M82-05).

- *Paper III*
  Ethics committee of Linköping University (Dnr M35-07 and Dnr 2011/201-31).

- *Paper IV*
  Ethics committees of Linköping University (Dnr M35-07, Dnr 2011/201-31) and Karolinska Institutet (Dnr 2007/791-31/3).

Results and discussion

This section summarizes the main findings of the thesis, while detailed results and a more extensive discussion can be found in respective paper (paper I-IV).

Paper I

We investigated the *CLDN* genes as candidate genes that may contribute to the increased intestinal permeability of patients with IBD. Based on a pilot-study of Swedish IBD patients, one SNP marker per gene (*CLDN1, CLDN2,* and *CLDN4*) was selected for investigation in the total study population.

Using a case-control study approach with *Swedish patients and controls*, we evaluated the associations between each of the three markers in the genetic region of *CLDN1, CLDN2,* or *CLDN4* and the IBD phenotypes (IBD, CD, or UC; Table 4). The strongest significant association was observed between the *CLDN2* marker (rs12014762, susceptibility allele C) and CD. Another significant association was identified between the *CLDN1* marker (rs7620166, susceptibility allele T) and overall IBD, while no significant associations were identified for the *CLDN4* marker with any of the phenotypes (IBD, CD, or UC).

The same SNP markers were also included in a family-based study of *European families of non-Swedish origin*. The significant findings observed in the case-control study with Swedish patients, were not replicated by the family-based approach in the non-Swedish cohort. Regional heterogeneity of allele frequencies among different sub-populations may explain the discrepant findings. Such heterogeneity of allele frequencies has previously been observed and described by several groups (Ellinghaus *et al.*, 2016;

Jostins *et al.*, 2012; Liu *et al.*, 2015). On the other hand, the divergent results may be explained by a false-positive finding in the case-control study of Swedish participants or alternatively by a false-negative finding in the family-based approach. The use of homogenous populations should enable the identification of genetic risk factors that are more prevalent and important in a specific population.

Table 4. Case-control approach[1] and family-based[2] genetic association studies[3] with respect to IBD-phenotype.

		rs7620166 (*CLDN1*)		**rs12014762 (*CLDN2*)**		**rs8629 (*CLDN4*)**	
		allelic OR (95% CI)	*p*-value	allelic OR (95% CI)	*p*-value	allelic OR (95% CI)	*p*-value
IBD	Swedish case-control	1.33 (1.04–1.72)	0.025	1.39 (0.95–2.01)	0.083	1.21 (0.89–1.65)	0.225
	Non-Swedish fam.	0.87 (0.72–1.06)	0.177	1.25 (0.89–1.77)	0.195	1.09 (0.88–1.33)	0.432
CD	Swedish case-control	1.17 (0.86–1.60)	0.319	1.98 (1.17–3.35)	0.007	1.25 (0.84–1.85)	0.258
	Non-Swedish fam.	0.80 (0.64–1.00)	0.052	1.37 (0.91–2.07)	0.126	1.14 (0.89–1.46)	0.287
UC	Swedish case-control	1.35 (0.98–1.84)	0.064	1.27 (0.80–2.02)	0.304	1.18 (0.80–1.73)	0.409
	Non-Swedish fam.	1.19 (0.77–1.84)	0.436	0.91 (0.39–2.14)	0.827	1.15 (0.75–1.77)	0.512

[1]Results from the case-control stud are based on 191, 103 and 102 cases of IBD, CD, and UC, respectively as well as 333 controls.

[2]The family-based association studies included 463 European families. For the SNP-markers rs7620166 (CLDN1), rs12014762 (CLDN2), and rs8629 (CLDN4), genotyping was unsuccessful for 5 (including 2 CD), 9 (including 3 CD), and 7 (including 2 CD) samples, respectively.

[3]OR, its related 95-% CI, and p-values (based on log likelihood ratio chi-square statistics) were calculated for the T allele of rs7620166, the C allele of rs12014762, and the C allele of rs8629.

CLDN1 and CLDN4 have been associated with a tight TJ structure, whereas CLDN2 expression results in a leakier epithelial layer (Furuse *et al.*, 2001; Krug *et al.*, 2014). The expression of the CLDNs may have different patterns within the GI tract, and these patterns may contribute to differences in epithelial integrity. Furthermore, the expression of CLDN2 differs along the crypt-to-villi axis, where the protein is restricted to the crypts in the small as well as colonic intestine of the rat (Rahner *et al.*, 2001). The expression of *CLDN2* mRNA as well as protein is increased in inflamed intestine, such as in cases of CD and UC, compared to non-inflamed intestine. Also, higher CLDN2 expression is observed in patients with UC compared to patients with CD (Prasad *et al.*, 2005; Weber *et al.*, 2008). Furthermore, studies have suggested that differences in

epithelial permeability stem from compositional differences in the TJ-structure (Luissint *et al.*, 2016).

Based on the most significant association detected in the case-control study, we re-sequenced *CLDN2* as a mean to identify novel sequence variants of importance for the development of IBD. With this approach, two novel polymorphisms (rs62605981 and rs72466477), which are located in the promotor region of the *CLDN2* gene (Sakaguchi *et al.*, 2002) were identified. Neither of these markers were, however, significantly associated with either IBD overall or with CD or UC.

MORC4 is located in the same genetic region as *CLDN2* and harbors a nonsynonymous SNP (rs6622126). This nonsynonymous SNP was investigated for genetic association to IBD, CD, or UC. A significant association was identified between the *MORC4* SNP (rs6622126, susceptibility allele G, p=0.018) and CD, but not between the SNP and IBD overall or UC. MORC4 has been identified as a putative interacting protein involved in the regulation of TGF-β (Colland *et al.*, 2004). Increased intestinal expression of *TGF-β* has been observed in CD patients, and TGF-β is involved in maintenance and recovery of the intestinal epithelial barrier function (Howe *et al.*, 2005; Planchon *et al.*, 1999; Suenaert *et al.*, 2010). In addition, the integrity of the epithelial barrier correlates with the TGF-β expression level, indicating a protective effect of TGF-β (Suenaert *et al.*, 2010).

Paper II

Non-HLA variants of genes encoding proteins involved in immunity seem to play a role in the outcome of stem cell transplantation (Cavet *et al.*, 1999; Dickinson and Charron, 2005; Dickinson and Holler, 2008; Elmaagacli *et al.*, 2006; Ferrara *et al.*, 2009; Granell *et al.*, 2008; Lin *et al.*, 2003; Sivula *et al.*, 2012). We therefore investigated if several non-HLA markers, including *CLDN2, MORC4,* and *NOD2*, were associated with the outcome after SCT (overall mortality and aGVHD; selected polymorphisms are described in Table 3).

Genetic association with respect to recipient genetic makeup

Significant associations were observed between overall mortality and genetic variation in *MORC4* (rs6622126, susceptibility allele A) as well as in *CD14* (rs2569190, susceptibility allele T) (Table 5). Weaker associations were also observed between overall mortality and genetic variation in *TLR4* (rs4986790, susceptibility allele A) as well as the presence of at least one of the three *NOD2* susceptibility alleles (rs2066844; susceptibility allele T; rs2066845, susceptibility allele C; rs2066847, susceptibility allele C). Since *MORC4* and *CLDN2* are located on the X chromosome, men and women were further analyzed separately. A significant association was observed between overall mortality and *MORC4* (allele A) in males but not in females.

No associations were observed between overall mortality and any of the SNPs in the genetic regions of *NOD2* (each marker separately)*, SLC22A4, SLC22A5, CARD8, NLRP3,* or *CLDN2.* Genetic variation of *CD14* (rs2569190, susceptibility allele T) was the only variation identified in significant association with aGVHD.

In paper I, we described an association between genetic variation in *MORC4* (rs6622126, susceptibility allele G) and CD in a Swedish population. In contrast, overall mortality among SCT patients was associated with the opposite allele (A) of the *MORC4* SNP compared to the susceptibility allele described in relation to CD. Dissimilar effects of alleles of a SNP marker have previously been shown, *i.e.* the A allele of a *NOD2* SNP

marker (rs5743289) was associated with CD whereas the G allele was associated with leprosy (Jostins *et al.*, 2012). Moreover, an association has also been reported for this genetic region (*CLDN2-MORC4)* and chronic pancreatitis (Giri *et al.*, 2016; Whitcomb *et al.*, 2012). MORC4 has been identified as a putative interacting protein involved in the TGF-β regulation (Colland *et al.*, 2004). Together, these results indicate that genetic variation within *MORC4* may predispose to inflammatory conditions in different tissues.

Based on the heterogeneity of the cohort, logistic regression was used to control for putative confounders, such as diagnosis, various conditioning protocols, GVHD prophylaxis, relationship between recipient and donor, and age at SCT, one factor at a time. The associations between *MORC4* (male recipients), *CD14*, *TLR4,* and overall mortality remained statistically significant. The logistic regression did, however, not confirm the significant association between *NOD2* combined (rs2066844; susceptibility allele T; rs2066845, susceptibility allele C; rs2066847, susceptibility allele C) and overall mortality. The significant association between *CD14* and aGVHD was confirmed by logistic regression, and diagnosis also contributed to the outcome.

CD14, together with TLR4, binds bacterial lipopolysaccharide and is further involved in the activation and regulation of NFκB (da Silveira Cruz-Machado *et al.*, 2010). Therefore, the observed association between genetic variation of *CD14* and *TLR4* and the outcome after SCT suggests a potential dysregulated immune response due to improper NFκB regulation. Moreover, higher level of circulating blood $CD14^+$ cells has previously been linked to the outcome after SCT (Arpinati *et al.*, 2007).

Genetic association with respect to donor genetic makeup

With respect to the donor genotype, a significant association was observed between the SNP marker in *MORC4* and overall mortality (Table 5). No additional associations were identified between overall mortality and any of the variations in the genetic regions of *CD14, TLR4, NOD2, SLC22A4, SLC22A5, CARD8, NLRP3,* or *CLDN2*, and no donor SNPs were significantly associated with aGVHD. Furthermore, logistic regression confirmed the significant association between donor *MORC4* genotype and overall mortality in male recipients.

Genetic association with respect to combined recipient and donor genotype

The effect of the combined allele composition of the donor and recipient for each gene was investigated in relation to aGVHD and overall mortality. A significant association was observed between overall recipient mortality and the *MORC4* SNP (rs6622126, susceptibility allele A, in male pairs) and the *CD14* SNP (rs2569190, susceptibility allele T, in all paired cases). Logistic regression confirmed the association between *MORC4* variation and mortality but not between the *CD14* variant and overall mortality.

One-year survival

To confirm the significant findings of the association studies, Kaplan-Meier survival curves were generated to analyze one-year mortality after SCT. The overall mortality in male patients was significantly increased in recipients carrying the *MORC4* susceptibility allele (rs6622126, A) compared to that in recipients carrying the non-susceptibility allele (G) (p=0.013; Figure 12). Additionally, increased recipient mortality was observed in patients who carried one susceptibility allele of *NOD2* (rs2066844; susceptibility allele T; rs2066845, susceptibility allele C; rs2066847, susceptibility allele C), compared to carrying of none (p=0.045; Figure 13).

Although this study lacks details about the transplant-related mortality, the Kaplan-Meier analysis of *MORC4* in relation to one-year mortality after SCT indicates a decreased survival in the short term after transplantation. This increased overall mortality may be a consequence of dysfunctional immunological reactions after SCT.

