Clinical and molecular effects of guanylate cyclase C-activation

Rune Rose Tronstad

Thesis for the degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2019

UNIVERSITY OF BERGEN

Clinical and molecular effects of guanylate cyclase C-activation

Rune Rose Tronstad

Thesis for the degree of Philosophiae Doctor (PhD)
at the University of Bergen

Date of defense: 06.06.2019

Year: 2019

Title: Clinical and molecular effects of guanylate cyclase C-activation

Name: Rune Rose Tronstad

Print: Skipnes Kommunikasjon / University of Bergen

Contents

Scientific environment

This project started out as a co-operation between clinicians and geneticists at the University of Bergen, and during its course it has involved other national and international research groups, above all the Norwegian Centre for Primary Sclerosing Cholangitis (NoPSC), Oslo University Hospital, Rikshospitalet lead by professor Tom Hemming Karlsen.

Patients with Familial *GUCY2C* diarrhoea syndrome were followed up at the Department of Paediatrics, the Department of Gastroenterology and at the Department of Medical Genetics, Haukeland University Hospital. Much of the genomic investigations depended on the research infrastructure provided by the Genomics Core Facility at the University of Bergen.

During the project I twice a year attended the NoPSC guest professor meetings, where I presented, discussed and got advice on my research with world leading experts on microbiota and genetics in inflammatory bowel disease (IBD). These meetings provided the basis for our co-operation with the Institute of Clinical Molecular Biology, University of Kiel, Germany, where genotyping of IBD genetic risk variants was performed.

The project also depended on co-operation with the Indian Institute of Science, under prof. S. Visweswariah, Bangalore, India, which contributed with functional studies of the *GUCY2C* mutation in a cell-line.

Acknowledgements

First and foremost I thank my main supervisor Torunn Fiskerstrand for giving me the opportunity to do this exciting work, for guiding me through and for sharing your immense knowledge and passion for science. Sadly you passed away on the first day of the New Year. I will profoundly miss you, your clear opinions and our inspiring discussions. Thanks for making the courageous choice of being my supervisor and for the huge amount of work and time spent in the process.

Next my thanks go to Tom Hemming Karlsen, my co-supervisor and dear friend. Ever since our time as students I have enjoyed your unlimited generosity and marvelled at the enthusiasm and dedication you display when exploring new concepts. Thank you for inviting me to take part in the research environment of NoPSC, and for helping me through this project.

I am thankful to our patients and the healthy controls for their participation.

Thanks to my co-supervisor Per Morten Knappskog who played a crucial role in planning, co-ordination and facilitation of the research projects.

Stephanie Le Hellard, also co-supervisor guided me with lots of humour through some of the most difficult aspects of this thesis.

I thank Hans-Richard Brattbakk, Christine Stansberg, Tatiana Polushina and Kari Merete Ersland at the Genomics Core Facility for your work and your patient efforts to enlighten me.

I am deeply thankful to our co-workers at the NoPSC in Oslo for your huge contribution to this project.

Especially I would like to thank Johannes Hov for tuition on the microbiota, scientific writing and the thoughtful way you supervised the creation of paper III. Martin Kummen provided crucial aid on paper III from the lab to the writing process. Your thesis on PSC and the microbiota was of great help and inspiration through this project. My gratitude goes also to their friendly colleagues Kristian Holm who did much of the bioinformatics and Silje Fjellsgård Jørgensen who assisted on the gene expression study. Thanks also to Hege Dahlen Sollid and Mona Bjørnstad for help with coordination of the study.

I thank the wonderful staff of the Department of Medical Genetics for all your help. My special gratitude goes to Hilde Rusaas who took care of much of the laboratory work, always kindly and always efficiently, and to Kjersti Alfheim Boge for keeping track e.g. of my belongings.

Many thanks go to Hilde Løland von Volkmann, a phenomenal adult gastroenterologist, for your co-operation, for sharing my interest in *GUCY2C,* for lending your listening ear and for our many enthusiastic discussions. My gratitude extends to your colleagues Khanh Do-Cong Pham, Nils Hovdenak, Birgitte Emken, Kim Nylund, Tarek Mazzawi, Trygve Hausken and Odd Helge Gilja.

Thanks to Kurt Hanevik for recruitment of participants and help in writing.