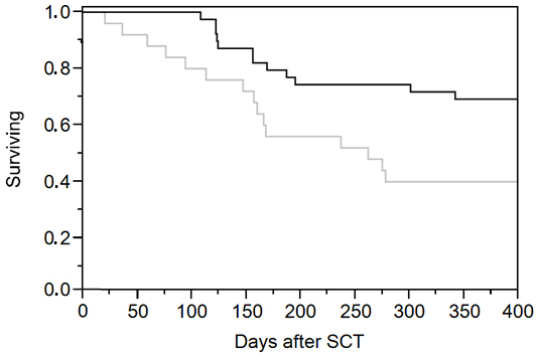


Figure 12. Male survival after SCT according to genetic variation of MORC4 (rs6622126). The grey line indicates the A allele, and the black line indicates the G allele (p=0.013). An increased overall mortality was observed in recipients carrying the susceptibility allele (rs6622126, A) of MORC4 compared to the non-susceptibility allele (G).

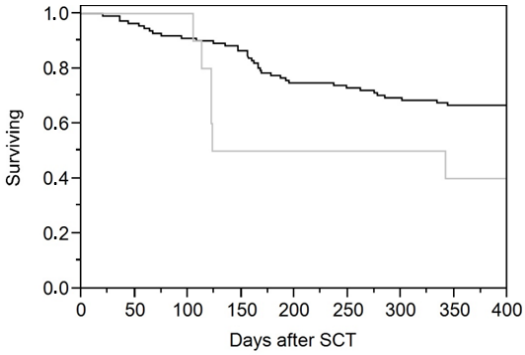


Figure 13. Combined male and female survival after SCT in relation to carriage of one susceptibility allele of NOD2 (p=0.045) rs2066844, rs2066845, and rs2066847). The grey line indicates the carriage of one allele, and the black line indicates carriage of no susceptibility allele.

Table 5. Genetic association between recipient genetic makeup, donor genetic makeup, or combined genetic makeup and the outcome after allogeneic SCT.

Gene	*SNP (rs nr)*	*Mortality allelic OR (95% CI)*	*p-value*	*Allele*	*aGVHD allelic OR (95% CI)*	*p-value*	*Allele*
Recipient genetic makeup							
MORC4	rs6622126	2.03 (1.07-3.86)	0.029	A	1.42 (0.72-2.82)	0.314	A
Female	rs6622126	1.53 (0.68-3.45)	0.304	A	1.66 (0.65-4.25)	0.286	A
Male	rs6622126	4.21 (1.31-13.48)	0.012	A	1.33 (0.47-3.78)	0.588	A
CD14	rs2569190	1.75 (1.05-2.91)	0.031	T	1.86 (1.09-3.20)	0.023	T
TLR4	rs4986790	2.94 (0.99-8.73)	0.043	A	1.81 (0.65-5.04)	0.271	G
NOD2 (R702W, G908R, 3020insC)		4.21 (0.90-19.62)	0.048	Any risk allele	1.16 (0.34-3.97)	0.753	Any risk allele
Donor genetic makeup							
MORC4	rs6622126	1.80 (0.96-3.36)	0.064	A	0.83 (0.43-1.59)	0.577	A
Female	rs6622126	1.17 (0.54-2.52)	0.695	A	0.82 (0.37-1.81)	0.620	A
Male	rs6622126	5.94 (1.68-20.99)	0.004	A	1.04 (0.31-3.43)	0.951	A
Combined donor and recipient allele composition							
MORC4 (male couple)	rs6622126	7.20 (2.14-24.23)	0.001	A	1.42 (0.50-4.05)	0.510	A
CD14	rs2569190	1.53 (1.07-2.19)	0.018	T	1.36 (0.93-1.98)	0.113	T

Paper III

Using a Swedish population, we investigated genetic variation of several TJ-related genes (Table 3) in relation to IBD. Also ileal and colonic gene expression was analyzed in relation to genotype, inflammatory status, phenotype, and ongoing medical treatment.

Of the 64 identified genetic markers, 12 were excluded due to failed genotyping or absence of Hardy-Weinberg equilibrium at those loci. The strongest significant associations were observed between IBD and SNP markers in *MAGI2* (rs6962966; susceptibility allele A) as well as *MAGI3* (rs1343126; susceptibility allele T) (Table 6; significant SNP markers). These markers were also associated with CD as well as UC, independently, although the *MAGI3* SNP marker was only borderline significantly associated with CD. In addition, significant associations were observed between additional SNP markers and IBD as well as UC (Table 6).

To the best of our knowledge, we have for the first time demonstrated an association between *MAGI3* and IBD. Previously, Jostins *et al.* (2012) identified a genetic association between this region and CD, although they primarily highlighted *PTPN22*. The associated *MAGI2* SNP marker (rs6962966) was previously described by Wapenaar *et al.* (2008) in association with UC; however, these investigators described an association between UC and the opposite allele (G), whereas we revealed a significant association between IBD and the (A) allele. We currently have no reliable explanation for this discrepancy.

In our association study, $p<0.05$ was considered significant; however, even though we applied a more stringent p-value (Bonferroni adjustment based on the number of analyzed genes, $p<0.008$), the association between the *MAGI3* rs1343126 and *MAGI2* rs6962966 SNPs and IBD remained significant. The significant findings ($p<0.05$) in the case-control approach were further followed up by analyzing for genotype-gene expression relationships in a second independent cohort.

All genes (Table 3), except *TJP1,* were expressed at equal levels in biopsies from different segments of the colon (cecum, ascending colon, transverse colon, descending colon, sigmoid colon, and rectum). *TJP* exhibited only slightly higher expression in the sigmoid colon than in the ascending colon (p=0.006; fc=1.16). Because of these equivalent levels, the colonic biopsies were treated as biological replicates.

Carriage of the *MAGI3* rs6689879 UC susceptibility allele contributed to increased *MAGI3* expression in ileal non-IBD mucosa, and this increased *MAGI3* expression may contribute to the observed restriction of inflammation to the colon in UC. In fact, all observed relationships, both significant and nominally significant, between gene expression levels and susceptibility alleles were identified in mucosal biopsies from non-inflamed non-IBD controls. It is unlikely that the genotype-gene expression relationship in non-inflamed IBD biopsies was concealed by subclinical inflammation, since these biopsies were assessed as histopathologically normal. It is possible though that gene expression levels were affected by other factors, such as miRNA, that were present in the non-inflamed mucosa from IBD patients (Fasseu *et al.*, 2010; Peloquin *et al.*, 2016; Planell *et al.*, 2013).

Colonic biopsies from inflamed mucosa from IBD and UC patients expressed lower levels of *MAGI3*, *PTEN*, and *TJP1*, than those from non-inflamed mucosa. MAGI3 is involved in the suppression of the PI3K/Akt and Wnt/β-catenin signaling pathways (Figure 14) (Ma *et al.*, 2015a; Ma *et al.*, 2015b). Both of these signaling pathways have been shown to promote inflammation (Huang *et al.*, 2011; Keerthivasan *et al.*, 2014) and increased activation of PI3K/Akt inhibits the proliferation of T cells towards T_{reg} cells (Sauer *et al.*, 2008). Thus, regulation of these pathways may constitute a mechanism by which decreased level of *MAGI3* promotes colonic inflammation.

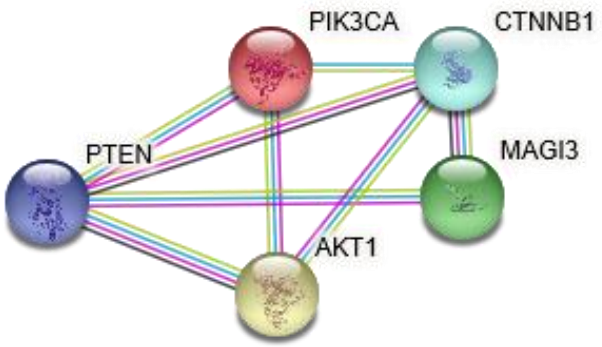


Figure 14. An illustration of the involvement of MAGI3 in the suppression of PTEN and CTNNB1. The network was identified using the STRING search tool (String consortium).

The expression level of *PTEN* was also decreased in inflamed colonic mucosa of CD patients compared to non-inflamed CD mucosa (Figure 15; *complete separation* of *PTEN* expression), suggesting *PTEN* as a marker for inflammatory response in the colonic mucosa of CD patients. PTEN is described as a tumor suppressor molecule inhibiting inflammatory response via PI3K/Akt pathway (Tokuhira *et al.*, 2015), and previously decreased expression level of *PTEN* was observed in intestinal mucosal lymphocytes in CD, compared to controls (Long *et al.*, 2013).

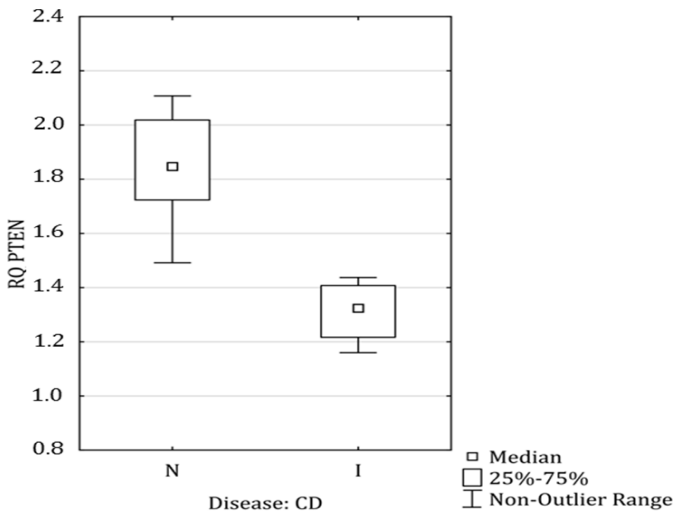


Figure 15. Box plot illustration of PTEN expression level in inflamed (I) colonic mucosa from patients with CD and non-inflamed (N) colonic mucosa.

PTPN22 was included in our study, since this gene is located in the same genetic region as *MAGI3* and has been described in relation to IBD previously (Diaz-Gallo *et al.*, 2011; Jostins *et al.*, 2012; Rivas *et al.*, 2011). Biopsies from inflamed colonic mucosa from IBD

patients exhibited significantly higher levels of *PTPN22* expression than biopsies from non-inflamed colonic mucosa. PTPN22, which negatively regulates T cell activation (Sharp *et al.*, 2015), is expressed at higher levels in immune cells than in non-immune cells (Arimura and Yagi, 2010). Since the expression level of *PTPN22* was determined using RNA from homogenized biopsies, it is not possible to determine the exact cellular origin of this expression. The increased number of immune cells in the inflamed mucosa may, however, have contributed to the increased *PTPN22* gene expression level in colonic IBD mucosa.

Table 6. SNP markers observed to be significantly associated with CD, UC, or IBD (overall).

Gene	NCBI SNP (rs number)	Allele	IBD (CD and UC)				CD				UC			
			p value	OR	Lower CL	Upper CL	*p* value	OR	Lower CL	Upper CL	*p* value	OR	Lower CL	Upper CL
F11R	rs7546890	T	0.414	1.10	0.88	1.36	0.381	0.88	0.66	1.17	**0.043**	1.32	1.01	1.74
MAGI2	rs7803276	C	0.294	1.12	0.90	1.39	0.550	0.92	0.69	1.22	**0.031**	1.35	1.03	1.77
MAGI2	rs6962966	G	**0.004**	0.72	0.58	0.90	**0.008**	0.69	0.52	0.91	**0.045**	0.76	0.58	0.99
MAGI3	rs6689879	T	0.211	0.86	0.68	1.09	0.896	1.02	0.75	1.39	**0.043**	0.75	0.56	0.99
MAGI3	rs1343126	T	**0.004**	1.43	1.12	1.84	0.051	1.38	1.00	1.91	**0.011**	1.48	1.10	2.01
MAGI3	rs12119076	C	**0.022**	1.33	1.04	1.69	0.088	1.31	0.96	1.78	0.050	1.34	1.00	1.80
PTEN	rs1234224	G	0.156	1.18	0.94	1.47	0.841	1.03	0.77	1.38	**0.046**	1.32	1.00	1.74
TJP1	rs260526	G	0.060	1.48	0.98	2.23	0.731	1.09	0.67	1.78	**0.010**	2.12	1.18	3.82

Bold type indicates statistical significance ($p<0.05$).