I am deeply grateful to my colleagues at the Department of Paediatrics, who make up an exceptionally good work environment. This Department is hold together and skilfully headed by Britt Skadberg.

In particular I owe a lot to Edda Olafsdottir, who once invited me to specialise in paediatric gastroenterology and has remained my most cherished colleague, supervisor and support ever after.

My former tutor during paediatric training, Damien Brackman triggered me to embark on this project. Having examined one of my patients, you jotted down a family tree with a couple of names and telephone numbers, gave it to me and said "Voilà, Rune, this is the beginning of a good story."

This work also involved the Department of Radiology and the Department of Clinical Biochemistry that both have their paediatric branches staffed by dedicated professionals.

Lastly I would like to express my infinite gratitude to my family. I am deeply thankful for the encouragement and support provided ever since childhood by my parents, Astri and Per, and for your eagerness to visit and look after your grandchildren. Thanks also to my mother-in-law Margit Lilleskog, for all your help and all the good moments you create.

I intensely admire my beautiful wife Eli Synnøve and our children Ingrid, Tarjei and Frid. Thank you for your love, your patience and all the adventures you share with me.

Abstract

We studied a large Norwegian family who had been affected by chronic diarrhoea and other symptoms from the digestive tract throughout several generations. Our genetic studies showed that family members with diarrhoea shared markers on the short arm of chromosome 12. In this part of the genome we discovered a new genetic variant in *GUCY2C*. This variant was present in 32 family members with diarrhoea, but not in the healthy family members. We performed functional assays of the *GUCY2C* variant in a cell-model. These studies showed that the genetic variant encodes an increased activity of guanylate cyclase C (GC-C), a protein known to be involved in infectious diarrhoea.

Apart from diarrhoea many family members had also experienced acute intestinal obstructions and/or inflammatory bowel disease (IBD) categorized as Crohn's disease (CD). To our knowledge the constellation of symptoms experienced by our patients had not been reported previously and we assumed that we were dealing with a new disorder. We called this disorder Familial *GUCY2C* diarrhoea syndrome (FGDS). The publication in 2012 of our research constituted the first report of a human disorder caused by a change in the *GUCY2C* gene.

The further work on this thesis aimed at discovering factors involved in the pathogenesis of FGDS with a main focus on its link to IBD.

In parallel with our project multinational case-control studies were charting the genetic basis for IBD, and by 2015 more than 200 genetic associations had been identified. While *GUCY2C* has not been flagged by these population-based studies we hypothesised that common genetic variants within the GC-C pathway may still contribute to the risk of developing IBD. Rather than assessing single genes, we tested whether an aggregation of genes within the GC-C pathway was associated with IBD. For this analysis we used genetic association statistics made publicly available from the largest IBD case-control studies. The GC-C gene list did indeed show significant enrichment of association in IBD.

We then looked at common IBD genetic risk variants in FGDS patients and found that genetic variants in the *NOD2* gene distinguished FGDS patients affected by IBD from those not developing IBD. *NOD2* encodes a protein involved in sensing and removal of bacteria in the intestinal wall.

We studied which genes were differentially expressed in the distal small bowel of FGDS compared to healthy controls as well as patients with CD. Down regulation of metallothionein genes was found in FGDS patients regardless of concomitant IBD when compared to healthy but not when compared to CD. Metallothioneins may serve as signal transducers in the interplay between the human host and its resident bacteria (microbiota). Paucity of these antioxidant proteins may perturb epithelial sensing and clearance of microbes and has been reported in IBD.

We hypothesised that *GUCY2C* related changes of the gut hydration could also impact on its bacterial composition. Using sequencing of the 16S ribosomal RNA gene we analysed the microbial composition in stool from adult FGDS patients, related and unrelated healthy controls as well as patients with IBD. Overall microbiota composition of FGDS patients was different from the other groups, but similar between healthy relatives and unrelated healthy controls. The microbiota of the FGDS patients displayed increased abundance of Enterobacteriaceae and loss of *Faecalibacterium*, findings that may have a pro-inflammatory potential and are found in IBD.