Paper IV

Since the etiology of MC is largely unknown, even if a genetic predisposition may be involved (Koskela *et al.*, 2011; Koskela *et al.*, 2008; Madisch *et al.*, 2011; Westerlind *et al.*, 2015). We investigated the genetic variation of several TJ-related genes (Table 3) in relation to MC. Subsequently, we analyzed colonic gene expression levels in relation to genotype, phenotype, and gender.

Seven out of the 63 SNP markers were excluded due to failed genotyping or deviation from the Hardy-Weinberg equilibrium. The strongest association was observed between the *PTEN* SNP marker (rs1234224; susceptibility allele G) and MC overall and CC (Table 7). PTEN is involved in the suppression of the PI3K/Akt pathway by dephosphorylation of phosphatidylinositol (3,4,5)-trisphosphate (PIP3), resulting in the biphosphate product phosphatidylinositol (4,5)-bisphosphate (PIP2) (Maehama and Dixon, 1998) (Figure 16).

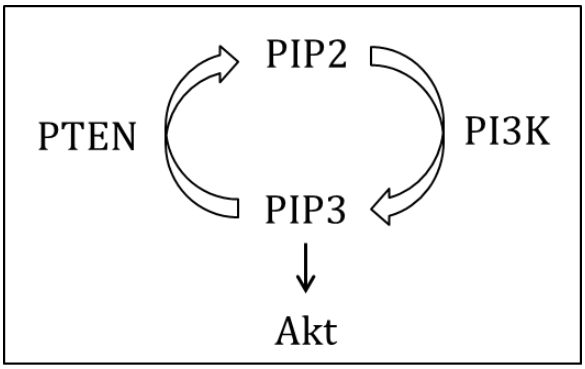


Figure 16. An illustration of PI3K/Akt regulation. PTEN inhibits the phosphorylation of PIP2 and thereby preventing the Akt activation.

We observed significantly decreased levels of *PTEN* expression in MC and CC colonic mucosa, compared to non-MC controls. Decreased PTEN expression has also been observed in association with pulmonary fibrosis, which is characterized by activation and proliferation of fibroblasts and collagen secretion (He *et al.*, 2012). Analogously, CC

is associated with a distinctive thickened sub-epithelial collagen layer (Langner *et al.*, 2015).

A *MAGI1* SNP marker (rs17417230) was associated with MC (Table 7), and decreased expression of *MAGI1* was further associated with MC, CC, and LC. Moreover, the expression levels of *MAGI1* and *PTEN* were positively correlated (p=4.10×10^{-4}, r_s=0.65) among the MC patients, but not among the non-MC controls (Figure 17). In line with this observation, recruitment of PTEN by MAGI-1b at AJs has been observed (Kotelevets *et al.*, 2005).

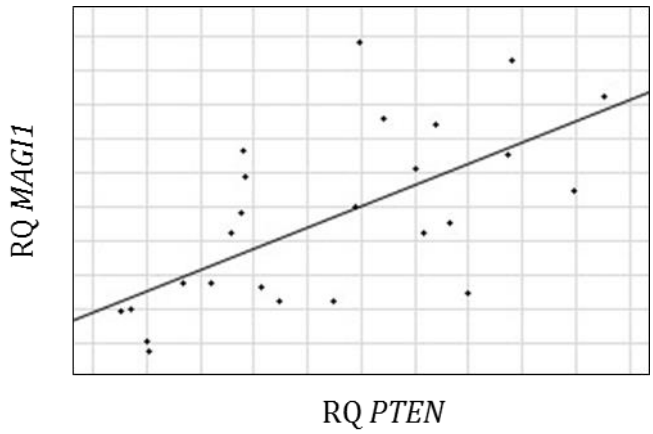


Figure 17. Correlation between RQ values of PTEN and MAGI1 in colonic mucosa from MC patients (p=4.10×10^{-4}, r_s=0.65).

A *F11R* SNP marker (rs790055) was significantly associated with CC (Table 7), and a nominally significant association (p=0.042) was observed between decreased *F11R* gene expression and MC overall. In addition to a thickened sub-epithelial collagen layer, CC is also associated with increased mononuclear inflammation in the lamina propria (Langner *et al.*, 2015). Reduced JAM-A expression is associated with increased infiltration of PMN in colonic mucosa in patients with colitis (Kucharzik *et al.*, 2001), and increased PMN infiltration and lymphoid aggregation are additionally observed in colonic mucosa from JAM-A$^{-/-}$ mice (Laukoetter *et al.*, 2007). Reduced levels of JAM-A have been described for inflamed CD and UC tissue compared to levels in controls (Vetrano *et al.*, 2008).

Table 7. SNP markers were observed to be significantly (Bonferroni adjustment based on the number of analyzed genes, p<0.008) associated with MC and CC.

Gene	NCBI SNP Reference	Allele	CC & LC				CC			
			p-value	OR	Lower CL	Upper CL	*p*-value	OR	Lower CL	Upper CL
F11R	rs790055	G	0.013	1.90	1.13	3.17	**0.007**	2.58	1.27	5.25
MAGI1	rs17417230	C	**0.006**	1.58	1.14	2.19	0.015	1.63	1.10	2.42
PTEN	rs1234224	G	**0.001**	1.70	1.23	2.34	**0.003**	1.79	1.22	2.62

Note: No markers were significantly associated with LC.

None of the selected SNPs affected the gene expression levels in either MC patients or non-MC controls. Nevertheless, a non-significant decreased *PTEN* expression in MC and CC patients was observed in relation to homozygosity for the susceptibility allele of the *PTEN* SNP marker rs1234224 (*p*-value=0.081 and *p*-value=0.103, respectively) (Figure 18). It is however possible that a larger study population would allow identification of more subtle correlations between SNP markers and their corresponding gene expression levels.

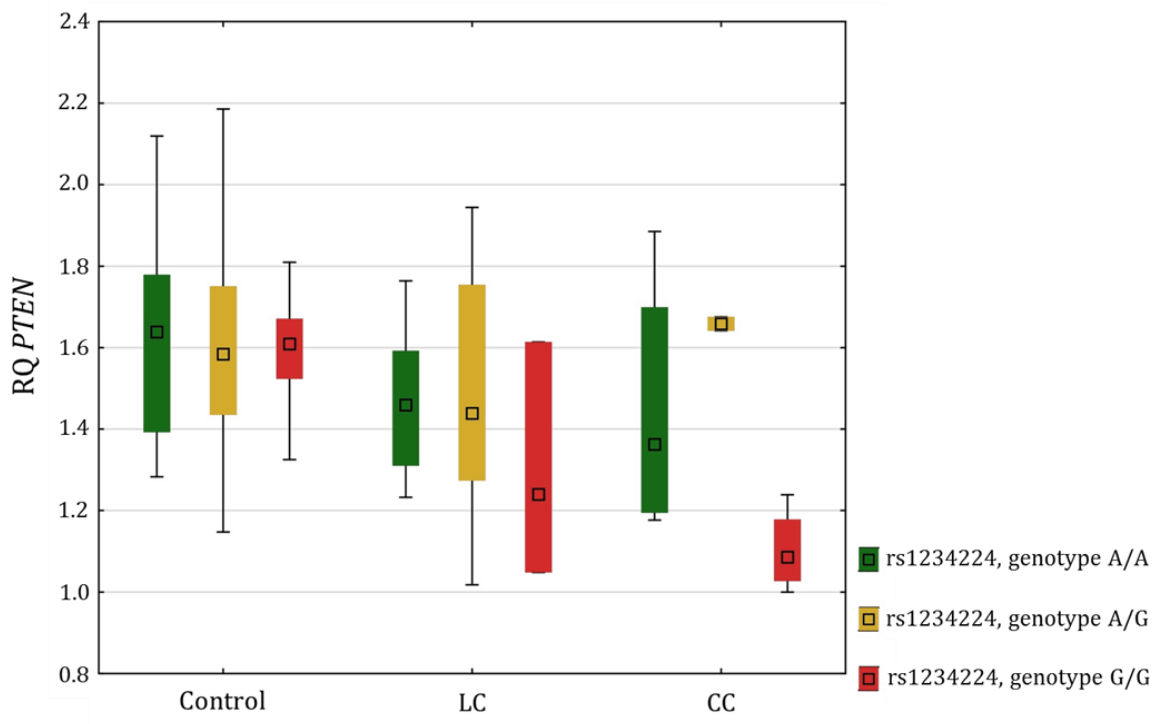


Figure 18. Box plot illustration of PTEN expression levels in colonic mucosa from non-MC controls, LC patients, and CC patients stratified based on rs1234224 genotype.

Thesis contribution

At the time when this work began (2007), increased intestinal paracellular permeability was established among IBD patients (Prasad *et al.*, 2005). It was also known that the composition of the TJ structure was involved in the regulation of paracellular permeability (Furuse *et al.*, 2001), and a number of studies pointed toward affected levels of TJ proteins among IBD patients (Kucharzik *et al.*, 2001; Prasad *et al.*, 2005). Nothing was, however, known regarding the importance of genetic variation of genes encoding proteins belonging to the TJ complex in relation to IBD.

This thesis have demonstrated that genetic variation of genes, encoding both transmembrane and intracellular TJ proteins, may predispose individuals to IBD and MC. Correlations between genotype and gene-expression have been described, as well as affected gene expression levels in relation to phenotypes and inflammation.

During this journey, a number of GWAS studies have been published and also supported that genetic variation of proteins involved in the regulation of the epithelial integrity may predispose individuals to IBD (Anderson *et al.*, 2011; Barrett *et al.*, 2008; Ellinghaus *et al.*, 2016; Franke *et al.*, 2010; Jostins *et al.*, 2012; Liu *et al.*, 2015; McGovern *et al.*, 2010). GWAS studies are, however, based on regionally heterogeneous populations. The papers of this thesis were mainly based on studies in homogenous patient cohorts to detect population-specific variations among IBD patients from Sweden. Therefore, our studies were able to uncover regionally relevant genetic variation of TJ structures in relation to IBD and MC in a meaningful way. Previous results from other groups together with our own results are summarized in Table 8.

The general conclusion presented in this thesis is that genes encoding proteins involved in regulation of the paracellular permeability, with focus on the TJ structure, may

predispose individuals to intestinal inflammation, such as IBD and MC. Important findings are that *MAGI3* is related to the development of IBD and that *MORC4* predisposes to CD and an increased risk of one-year mortality after SCT for hematological malignancy. Overall the findings underscore a polygenetic influence in the pathophysiology of classical IBD and MC.

Table 8. Summary of genetic variations in IBD of genes encoding TJ proteins.

Gene	**Chromosome**	**Author**
CLDN2	Xq22.3	Söderman *et. al*. this thesis
F11R	1q23.3	Norén *et. al*. this thesis
GNA12	7p22.3-p22.2	Anderson *et al.* (2011)
MAGI2	7q21.11	Wapenaar *et al.* (2008) McGovern *et al.* (2009) Norén *et. al*. this thesis
MAGI3	1p13.2	Norén *et. al*. this thesis
MYO9B	19p13.11	Cooney *et al*. (2009) Li *et al.* (2016)
PTEN	10q23.31	Norén *et. al*. this thesis
TJP1	15q13.1	Norén *et. al*. this thesis

Conclusions

Based on the observations described in papers I-IV, the following conclusions can be drawn:

- Genes encoding proteins involved in the regulation of paracellular permeability via TJ structures may predispose to CD, UC, and MC (paper I, III, and IV).

- Genetic variation of *CLDN2* may be a useful marker for predisposition to CD. This impact may occur directly through CLDN2 or indirectly via a putative link between MORC4 and TGF-β regulation (paper I).

- Genetic variation of *MORC4* may predispose to CD (paper I) and an increased risk of one-year mortality after SCT for hematological malignancy (paper II).

- Non-HLA genetic markers involved in the etiology of IBD (*TLR4* and *CD14*) may impact the outcome (acute GVHD and overall mortality) after allogeneic SCT for hematological malignancy (paper II).

- Genetic variation of *MAGI3* is associated with IBD (paper III).

- *PTEN* may act as a marker for the inflammatory response in colonic mucosa of patients with CD (paper III).

- Genetic variation of *F11R, MAGI1*, and *PTEN* is associated with MC (paper IV).

Acknowledgements

*First of all, I would like to take the opportunity to thank **all the people,** who in one way or another have supported me during this journey.*

Sven Almer, min huvudhandledare. Tack för att du gav mig möjlighet att arbeta med forskning. Jag vill tacka för all tid och all stöttning du gett mig, dagar som sena kvällar. Tack för din gränslösa entusiasm och för att du delar med dig av din kunskap inom gastroenterologi och vetenskap. TACK!

Jan Söderman, min biträdande handledare. Tack Jan för att du trott på mig och för att din dörr alltid varit öppen. Tack för att du accepterat (vare sig du velat eller inte) mina mentala utbrott som resulterat i både skratt och tårar. Jag är tacksam över att du många gånger visat mig vägen; ”Fokus Bettan!”, då mina tankar flugit iväg mot andra spännande frågeställningar. Tack för givande diskussioner och för att du alltid delar med dig av din kunskap. Jag är otroligt tacksam och glad över den vänskap som vuxit fram genom åren och jag ser med glädje fram emot ytterligare samarbete tillsammans med dig. Jag ser även fram emot att springa halvmaraton i Jönköping 2017!

Tack **Hanna Odén Poulsen** och **Marie-Louise Lennartsson** för stöttning, hejarop och visdom!

Tack **Lena Svensson**, min kollega och vän från Linköping! Vad hade mina veckor på forskningslabbet i Paris varit utan dig? Ser fram emot att möta våren (2017) i Paris.

Tack **Sara**! Gospel och vetenskap förenade oss i en bil mellan Linköping och Jönköping. Tack **Vera**! Resan började på en busshållsplats (1999) och vi var båda lite nervösa inför uppropet på Hälsohögskolan. Vi skulle precis börja studera till Biomedicinsk analytiker.

Jag är otroligt tacksam för att jag får ha **er** i mitt liv! Ni har genom åren stöttat mig på alla sätt man kan önska sig. Ni har aldrig slutat tro på mig. TACK!

Olaf, tack för all choklad (och all stöttning!) som jag blivit tilldelad genom åren.

Tack till **"the girls at the lab"** – bland andra – Lisa, Marita, Levar, Renate, Hanna, Malin, Linda, Kiki, Emma och Sofie. Speciellt tack till **er** som faktiskt korrekturläste min avhandling och bidrog med värdefulla tankar och idéer. Tack för alla hejarop i både medgång och motgång. Tack för att ni delar med er av er kunskap. *Tillsammans är vi starka!*

Alla ni **vänner** utanför jobbet som ser mig för den jag alltid varit. Ni som tvingat mig att diskutera annat än vetenskap, *p*-värden och genetiska associationer. *Ingen nämnd, ingen glömd!*

Ett stor tack till mina föräldrar, **Ewa och Carl-Magnus**. Tack till syster **Chatarina** med familj. Tack för att **ni** alltid trott på mig. Tack för allt stöd och all support genom åren. Jag är er evigt tacksam! Jag älskar er!

Jag vill även rikta ett stort tack till mina svärföräldrar, **Lena** och **Hans**, för all hjälp och allt stöd med barnen. Hur hade vardagen fungerat utan er? Tack till min svägerska **Camilla** med familj. Tack för alla hejarop.

Roger, min älskade make och min allra käraste vän. Utan dig hade jag inte vågat ta ett enda steg. Jag lovar att jag framöver kommer vara mer närvarande än vad jag varit den senaste tiden. **Albin, Alfred** och **Elna**, våra älskade guldklimpar. Ni har visat mig att livet består av annat än jobb och forskning. Tack för att ni förgyller varje minut av mitt liv och tack för all er kärlek! *Jag älskar er fyra villkorslöst!*

This thesis was mainly supported by Futurum-the Academy for Healthcare, Region Jönköping County and FORSS, the Medical Research Council of South-Eastern Sweden.

References

Al-Sadi, R., Khatib, K., Guo, S., Ye, D., Youssef, M. & Ma, T. 2011. Occludin regulates macromolecule flux across the intestinal epithelial tight junction barrier. *Am J Physiol Gastrointest Liver Physiol,* 300**,** G1054-64.

Al-Sadi, R.M. & Ma, T.Y. 2007. IL-1beta causes an increase in intestinal epithelial tight junction permeability. *J Immunol,* 178**,** 4641-9.

Almer, S., Franzen, L., Olaison, G., Smedh, K. & Strom, M. 1993. Increased absorption of polyethylene glycol 600 deposited in the colon in active ulcerative colitis. *Gut,* 34**,** 509-13.

Ananthakrishnan, A.N. 2015. Environmental risk factors for inflammatory bowel diseases: a review. *Dig Dis Sci,* 60**,** 290-8.

Anderson, C.A., Boucher, G., Lees, C.W., Franke, A., D'amato, M., Taylor, K.D., Lee, J.C., Goyette, P., Imielinski, M., Latiano, A.*, et al.* 2011. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet,* 43**,** 246-52.

Ardlie, K.G., Kruglyak, L. & Seielstad, M. 2002. Patterns of linkage disequilibrium in the human genome. *Nat Rev Genet,* 3**,** 299-309.

Arimura, Y. & Yagi, J. 2010. Comprehensive expression profiles of genes for protein tyrosine phosphatases in immune cells. *Sci Signal,* 3**,** rs1.

Arpinati, M., Chirumbolo, G., Saunthararajah, Y., Stanzani, M., Bonifazi, F., Bandini, G., Baccarani, M. & Rondelli, D. 2007. Higher numbers of blood CD14+ cells before starting conditioning regimen correlate with greater risk of acute graft-versus-host disease in allogeneic stem cell transplantation from related donors. *Biol Blood Marrow Transplant,* 13**,** 228-34.

Bager, P., Simonsen, J., Nielsen, N.M. & Frisch, M. 2012. Cesarean section and offspring's risk of inflammatory bowel disease: a national cohort study. *Inflamm Bowel Dis,* 18**,** 857-62.

Ball, L.M., Egeler, R.M. & Party, E.P.W. 2008. Acute GvHD: pathogenesis and classification. *Bone Marrow Transplant,* 41 Suppl 2**,** S58-64.

Barmeyer, C., Erko, I., Fromm, A., Bojarski, C., Allers, K., Moos, V., Zeitz, M., Fromm, M. & Schulzke, J.D. 2012. Ion transport and barrier function are disturbed in microscopic colitis. *Ann N Y Acad Sci,* 1258**,** 143-8.

Baron, S., Turck, D., Leplat, C., Merle, V., Gower-Rousseau, C., Marti, R., Yzet, T., Lerebours, E., Dupas, J.L., Debeugny, S.*, et al.* 2005. Environmental risk factors in paediatric inflammatory bowel diseases: a population based case control study. *Gut,* 54**,** 357-63.

Barrett, J.C., Hansoul, S., Nicolae, D.L., Cho, J.H., Duerr, R.H., Rioux, J.D., Brant, S.R., Silverberg, M.S., Taylor, K.D., Barmada, M.M.*, et al.* 2008. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet,* 40**,** 955-62.

Bernstein, C.N., Banerjee, A., Targownik, L.E., Singh, H., Ghia, J.E., Burchill, C., Chateau, D. & Roos, L.L. 2016. Cesarean Section Delivery Is Not a Risk Factor for Development of Inflammatory Bowel Disease: A Population-based Analysis. *Clin Gastroenterol Hepatol,* 14**,** 50-7.

Bjarnason, I. 1994. Intestinal permeability. *Gut,* 35**,** S18-22.

Bonderup, O.K., Wigh, T., Nielsen, G.L., Pedersen, L. & Fenger-Gron, M. 2015. The epidemiology of microscopic colitis: a 10-year pathology-based nationwide Danish cohort study. *Scand J Gastroenterol,* 50**,** 393-8.

Bosi, E., Molteni, L., Radaelli, M.G., Folini, L., Fermo, I., Bazzigaluppi, E., Piemonti, L., Pastore, M.R. & Paroni, R. 2006. Increased intestinal permeability precedes clinical onset of type 1 diabetes. *Diabetologia,* 49**,** 2824-7.

Burgel, N., Bojarski, C., Mankertz, J., Zeitz, M., Fromm, M. & Schulzke, J.D. 2002. Mechanisms of diarrhea in collagenous colitis. *Gastroenterology,* 123**,** 433-43.

Burisch, J., Pedersen, N., Cukovic-Cavka, S., Brinar, M., Kaimakliotis, I., Duricova, D., Shonova, O., Vind, I., Avnstrom, S., Thorsgaard, N.*, et al.* 2014a. East-West gradient in the incidence of inflammatory bowel disease in Europe: the ECCO-EpiCom inception cohort. *Gut,* 63**,** 588-97.

Burisch, J., Pedersen, N., Cukovic-Cavka, S., Turk, N., Kaimakliotis, I., Duricova, D., Bortlik, M., Shonova, O., Vind, I., Avnstrom, S.*, et al.* 2014b. Environmental factors in a

population-based inception cohort of inflammatory bowel disease patients in Europe--an ECCO-EpiCom study. *J Crohns Colitis,* 8**,** 607-16.

Cardon, L.R. & Palmer, L.J. 2003. Population stratification and spurious allelic association. *Lancet,* 361**,** 598-604.

Carratu, R., Secondulfo, M., De Magistris, L., Iafusco, D., Urio, A., Carbone, M.G., Pontoni, G., Carteni, M. & Prisco, F. 1999. Altered intestinal permeability to mannitol in diabetes mellitus type I. *J Pediatr Gastroenterol Nutr,* 28**,** 264-9.

Cavanaugh, J. 2006. NOD2: ethnic and geographic differences. *World J Gastroenterol,* 12**,** 3673-7.

Cavet, J., Middleton, P.G., Segall, M., Noreen, H., Davies, S.M. & Dickinson, A.M. 1999. Recipient tumor necrosis factor-alpha and interleukin-10 gene polymorphisms associate with early mortality and acute graft-versus-host disease severity in HLA-matched sibling bone marrow transplants. *Blood,* 94**,** 3941-6.

Chuang, L.Y., Yang, C.H., Tsui, K.H., Cheng, Y.H., Chang, P.L., Wen, C.H. & Chang, H.W. 2008. Restriction enzyme mining for SNPs in genomes. *Anticancer Res,* 28**,** 2001-7.

Clarke, G.M., Anderson, C.A., Pettersson, F.H., Cardon, L.R., Morris, A.P. & Zondervan, K.T. 2011. Basic statistical analysis in genetic case-control studies. *Nat Protoc,* 6**,** 121-33.

Clayburgh, D.R., Shen, L. & Turner, J.R. 2004. A porous defense: the leaky epithelial barrier in intestinal disease. *Lab Invest,* 84**,** 282-91.

Cleynen, I., Boucher, G., Jostins, L., Schumm, L.P., Zeissig, S., Ahmad, T., Andersen, V., Andrews, J.M., Annese, V., Brand, S*., et al.* 2016. Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. *Lancet,* 387**,** 156-67.

Cleynen, I. & Vermeire, S. 2015. The genetic architecture of inflammatory bowel disease: past, present and future. *Curr Opin Gastroenterol,* 31**,** 456-63.