In conclusion we here identified activating mutations of the *GUCY2C* gene as the cause of familial diarrhoea and defined a new disorder, Familial *GUCY2C* diarrhoea syndrome. Our work indicates that genetic testing for mutations in *GUCY2C* should be considered in patients presenting with early onset secretory diarrhoea. Apart from diarrhoea we also highlighted clinical complications potentially linked to hyperactivation of *GUCY2C*, such as IBD, intestinal obstruction and dysfunctional gut peristalsis. Our subsequent studies suggested that *GUCY2C* is involved in the crosstalk between the gut mucosa and the adjacent microbiota, providing a potential clue to the link between *GUCY2C* and IBD.

List of publications

1) Fiskerstrand T, Arshad N, Haukanes BI, Tronstad RR, Pham KD, Johansson S, Håvik B, Tønder SL, Levy SE, Brackman D, Boman H, Biswas KH, Apold J, Hovdenak N, Visweswariah SS, Knappskog PM. *Familial diarrhea syndrome caused by an activating GUCY2C mutation.* N Engl J Med. 2012 Apr 26;366(17):1586-95.

   Reprints were made with permission from the Massachusetts Medical Society.

2) Tronstad RR, Polushina T, Brattbakk HR, Stansberg C, von Volkmann HL, Hanevik K, Ellinghaus E, Jørgensen SF, Ersland KM, Pham KD, Gilja OH, Hovdenak N, Hausken T, Vatn MH, Franke A, Knappskog PM, Le Hellard S, Karlsen TH, Fiskerstrand T. *Genetic and transcriptional analysis of inflammatory bowel disease-associated pathways in patients with GUCY2C-linked familial diarrhea.* Scand J Gastroenterol. 2018 Oct 24:1-10.

   Reprints were made with permission from the Taylor & Francis Group.

3) Tronstad RR, Kummen M, Holm K, von Volkmann HL, Anmarkrud JA, Høivik ML, Moum B, Gilja OH, Hausken T, Baines J, Karlsen TH, Fiskerstrand T, Hov JR. *Guanylate Cyclase C Activation Shapes the Intestinal Microbiota in Patients with Familial Diarrhea and Increased Susceptibility for Crohn's Disease.* Inflamm Bowel Dis. 2017 Oct;23(10):1752-1761

   Reprints were made with permission from Oxford University Press.

Abbreviations

AMPK	Adenosine mono phosphate activated kinase
BMI	Body mass index
BMR	Basal metabolic rate
cAMP	Cyclic adenosine monophosphate
CD	Crohn's disease
CDD	Congenital diarrheal disorder
CSD	Congenital sodium diarrhoea
cDNA	Complementary DNA
cGMP	Cyclic guanosine monophosphate
CFTR	Cystic fibrosis transmembrane conductance regulator
DNA	Deoxyribonucleic acid
EI	Energy intake
ETEC	Enterotoxigenic *Escherichia Coli*
FDR	False discovery rate
FFQ	Food frequency questionnaire
FGDS	Familial *GUCY2C* diarrhoea syndrome
GC-C	Guanylate cyclase C
GO	Gene Ontology
GSEA	Gene set enrichment analysis
GUCY2C	Gene encoding guanylate cyclase C (GC-C)
GWAS	Genome-wide association study (-ies)
HC	Healthy control
IHC	Immunohistochemistry

IBD	Inflammatory bowel disease
IBS-D	Irritable bowel syndrome with diarrhoea
IBS-C	Irritable bowel syndrome with constipation
LD	Linkage disequilibrium
mRNA	Messenger RNA
NHE3	Sodium-hydrogen exchanger 3 (Encoded by *SLC9A3*)
NOD2	Nucleotide-binding oligomerisation domain-containing 2
PCR	Polymerase chain reaction
PKA	Protein Kinase A
qPCR	Quantitative polymerase chain reaction
RNA	Ribonucleic acid
SAM	Significance analysis of microarrays
SLC9A3	Gene encoding NHE3
SNP	Single nucleotide polymorphism
ST	Heat stable toxin
UC	Ulcerative colitis
WES	Whole exome sequencing
WGS	Whole genome sequencing

Introduction

Diarrhoea takes a high toll on health worldwide and is a major cause of childhood death in low-and middle-income societies.[1] Most episodes of diarrhoea are self-limited and are commonly elicited by heat stable bacterial toxins binding to the epithelial enzyme linked receptor guanylate cyclase C (GC-C).[2] This thesis deals with our discovery and characterization of a novel human diarrheal disorder and its cause, an inheritable change in GC-C function.