Cohen, J.L. & Boyer, O. 2006. The role of CD4+CD25hi regulatory T cells in the physiopathogeny of graft-versus-host disease. *Curr Opin Immunol,* 18**,** 580-5.

Colland, F., Jacq, X., Trouplin, V., Mougin, C., Groizeleau, C., Hamburger, A., Meil, A., Wojcik, J., Legrain, P. & Gauthier, J.M. 2004. Functional proteomics mapping of a human signaling pathway. *Genome Res,* 14**,** 1324-32.

Cooney, R., Cummings, J.R., Pathan, S., Beckly, J., Geremia, A., Hancock, L., Guo, C., Morris, A. & Jewell, D.P. 2009. Association between genetic variants in myosin IXB and Crohn's disease. *Inflamm Bowel Dis,* 15**,** 1014-21.

Corridoni, D., Arseneau, K.O. & Cominelli, F. 2014. Inflammatory bowel disease. *Immunol Lett,* 161**,** 231-5.

Cosnes, J., Gower-Rousseau, C., Seksik, P. & Cortot, A. 2011. Epidemiology and natural history of inflammatory bowel diseases. *Gastroenterology,* 140**,** 1785-94.

Crawford, D.C. & Nickerson, D.A. 2005. Definition and clinical importance of haplotypes. *Annu Rev Med,* 56**,** 303-20.

Crawford, J.M. 1999. *Robbins pathologic basis of disease*, W.B. Saunders Company.

Cui, W., Li, L.X., Sun, C.M., Wen, Y., Zhou, Y., Dong, Y.L. & Liu, P. 2010. Tumor necrosis factor alpha increases epithelial barrier permeability by disrupting tight junctions in Caco-2 cells. *Braz J Med Biol Res,* 43**,** 330-7.

Cuthbert, A.P., Fisher, S.A., Mirza, M.M., King, K., Hampe, J., Croucher, P.J., Mascheretti, S., Sanderson, J., Forbes, A., Mansfield, J.*, et al.* 2002. The contribution of NOD2 gene mutations to the risk and site of disease in inflammatory bowel disease. *Gastroenterology,* 122**,** 867-74.

Da Silveira Cruz-Machado, S., Carvalho-Sousa, C.E., Tamura, E.K., Pinato, L., Cecon, E., Fernandes, P.A., De Avellar, M.C., Ferreira, Z.S. & Markus, R.P. 2010. TLR4 and CD14 receptors expressed in rat pineal gland trigger NFKB pathway. *J Pineal Res,* 49**,** 183-92.

Dave, M., Papadakis, K.A. & Faubion, W.A., Jr. 2014. Immunology of inflammatory bowel disease and molecular targets for biologics. *Gastroenterol Clin North Am,* 43**,** 405-24.

Diaz-Gallo, L.M., Espino-Paisan, L., Fransen, K., Gomez-Garcia, M., Van Sommeren, S., Cardena, C., Rodrigo, L., Mendoza, J.L., Taxonera, C., Nieto, A.*, et al.* 2011. Differential association of two PTPN22 coding variants with Crohn's disease and ulcerative colitis. *Inflamm Bowel Dis,* 17**,** 2287-94.

Dickinson, A.M. & Charron, D. 2005. Non-HLA immunogenetics in hematopoietic stem cell transplantation. *Curr Opin Immunol,* 17**,** 517-25.

Dickinson, A.M. & Holler, E. 2008. Polymorphisms of cytokine and innate immunity genes and GVHD. *Best Pract Res Clin Haematol,* 21**,** 149-64.

Dudbridge, F. 2008. Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. *Hum Hered,* 66**,** 87-98.

Ek, W.E., D'amato, M. & Halfvarson, J. 2014. The history of genetics in inflammatory bowel disease. *Ann Gastroenterol,* 27**,** 294-303.

Ellinghaus, D., Jostins, L., Spain, S.L., Cortes, A., Bethune, J., Han, B., Park, Y.R., Raychaudhuri, S., Pouget, J.G., Hubenthal, M.*, et al.* 2016. Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. *Nat Genet,* 48**,** 510-8.

Elmaagacli, A.H., Koldehoff, M., Hindahl, H., Steckel, N.K., Trenschel, R., Peceny, R., Ottinger, H., Rath, P.M., Ross, R.S., Roggendorf, M.*, et al.* 2006. Mutations in innate immune system NOD2/CARD 15 and TLR-4 (Thr399Ile) genes influence the risk for severe acute graft-versus-host disease in patients who underwent an allogeneic transplantation. *Transplantation,* 81**,** 247-54.

Fanning, A.S., Jameson, B.J., Jesaitis, L.A. & Anderson, J.M. 1998. The tight junction protein ZO-1 establishes a link between the transmembrane protein occludin and the actin cytoskeleton. *J Biol Chem,* 273**,** 29745-53.

Farquhar, M.G. & Palade, G.E. 1963. Junctional complexes in various epithelia. *J Cell Biol,* 17**,** 375-412.

Fasseu, M., Treton, X., Guichard, C., Pedruzzi, E., Cazals-Hatem, D., Richard, C., Aparicio, T., Daniel, F., Soule, J.C., Moreau, R.*, et al.* 2010. Identification of restricted subsets of mature microRNA abnormally expressed in inactive colonic mucosa of patients with inflammatory bowel disease. *PLoS One,* 5.

Ferrara, J.L., Levine, J.E., Reddy, P. & Holler, E. 2009. Graft-versus-host disease. *Lancet,* 373**,** 1550-61.

Foersch, S., Waldner, M.J. & Neurath, M.F. 2013. Innate and adaptive immunity in inflammatory bowel diseases. *Dig Dis,* 31**,** 317-20.

Franke, A., Mcgovern, D.P., Barrett, J.C., Wang, K., Radford-Smith, G.L., Ahmad, T., Lees, C.W., Balschun, T., Lee, J., Roberts, R.*, et al.* 2010. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet,* 42**,** 1118-25.

Frolkis, A., Dieleman, L.A., Barkema, H.W., Panaccione, R., Ghosh, S., Fedorak, R.N., Madsen, K., Kaplan, G.G. & Alberta, I.B.D.C. 2013. Environment and the inflammatory bowel diseases. *Can J Gastroenterol,* 27**,** e18-24.

Frolova, L., Drastich, P., Rossmann, P., Klimesova, K. & Tlaskalova-Hogenova, H. 2008. Expression of Toll-like receptor 2 (TLR2), TLR4, and CD14 in biopsy samples of patients with inflammatory bowel diseases: upregulated expression of TLR2 in terminal ileum of patients with ulcerative colitis. *J Histochem Cytochem,* 56**,** 267-74.

Frueh, F.W. & Noyer-Weidner, M. 2003. The use of denaturing high-performance liquid chromatography (DHPLC) for the analysis of genetic variations: impact for diagnostics and pharmacogenetics. *Clin Chem Lab Med,* 41**,** 452-61.

Furuse, M., Fujita, K., Hiiragi, T., Fujimoto, K. & Tsukita, S. 1998. Claudin-1 and -2: novel integral membrane proteins localizing at tight junctions with no sequence similarity to occludin. *J Cell Biol,* 141**,** 1539-50.

Furuse, M., Furuse, K., Sasaki, H. & Tsukita, S. 2001. Conversion of zonulae occludentes from tight to leaky strand type by introducing claudin-2 into Madin-Darby canine kidney I cells. *J Cell Biol,* 153**,** 263-72.

Furuse, M., Hirase, T., Itoh, M., Nagafuchi, A., Yonemura, S., Tsukita, S. & Tsukita, S. 1993. Occludin: a novel integral membrane protein localizing at tight junctions. *J Cell Biol,* 123**,** 1777-88.

Geremia, A., Biancheri, P., Allan, P., Corazza, G.R. & Di Sabatino, A. 2014. Innate and adaptive immunity in inflammatory bowel disease. *Autoimmun Rev,* 13**,** 3-10.

Giri, A.K., Midha, S., Banerjee, P., Agrawal, A., Mehdi, S.J., Dhingra, R., Kaur, I., G, R.K., Lakhotia, R., Ghosh, S.*, et al.* 2016. Common Variants in CLDN2 and MORC4 Genes Confer Disease Susceptibility in Patients with Chronic Pancreatitis. *PLoS One,* 11**,** e0147345.

Granell, M., Urbano-Ispizua, A., Pons, A., Arostegui, J.I., Gel, B., Navarro, A., Jansa, S., Artells, R., Gaya, A., Talarn, C.*, et al.* 2008. Common variants in NLRP2 and NLRP3 genes are strong prognostic factors for the outcome of HLA-identical sibling allogeneic stem cell transplantation. *Blood,* 112**,** 4337-42.

Groschwitz, K.R. & Hogan, S.P. 2009. Intestinal barrier function: molecular regulation and disease pathogenesis. *J Allergy Clin Immunol,* 124**,** 3-20; quiz 21-2.

Guillemot, L., Paschoud, S., Pulimeno, P., Foglia, A. & Citi, S. 2008. The cytoplasmic plaque of tight junctions: a scaffolding and signalling center. *Biochim Biophys Acta,* 1778**,** 601-13.

Gullberg, E. & Soderholm, J.D. 2006. Peyer's patches and M cells as potential sites of the inflammatory onset in Crohn's disease. *Ann N Y Acad Sci,* 1072**,** 218-32.

Halfvarson, J., Bodin, L., Tysk, C., Lindberg, E. & Jarnerot, G. 2003. Inflammatory bowel disease in a Swedish twin cohort: a long-term follow-up of concordance and clinical characteristics. *Gastroenterology,* 124**,** 1767-73.

Halme, L., Paavola-Sakki, P., Turunen, U., Lappalainen, M., Farkkila, M. & Kontula, K. 2006. Family and twin studies in inflammatory bowel disease. *World J Gastroenterol,* 12**,** 3668-72.

He, Z., Gao, Y., Deng, Y., Li, W., Chen, Y., Xing, S., Zhao, X., Ding, J. & Wang, X. 2012. Lipopolysaccharide induces lung fibroblast proliferation through Toll-like receptor 4 signaling and the phosphoinositide3-kinase-Akt pathway. *PLoS One,* 7**,** e35926.

Helander, H.F. & Fandriks, L. 2014. Surface area of the digestive tract - revisited. *Scand J Gastroenterol,* 49**,** 681-9.

Henriksen, M., Jahnsen, J., Lygren, I., Sauar, J., Schulz, T., Stray, N., Vatn, M.H., Moum, B. & Ibsen Study, G. 2006. Change of diagnosis during the first five years after onset of inflammatory bowel disease: results of a prospective follow-up study (the IBSEN Study). *Scand J Gastroenterol,* 41**,** 1037-43.

Hilsden, R.J., Meddings, J.B. & Sutherland, L.R. 1996. Intestinal permeability changes in response to acetylsalicylic acid in relatives of patients with Crohn's disease. *Gastroenterology,* 110**,** 1395-403.

Hollander, D., Vadheim, C.M., Brettholz, E., Petersen, G.M., Delahunty, T. & Rotter, J.I. 1986. Increased intestinal permeability in patients with Crohn's disease and their relatives. A possible etiologic factor. *Ann Intern Med,* 105**,** 883-5.

Hong, E.P. & Park, J.W. 2012. Sample size and statistical power calculation in genetic association studies. *Genomics Inform,* 10**,** 117-22.

Howe, K.L., Reardon, C., Wang, A., Nazli, A. & Mckay, D.M. 2005. Transforming growth factor-beta regulation of epithelial tight junction proteins enhances barrier function and blocks enterohemorrhagic Escherichia coli O157:H7-induced increased permeability. *Am J Pathol,* 167**,** 1587-97.

Hu, Y.J., Wang, Y.D., Tan, F.Q. & Yang, W.X. 2013. Regulation of paracellular permeability: factors and mechanisms. *Mol Biol Rep,* 40**,** 6123-42.

Huang, X.L., Xu, J., Zhang, X.H., Qiu, B.Y., Peng, L., Zhang, M. & Gan, H.T. 2011. PI3K/Akt signaling pathway is involved in the pathogenesis of ulcerative colitis. *Inflamm Res,* 60**,** 727-34.

Huang, Y. & Chen, Z. 2016. Inflammatory bowel disease related innate immunity and adaptive immunity. *Am J Transl Res,* 8**,** 2490-7.

Hugot, J.P., Chamaillard, M., Zouali, H., Lesage, S., Cezard, J.P., Belaiche, J., Almer, S., Tysk, C., O'morain, C.A., Gassull, M.*, et al.* 2001. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature,* 411**,** 599-603.

Hymes, S.R., Alousi, A.M. & Cowen, E.W. 2012. Graft-versus-host disease: part I. Pathogenesis and clinical manifestations of graft-versus-host disease. *J Am Acad Dermatol,* 66**,** 515 e1-18; quiz 533-4.

Imhann, F., Vich Vila, A., Bonder, M.J., Fu, J., Gevers, D., Visschedijk, M.C., Spekhorst, L.M., Alberts, R., Franke, L., Van Dullemen, H.M.*, et al.* 2016. Interplay of host genetics and gut microbiota underlying the onset and clinical presentation of inflammatory bowel disease. *Gut***,** Doi: 10.1136/gutjnl-2016-312135.

Inai, T., Kobayashi, J. & Shibata, Y. 1999. Claudin-1 contributes to the epithelial barrier function in MDCK cells. *Eur J Cell Biol,* 78**,** 849-55.

Itoh, M., Nagafuchi, A., Moroi, S. & Tsukita, S. 1997. Involvement of ZO-1 in cadherin-based cell adhesion through its direct binding to alpha catenin and actin filaments. *J Cell Biol,* 138**,** 181-92.

Jarnerot, G., Hertervig, E., Granno, C., Thorhallsson, E., Eriksson, S., Tysk, C., Hansson, I., Bjorknas, H., Bohr, J., Olesen, M.*, et al.* 2001. Familial occurrence of microscopic colitis: a report on five families. *Scand J Gastroenterol,* 36**,** 959-62.

Jostins, L., Ripke, S., Weersma, R.K., Duerr, R.H., Mcgovern, D.P., Hui, K.Y., Lee, J.C., Schumm, L.P., Sharma, Y., Anderson, C.A.*, et al.* 2012. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature,* 491**,** 119-24.

Kabakchiev, B. & Silverberg, M.S. 2013. Expression quantitative trait loci analysis identifies associations between genotype and gene expression in human intestine. *Gastroenterology,* 144**,** 1488-96, 1496 e1-3.

Keerthivasan, S., Aghajani, K., Dose, M., Molinero, L., Khan, M.W., Venkateswaran, V., Weber, C., Emmanuel, A.O., Sun, T., Bentrem, D.J.*, et al.* 2014. beta-Catenin

promotes colitis and colon cancer through imprinting of proinflammatory properties in T cells. *Sci Transl Med,* 6**,** 225ra28.

Keita, A.V., Salim, S.Y., Jiang, T., Yang, P.C., Franzen, L., Soderkvist, P., Magnusson, K.E. & Soderholm, J.D. 2008. Increased uptake of non-pathogenic E. coli via the follicle-associated epithelium in longstanding ileal Crohn's disease. *J Pathol,* 215**,** 135-44.

Khatkar, M.S., Collins, A., Cavanagh, J.A., Hawken, R.J., Hobbs, M., Zenger, K.R., Barris, W., Mcclintock, A.E., Thomson, P.C., Nicholas, F.W.*, et al.* 2006. A first-generation metric linkage disequilibrium map of bovine chromosome 6. *Genetics,* 174**,** 79-85.

Kirsner, J.B. & Spencer, J.A. 1963. Family Occurrences of Ulcerative Colitis, Regional Enteritis, and Ileocolitis. *Ann Intern Med,* 59**,** 133-44.

Klement, E., Cohen, R.V., Boxman, J., Joseph, A. & Reif, S. 2004. Breastfeeding and risk of inflammatory bowel disease: a systematic review with meta-analysis. *Am J Clin Nutr,* 80**,** 1342-52.

Koskela, R.M., Karttunen, T.J., Niemela, S.E., Lehtola, J.K., Bloigu, R.S. & Karttunen, R.A. 2011. Cytokine gene polymorphism in microscopic colitis association with the IL-6-174 GG genotype. *Eur J Gastroenterol Hepatol,* 23**,** 607-13.

Koskela, R.M., Karttunen, T.J., Niemela, S.E., Lehtola, J.K., Ilonen, J. & Karttunen, R.A. 2008. Human leucocyte antigen and TNFalpha polymorphism association in microscopic colitis. *Eur J Gastroenterol Hepatol,* 20**,** 276-82.

Kostic, A.D., Xavier, R.J. & Gevers, D. 2014. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology,* 146**,** 1489-99.

Kotelevets, L., Van Hengel, J., Bruyneel, E., Mareel, M., Van Roy, F. & Chastre, E. 2005. Implication of the MAGI-1b/PTEN signalosome in stabilization of adherens junctions and suppression of invasiveness. *FASEB J,* 19**,** 115-7.

Krug, S.M., Schulzke, J.D. & Fromm, M. 2014. Tight junction, selective permeability, and related diseases. *Semin Cell Dev Biol,* 36**,** 166-76.

Kucharzik, T., Walsh, S.V., Chen, J., Parkos, C.A. & Nusrat, A. 2001. Neutrophil transmigration in inflammatory bowel disease is associated with differential expression of epithelial intercellular junction proteins. *Am J Pathol,* 159**,** 2001-9.

Landy, J., Ronde, E., English, N., Clark, S.K., Hart, A.L., Knight, S.C., Ciclitira, P.J. & Al-Hassi, H.O. 2016. Tight junctions in inflammatory bowel diseases and inflammatory bowel disease associated colorectal cancer. *World J Gastroenterol,* 22**,** 3117-26.

Langner, C., Aust, D., Ensari, A., Villanacci, V., Becheanu, G., Miehlke, S., Geboes, K., Munch, A., Working Group of Digestive Diseases of the European Society Of, P. & The European Microscopic Colitis, G. 2015. Histology of microscopic colitis-review with a practical approach for pathologists. *Histopathology,* 66**,** 613-26.

Laukoetter, M.G., Nava, P., Lee, W.Y., Severson, E.A., Capaldo, C.T., Babbin, B.A., Williams, I.R., Koval, M., Peatman, E., Campbell, J.A.*, et al.* 2007. JAM-A regulates permeability and inflammation in the intestine in vivo. *J Exp Med,* 204**,** 3067-76.

Lee, S.H. 2015. Intestinal permeability regulation by tight junction: implication on inflammatory bowel diseases. *Intest Res,* 13**,** 11-8.

Lennard-Jones, J.E. 1989. Classification of inflammatory bowel disease. *Scand J Gastroenterol Suppl,* 170**,** 2-6; discussion 16-9.

Lewis, C.M. 2002. Genetic association studies: design, analysis and interpretation. *Brief Bioinform,* 3**,** 146-53.

Li, P., Yang, X.K., Wang, X., Zhao, M.Q., Zhang, C., Tao, S.S., Zhao, W., Huang, Q., Li, L.J., Pan, H.F.*, et al.* 2016. A meta-analysis of the relationship between MYO9B gene polymorphisms and susceptibility to Crohn's disease and ulcerative colitis. *Hum Immunol,* 77**,** 990-6.

Liang, G.H. & Weber, C.R. 2014. Molecular aspects of tight junction barrier function. *Curr Opin Pharmacol,* 19**,** 84-9.

Lin, M.T., Storer, B., Martin, P.J., Tseng, L.H., Gooley, T., Chen, P.J. & Hansen, J.A. 2003. Relation of an interleukin-10 promoter polymorphism to graft-versus-host disease and survival after hematopoietic-cell transplantation. *N Engl J Med,* 349**,** 2201-10.

Lindgren, S. & Löfberg, R. 2011. *Gastroenterologi och hepatologi*, Författarna och Liber AB.

Liu, J.Z., Van Sommeren, S., Huang, H., Ng, S.C., Alberts, R., Takahashi, A., Ripke, S., Lee, J.C., Jostins, L., Shah, T.*, et al.* 2015. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet,* 47**,** 979-86.

Liu, T.C. & Stappenbeck, T.S. 2016. Genetics and Pathogenesis of Inflammatory Bowel Disease. *Annu Rev Pathol,* 11**,** 127-48.

Livak, K.J. & Schmittgen, T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods,* 25**,** 402-8.

Loddo, I. & Romano, C. 2015. Inflammatory Bowel Disease: Genetics, Epigenetics, and Pathogenesis. *Front Immunol,* 6**,** 551.

Long, S.H., He, Y., Chen, M.H., Cao, K., Chen, Y.J., Chen, B.L., Mao, R., Zhang, S.H., Zhu, Z.H., Zeng, Z.R.*, et al.* 2013. Activation of PI3K/Akt/mTOR signaling pathway triggered by PTEN downregulation in the pathogenesis of Crohn's disease. *J Dig Dis,* 14**,** 662-9.

Luissint, A.C., Nusrat, A. & Parkos, C.A. 2014. JAM-related proteins in mucosal homeostasis and inflammation. *Semin Immunopathol,* 36**,** 211-26.

Luissint, A.C., Parkos, C.A. & Nusrat, A. 2016. Inflammation and the Intestinal Barrier: Leukocyte-Epithelial Cell Interactions, Cell Junction Remodeling, and Mucosal Repair. *Gastroenterology,* 151**,** 616-32.

Ma, Q., Yang, Y., Feng, D., Zheng, S., Meng, R., Fa, P., Zhao, C., Liu, H., Song, R., Tao, T.*, et al.* 2015a. MAGI3 negatively regulates Wnt/beta-catenin signaling and suppresses malignant phenotypes of glioma cells. *Oncotarget,* 6**,** 35851-65.

Ma, Q., Zhang, Y., Meng, R., Xie, K.M., Xiong, Y., Lin, S., He, Z.L., Tao, T., Yang, Y., Zhao, J.Z.*, et al.* 2015b. MAGI3 Suppresses Glioma Cell Proliferation via Upregulation of PTEN Expression. *Biomed Environ Sci,* 28**,** 502-9.

Madisch, A., Hellmig, S., Schreiber, S., Bethke, B., Stolte, M. & Miehlke, S. 2011. Allelic variation of the matrix metalloproteinase-9 gene is associated with collagenous colitis. *Inflamm Bowel Dis,* 17**,** 2295-8.

Maehama, T. & Dixon, J.E. 1998. The tumor suppressor, PTEN/MMAC1, dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate. *J Biol Chem,* 273**,** 13375-8.

Magro, F., Langner, C., Driessen, A., Ensari, A., Geboes, K., Mantzaris, G.J., Villanacci, V., Becheanu, G., Borralho Nunes, P., Cathomas, G.*, et al.* 2013. European consensus on the histopathology of inflammatory bowel disease. *J Crohns Colitis,* 7**,** 827-51.

Mahid, S.S., Minor, K.S., Soto, R.E., Hornung, C.A. & Galandiuk, S. 2006. Smoking and inflammatory bowel disease: a meta-analysis. *Mayo Clin Proc,* 81**,** 1462-71.

Martin-Padura, I., Lostaglio, S., Schneemann, M., Williams, L., Romano, M., Fruscella, P., Panzeri, C., Stoppacciaro, A., Ruco, L., Villa, A.*, et al.* 1998. Junctional adhesion molecule, a novel member of the immunoglobulin superfamily that distributes at intercellular junctions and modulates monocyte transmigration. *J Cell Biol,* 142**,** 117-27.

Mccarthy, K.M., Skare, I.B., Stankewich, M.C., Furuse, M., Tsukita, S., Rogers, R.A., Lynch, R.D. & Schneeberger, E.E. 1996. Occludin is a functional component of the tight junction. *J Cell Sci,* 109 (Pt 9)**,** 2287-98.

Mcgovern, D.P., Gardet, A., Torkvist, L., Goyette, P., Essers, J., Taylor, K.D., Neale, B.M., Ong, R.T., Lagace, C., Li, C.*, et al.* 2010. Genome-wide association identifies multiple ulcerative colitis susceptibility loci. *Nat Genet,* 42**,** 332-7.

Mcgovern, D.P., Taylor, K.D., Landers, C., Derkowski, C., Dutridge, D., Dubinsky, M., Ippoliti, A., Vasiliauskas, E., Mei, L., Mengesha, E.*, et al.* 2009. MAGI2 genetic variation and inflammatory bowel disease. *Inflamm Bowel Dis,* 15**,** 75-83.

Mellander, M.R., Ekbom, A., Hultcrantz, R., Lofberg, R., Ost, A. & Bjork, J. 2016. Microscopic colitis: a descriptive clinical cohort study of 795 patients with collagenous and lymphocytic colitis. *Scand J Gastroenterol,* 51**,** 556-62.

Menard, S., Cerf-Bensussan, N. & Heyman, M. 2010. Multiple facets of intestinal permeability and epithelial handling of dietary antigens. *Mucosal Immunol,* 3**,** 247-59.

Mersha, T.B., Ding, L., He, H., Alexander, E.S., Zhang, X., Kurowski, B.G., Pilipenko, V., Kottyan, L., Martin, L.J. & Fardo, D.W. 2015. Impact of Population Stratification on Family-Based Association in an Admixed Population. *Int J Genomics,* 2015**,** 501617.

Mineta, K., Yamamoto, Y., Yamazaki, Y., Tanaka, H., Tada, Y., Saito, K., Tamura, A., Igarashi, M., Endo, T., Takeuchi, K.*, et al.* 2011. Predicted expansion of the claudin multigene family. *FEBS Lett,* 585**,** 606-12.

Molodecky, N.A., Soon, I.S., Rabi, D.M., Ghali, W.A., Ferris, M., Chernoff, G., Benchimol, E.I., Panaccione, R., Ghosh, S., Barkema, H.W.*, et al.* 2012. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology,* 142**,** 46-54 e42; quiz e30.

Moum, B., Vatn, M.H., Ekbom, A., Aadland, E., Fausa, O., Lygren, I., Sauar, J., Schulz, T. & Stray, N. 1996a. Incidence of ulcerative colitis and indeterminate colitis in four counties of southeastern Norway, 1990-93. A prospective population-based study. The Inflammatory Bowel South-Eastern Norway (IBSEN) Study Group of Gastroenterologists. *Scand J Gastroenterol,* 31**,** 362-6.

Moum, B., Vatn, M.H., Ekbom, A., Aadland, E., Fausa, O., Lygren, I., Stray, N., Sauar, J. & Schulz, T. 1996b. Incidence of Crohn's disease in four counties in southeastern

Norway, 1990-93. A prospective population-based study. The Inflammatory Bowel South-Eastern Norway (IBSEN) Study Group of Gastroenterologists. *Scand J Gastroenterol,* 31**,** 355-61.

Mowat, A.M. & Agace, W.W. 2014. Regional specialization within the intestinal immune system. *Nat Rev Immunol,* 14**,** 667-85.

Munch, A. & Langner, C. 2015. Microscopic colitis: clinical and pathologic perspectives. *Clin Gastroenterol Hepatol,* 13**,** 228-36.

Munch, A., Soderholm, J.D., Wallon, C., Ost, A., Olaison, G. & Strom, M. 2005. Dynamics of mucosal permeability and inflammation in collagenous colitis before, during, and after loop ileostomy. *Gut,* 54**,** 1126-8.

Myren, J., Gjone, E., Hertzberg, J.N., Rygvold, O., Semb, L.S. & Fretheim, B. 1971. Epidemiology of ulcerative colitis and regional enterocolitis (Crohn's disease) in Norway. *Scand J Gastroenterol,* 6**,** 511-4.

Nordenvall, B., Brostrom, O., Berglund, M., Monsen, U., Nordenstrom, J., Sorstad, J. & Hellers, G. 1985. Incidence of ulcerative colitis in Stockholm County 1955-1979. *Scand J Gastroenterol,* 20**,** 783-90.

Noth, R., Lange-Grumfeld, J., Stuber, E., Kruse, M.L., Ellrichmann, M., Hasler, R., Hampe, J., Bewig, B., Rosenstiel, P., Schreiber, S.*, et al.* 2011. Increased intestinal permeability and tight junction disruption by altered expression and localization of occludin in a murine graft versus host disease model. *BMC Gastroenterol,* 11**,** 109.

Ogura, Y., Bonen, D.K., Inohara, N., Nicolae, D.L., Chen, F.F., Ramos, R., Britton, H., Moran, T., Karaliuskas, R., Duerr, R.H.*, et al.* 2001. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature,* 411**,** 603-6.

Olesen, M., Eriksson, S., Bohr, J., Jarnerot, G. & Tysk, C. 2004. Microscopic colitis: a common diarrhoeal disease. An epidemiological study in Orebro, Sweden, 1993-1998. *Gut,* 53**,** 346-50.

Orholm, M., Binder, V., Sorensen, T.I., Rasmussen, L.P. & Kyvik, K.O. 2000. Concordance of inflammatory bowel disease among Danish twins. Results of a nationwide study. *Scand J Gastroenterol,* 35**,** 1075-81.

Orholm, M., Munkholm, P., Langholz, E., Nielsen, O.H., Sorensen, T.I. & Binder, V. 1991. Familial occurrence of inflammatory bowel disease. *N Engl J Med,* 324**,** 84-8.

Peeters, M., Geypens, B., Claus, D., Nevens, H., Ghoos, Y., Verbeke, G., Baert, F., Vermeire, S., Vlietinck, R. & Rutgeerts, P. 1997. Clustering of increased small intestinal permeability in families with Crohn's disease. *Gastroenterology,* 113**,** 802-7.

Peloquin, J.M., Goel, G., Kong, L., Huang, H., Haritunians, T., Sartor, R.B., Daly, M.J., Newberry, R.D., Mcgovern, D.P., Yajnik, V.*, et al.* 2016. Characterization of candidate genes in inflammatory bowel disease-associated risk loci. *JCI Insight,* 1**,** e87899.

Peterson, C.T., Sharma, V., Elmen, L. & Peterson, S.N. 2015. Immune homeostasis, dysbiosis and therapeutic modulation of the gut microbiota. *Clin Exp Immunol,* 179**,** 363-77.

Pettersson, F.H., Anderson, C.A., Clarke, G.M., Barrett, J.C., Cardon, L.R., Morris, A.P. & Zondervan, K.T. 2009. Marker selection for genetic case-control association studies. *Nat Protoc,* 4**,** 743-52.

Pisani, L.F., Tontini, G.E., Vecchi, M. & Pastorelli, L. 2016. Microscopic Colitis: What Do We Know About Pathogenesis? *Inflamm Bowel Dis,* 22**,** 450-8.

Planchon, S., Fiocchi, C., Takafuji, V. & Roche, J.K. 1999. Transforming growth factor-beta1 preserves epithelial barrier function: identification of receptors, biochemical intermediates, and cytokine antagonists. *J Cell Physiol,* 181**,** 55-66.

Planell, N., Lozano, J.J., Mora-Buch, R., Masamunt, M.C., Jimeno, M., Ordas, I., Esteller, M., Ricart, E., Pique, J.M., Panes, J.*, et al.* 2013. Transcriptional analysis of the intestinal mucosa of patients with ulcerative colitis in remission reveals lasting epithelial cell alterations. *Gut,* 62**,** 967-76.

Poritz, L.S., Harris, L.R., 3rd, Kelly, A.A. & Koltun, W.A. 2011. Increase in the tight junction protein claudin-1 in intestinal inflammation. *Dig Dis Sci,* 56**,** 2802-9.

Prasad, S., Mingrino, R., Kaukinen, K., Hayes, K.L., Powell, R.M., Macdonald, T.T. & Collins, J.E. 2005. Inflammatory processes have differential effects on claudins 2, 3 and 4 in colonic epithelial cells. *Lab Invest,* 85**,** 1139-62.

Premstaller, A. & Oefner, P.J. 2003. Denaturing high-performance liquid chromatography. *Methods Mol Biol,* 212**,** 15-35.

Pritchard, J.K. & Przeworski, M. 2001. Linkage disequilibrium in humans: models and data. *Am J Hum Genet,* 69**,** 1-14.

Probert, C.S., Jayanthi, V., Hughes, A.O., Thompson, J.R., Wicks, A.C. & Mayberry, J.F. 1993. Prevalence and family risk of ulcerative colitis and Crohn's disease: an

epidemiological study among Europeans and south Asians in Leicestershire. *Gut,* 34**,** 1547-51.

Rahner, C., Mitic, L.L. & Anderson, J.M. 2001. Heterogeneity in expression and subcellular localization of claudins 2, 3, 4, and 5 in the rat liver, pancreas, and gut. *Gastroenterology,* 120**,** 411-22.

Rivas, M.A., Beaudoin, M., Gardet, A., Stevens, C., Sharma, Y., Zhang, C.K., Boucher, G., Ripke, S., Ellinghaus, D., Burtt, N.*, et al.* 2011. Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. *Nat Genet,* 43**,** 1066-73.

Sairafi, D., Uzunel, M., Remberger, M., Ringden, O. & Mattsson, J. 2008. No impact of NOD2/CARD15 on outcome after SCT. *Bone Marrow Transplant,* 41**,** 961-4.

Sakaguchi, T., Gu, X., Golden, H.M., Suh, E., Rhoads, D.B. & Reinecker, H.C. 2002. Cloning of the human claudin-2 5'-flanking region revealed a TATA-less promoter with conserved binding sites in mouse and human for caudal-related homeodomain proteins and hepatocyte nuclear factor-1alpha. *J Biol Chem,* 277**,** 21361-70.

Salim, S.Y. & Soderholm, J.D. 2011. Importance of disrupted intestinal barrier in inflammatory bowel diseases. *Inflamm Bowel Dis,* 17**,** 362-81.

Sauer, S., Bruno, L., Hertweck, A., Finlay, D., Leleu, M., Spivakov, M., Knight, Z.A., Cobb, B.S., Cantrell, D., O'connor, E.*, et al.* 2008. T cell receptor signaling controls Foxp3 expression via PI3K, Akt, and mTOR. *Proc Natl Acad Sci U S A,* 105**,** 7797-802.

Sawada, N. 2013. Tight junction-related human diseases. *Pathol Int,* 63**,** 1-12.

Schmitz, H., Barmeyer, C., Fromm, M., Runkel, N., Foss, H.D., Bentzel, C.J., Riecken, E.O. & Schulzke, J.D. 1999. Altered tight junction structure contributes to the impaired epithelial barrier function in ulcerative colitis. *Gastroenterology,* 116**,** 301-9.

Secondulfo, M., De Magistris, L., Fiandra, R., Caserta, L., Belletta, M., Tartaglione, M.T., Riegler, G., Biagi, F., Corazza, G.R. & Carratu, R. 2001. Intestinal permeability in Crohn's disease patients and their first degree relatives. *Dig Liver Dis,* 33**,** 680-5.

Sharp, R.C., Abdulrahim, M., Naser, E.S. & Naser, S.A. 2015. Genetic Variations of PTPN2 and PTPN22: Role in the Pathogenesis of Type 1 Diabetes and Crohn's Disease. *Front Cell Infect Microbiol,* 5**,** 95.

Shivananda, S., Lennard-Jones, J., Logan, R., Fear, N., Price, A., Carpenter, L. & Van Blankenstein, M. 1996. Incidence of inflammatory bowel disease across Europe:

is there a difference between north and south? Results of the European Collaborative Study on Inflammatory Bowel Disease (EC-IBD). *Gut,* 39**,** 690-7.

Sivula, J., Cordova, Z.M., Tuimala, J., Jaatinen, T., Partanen, J., Volin, L. & Turpeinen, H. 2012. Toll-like receptor gene polymorphisms confer susceptibility to graft-versus-host disease in allogenic hematopoietic stem cell transplantation. *Scand J Immunol,* 76**,** 336-41.

Soderholm, J.D., Olaison, G., Lindberg, E., Hannestad, U., Vindels, A., Tysk, C., Jarnerot, G. & Sjodahl, R. 1999. Different intestinal permeability patterns in relatives and spouses of patients with Crohn's disease: an inherited defect in mucosal defence? *Gut,* 44**,** 96-100.

Stappenbeck, T.S. & Mcgovern, D.P. 2016. Paneth Cell Alterations in the Development and Phenotype of Crohn's Disease. *Gastroenterology***,** Doi: 10.1053/j.gastro.2016.10.003.

Stewart, M., Andrews, C.N., Urbanski, S., Beck, P.L. & Storr, M. 2011. The association of coeliac disease and microscopic colitis: a large population-based study. *Aliment Pharmacol Ther,* 33**,** 1340-9.

Stevenson, B.R., Siliciano, J.D., Mooseker, M.S. & Goodenough, D.A. 1986. Identification of ZO-1: a high molecular weight polypeptide associated with the tight junction (zonula occludens) in a variety of epithelia. *J Cell Biol,* 103**,** 755-66.

Storr, M.A. 2013. Microscopic colitis: epidemiology, pathophysiology, diagnosis and current management-an update 2013. *ISRN Gastroenterol,* 2013**,** 352718.

String Consortium. STRING [Online]. Available: http://string-db.org/ [Accessed 16 August 2016].

Suenaert, P., Maerten, P., Van Assche, G., Van Driessche, W., Geboes, K., Bulteel, V., Simaels, J., Augustijns, P., Ceuppens, J.L., Rutgeerts, P.*, et al.* 2010. Effects of T cell-induced colonic inflammation on epithelial barrier function. *Inflamm Bowel Dis,* 16**,** 1322-31.

Sunyaev, S., Ramensky, V., Koch, I., Lathe, W., 3rd, Kondrashov, A.S. & Bork, P. 2001. Prediction of deleterious human alleles. *Hum Mol Genet,* 10**,** 591-7.

Suzuki, T. 2013. Regulation of intestinal epithelial permeability by tight junctions. *Cell Mol Life Sci,* 70**,** 631-59.

Syvanen, A.C. 1999. From gels to chips: "minisequencing" primer extension for analysis of point mutations and single nucleotide polymorphisms. *Hum Mutat,* 13**,** 1-10.

Söderman, J., Berglind, L. & Almer, S. 2015. Gene Expression-Genotype Analysis Implicates GSDMA, GSDMB, and LRRC3C as Contributors to Inflammatory Bowel Disease Susceptibility. *Biomed Res Int,* 2015**,** 834805.

Tokuhira, N., Kitagishi, Y., Suzuki, M., Minami, A., Nakanishi, A., Ono, Y., Kobayashi, K., Matsuda, S. & Ogura, Y. 2015. PI3K/AKT/PTEN pathway as a target for Crohn's disease therapy (Review). *Int J Mol Med,* 35**,** 10-6.

Turner, J.R., Rill, B.K., Carlson, S.L., Carnes, D., Kerner, R., Mrsny, R.J. & Madara, J.L. 1997. Physiological regulation of epithelial tight junctions is associated with myosin light-chain phosphorylation. *Am J Physiol,* 273**,** C1378-85.

Tysk, C., Lindberg, E., Jarnerot, G. & Floderus-Myrhed, B. 1988. Ulcerative colitis and Crohn's disease in an unselected population of monozygotic and dizygotic twins. A study of heritability and the influence of smoking. *Gut,* 29**,** 990-6.

Umeda, K., Ikenouchi, J., Katahira-Tayama, S., Furuse, K., Sasaki, H., Nakayama, M., Matsui, T., Tsukita, S., Furuse, M. & Tsukita, S. 2006. ZO-1 and ZO-2 independently determine where claudins are polymerized in tight-junction strand formation. *Cell,* 126**,** 741-54.

Wall, J.D. & Pritchard, J.K. 2003. Haplotype blocks and linkage disequilibrium in the human genome. *Nat Rev Genet,* 4**,** 587-97.

Van Elburg, R.M., Uil, J.J., Mulder, C.J. & Heymans, H.S. 1993. Intestinal permeability in patients with coeliac disease and relatives of patients with coeliac disease. *Gut,* 34**,** 354-7.

Van Itallie, C.M., Fanning, A.S. & Anderson, J.M. 2003. Reversal of charge selectivity in cation or anion-selective epithelial lines by expression of different claudins. *Am J Physiol Renal Physiol,* 285**,** F1078-84.

Van Itallie, C.M., Holmes, J., Bridges, A., Gookin, J.L., Coccaro, M.R., Proctor, W., Colegio, O.R. & Anderson, J.M. 2008. The density of small tight junction pores varies among cell types and is increased by expression of claudin-2. *J Cell Sci,* 121**,** 298-305.

Wapenaar, M.C., Monsuur, A.J., Van Bodegraven, A.A., Weersma, R.K., Bevova, M.R., Linskens, R.K., Howdle, P., Holmes, G., Mulder, C.J., Dijkstra, G.*, et al.* 2008. Associations with tight junction genes PARD3 and MAGI2 in Dutch patients point to a common barrier defect for coeliac disease and ulcerative colitis. *Gut,* 57**,** 463-7.

Vatn, M.H. & Sandvik, A.K. 2015. Inflammatory bowel disease. *Scand J Gastroenterol,* 50**,** 748-62.

Weber, C.R., Nalle, S.C., Tretiakova, M., Rubin, D.T. & Turner, J.R. 2008. Claudin-1 and claudin-2 expression is elevated in inflammatory bowel disease and may contribute to early neoplastic transformation. *Lab Invest,* 88**,** 1110-20.

Vegh, Z., Burisch, J., Pedersen, N., Kaimakliotis, I., Duricova, D., Bortlik, M., Avnstrom, S., Vinding, K.K., Olsen, J., Nielsen, K.R.*, et al.* 2014. Incidence and initial disease course of inflammatory bowel diseases in 2011 in Europe and Australia: results of the 2011 ECCO-EpiCom inception cohort. *J Crohns Colitis,* 8**,** 1506-15.

Wehkamp, J., Harder, J., Weichenthal, M., Schwab, M., Schaffeler, E., Schlee, M., Herrlinger, K.R., Stallmach, A., Noack, F., Fritz, P.*, et al.* 2004. NOD2 (CARD15) mutations in Crohn's disease are associated with diminished mucosal alpha-defensin expression. *Gut,* 53**,** 1658-64.

Wehkamp, J., Salzman, N.H., Porter, E., Nuding, S., Weichenthal, M., Petras, R.E., Shen, B., Schaeffeler, E., Schwab, M., Linzmeier, R.*, et al.* 2005. Reduced Paneth cell alpha-defensins in ileal Crohn's disease. *Proc Natl Acad Sci U S A,* 102**,** 18129-34.

Wehkamp, J. & Stange, E.F. 2010. Paneth's disease. *J Crohns Colitis,* 4**,** 523-31.

Weisdorf, D. 2007. GVHD the nuts and bolts. *Hematology Am Soc Hematol Educ Program***,** 62-7.

Westerlind, H., Mellander, M.R., Bresso, F., Munch, A., Bonfiglio, F., Assadi, G., Rafter, J., Hubenthal, M., Lieb, W., Kallberg, H.*, et al.* 2015. Dense genotyping of immune-related loci identifies HLA variants associated with increased risk of collagenous colitis. *Gut*.

Vetrano, S. & Danese, S. 2009. The role of JAM-A in inflammatory bowel disease: unrevealing the ties that bind. *Ann N Y Acad Sci,* 1165**,** 308-13.

Vetrano, S., Rescigno, M., Cera, M.R., Correale, C., Rumio, C., Doni, A., Fantini, M., Sturm, A., Borroni, E., Repici, A.*, et al.* 2008. Unique role of junctional adhesion molecule-a in maintaining mucosal homeostasis in inflammatory bowel disease. *Gastroenterology,* 135**,** 173-84.

Whitcomb, D.C., Larusch, J., Krasinskas, A.M., Klei, L., Smith, J.P., Brand, R.E., Neoptolemos, J.P., Lerch, M.M., Tector, M., Sandhu, B.S.*, et al.* 2012. Common genetic variants in the CLDN2 and PRSS1-PRSS2 loci alter risk for alcohol-related and sporadic pancreatitis. *Nat Genet,* 44**,** 1349-54.

Wilkening, S., Chen, B., Bermejo, J.L. & Canzian, F. 2009. Is there still a need for candidate gene approaches in the era of genome-wide association studies? *Genomics,* 93**,** 415-9.

Wisner, D.M., Harris, L.R., 3rd, Green, C.L. & Poritz, L.S. 2008. Opposing regulation of the tight junction protein claudin-2 by interferon-gamma and interleukin-4. *J Surg Res,* 144**,** 1-7.

Wray, N.R. 2005. Allele frequencies and the r2 measure of linkage disequilibrium: impact on design and interpretation of association studies. *Twin Res Hum Genet,* 8**,** 87-94.

Www.Ncbi.Nlm.Nih.Gov/Projects/Snp. US National Library of Medicine National Institutes of Health [Online]. Available: www.ncbi.nlm.nih.gov/projects/SNP [Accessed 30 September 2016].

Yamamoto-Furusho, J.K., Mendivil, E.J. & Fonseca-Camarillo, G. 2012. Differential expression of occludin in patients with ulcerative colitis and healthy controls. *Inflamm Bowel Dis,* 18**,** E1999.

Yamazaki, K., Mcgovern, D., Ragoussis, J., Paolucci, M., Butler, H., Jewell, D., Cardon, L., Takazoe, M., Tanaka, T., Ichimori, T.*, et al.* 2005. Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease. *Hum Mol Genet,* 14**,** 3499-506.

Zanello, G., Kevans, D., Goethel, A., Silverberg, M., Tyler, A. & Croitoru, K. 2014. Genetics and innate and adaptive immunity in IBD. *Nestle Nutr Inst Workshop Ser,* 79**,** 41-55.

Zeissig, S., Burgel, N., Gunzel, D., Richter, J., Mankertz, J., Wahnschaffe, U., Kroesen, A.J., Zeitz, M., Fromm, M. & Schulzke, J.D. 2007. Changes in expression and distribution of claudin 2, 5 and 8 lead to discontinuous tight junctions and barrier dysfunction in active Crohn's disease. *Gut,* 56**,** 61-72.

Zihni, C., Mills, C., Matter, K. & Balda, M.S. 2016. Tight junctions: from simple barriers to multifunctional molecular gates. *Nat Rev Mol Cell Biol,* 17**,** 564-80.

Zondervan, K.T. & Cardon, L.R. 2007. Designing candidate gene and genome-wide case-control association studies. *Nat Protoc,* 2**,** 2492-501.

Zundler, S. & Neurath, M.F. 2015. Immunopathogenesis of inflammatory bowel diseases: functional role of T cells and T cell homing. *Clin Exp Rheumatol,* 33**,** S19-28.