Genetic linkage analysis and sequencing helped to identify a previously unknown mutation in the gene *GUCY2C* (which encodes GC-C) as the cause of childhood onset chronic diarrhoea in 32 members of a large Norwegian family. The mutation is a so-called "gain of function" that increases GC-C activity. We named the disorder Familial *GUCY2C* diarrhoea syndrome (FGDS) referring to the mutated gene and the concurrence of diarrhoea with other intestinal symptoms experienced by these patients.

Finding the causative mutation of this rare monogenic disorder constituted the initial element of this thesis, offering the prospect of new insights along two main avenues. Firstly existing and future knowledge about GC-C could be exploited to better understand the disorder, and secondly our studies of FGDS may reveal new knowledge on *GUCY2C,* the function of GC-C and more broadly about the pathways related to GC-C signalling.[3] More than 20% of the FGDS patients had been diagnosed with Crohn's disease (CD), one of the two main types of the inflammatory bowel diseases (IBD). This observation strongly suggested that GC-C deregulation influences the risk for development of IBD. After our initial discovery and characterization of FGDS we aimed at identifying factors that could help to understand the clinical presentation and variable complications of this disorder, focussing on the link between GC-C deregulation and IBD. This had previously only been done in mice.[4, 5]

Our project coincided in time with large international case-control studies (genome-wide association studies, GWAS) that were fruitful in mapping the genetic basis of IBD, opening similar opportunities for disclosing functional implications of the genetic associations. In IBD as a complex genetic trait, most of the individual genetic risk factors are common in the population and the impact on disease development of each variant may be modest and operate only when interacting in complex manners with other genetic and non–genetic factors. Our study aimed at helping to bridge this gap by further characterizing the effects of a *GUCY2C* mutation, which high functional impact was evident clinically as well as in the laboratory. We also evaluated whether these effects could be relevant for development of IBD in nearly one in four adult FGDS patients. Finally, the study of patients with this GC-C activating disorder might clarify unknown biological effects of the recently approved GC-C activating drugs.[4, 6, 7, 8]

This thesis gives first an introduction that presents an overview of diarrheal disorders, emphasising on IBD, the role of genetics and the intestinal microbiota. A vocabulary is provided. Aims for the different subprojects are presented followed by a section covering methodological considerations. Next a summary of findings is presented. Lastly the discussion examines the conclusions drawn in the subprojects. A section is dedicated to how the study might help to guide management of FGDS and ultimately suggestions for topics of further studies are given.

1.1 Diarrhoea

Diarrhoea, defined as the passage of 3 or more liquid stools per day, results from an imbalance of secretion and absorption of fluids in the gastrointestinal tract. Although most episodes of diarrhoea are self-limiting, electrolyte and fluid losses can be large enough to result in cardiovascular collapse and death. Diarrhoea may also impair uptake of micro- and macronutrients, leading to malnutrition, impaired immune defences and a vicious circle of recurrent or persistent diarrhoea.[1] Worldwide an estimated 800.000 pre-school children die each year as a consequence of diarrhoea, the vast majority in low- and middle-income societies.[1, 9] Prevention and treatment of

diarrhoea are thus highly prioritized measures to improve global public health. The most frequent infectious agents underlying moderate to severe diarrhoea and hence important targets of interventions are rotavirus, enterotoxigenic *Escherichia coli* (ETEC), cryptosporidium and *Shigella*.[1] By mimicking intestinal peptides, heat stable toxins produced by ETEC activate GC-C and thus increase formation of the second messenger cyclic guanosine monophosphate (cGMP), eliciting a chain of reactions resulting in net secretion of ions and fluids across the intestinal epithelium.[10] Two crucial steps in this cascade, the inhibition of sodium/ hydrogen exchange (NHE) channels and the opening of the cystic fibrosis conductance regulator (CFTR) are shared by other diarrhoea inducing agents (e.g. *Vibrio Cholera*) employing the second messenger cyclic adenosine monophosphate (cAMP).[11]

The current study was performed in an affluent society in which improvements in nutrition, hygiene and health care during the last century have reduced the incidence and severity of infectious diarrhoea.[12] (https://ourworldindata.org/health-meta)

Life style changes in the same period have been associated with an increased prevalence of other diarrheal disorders such as coeliac disease and IBD.[13, 14] Patients with coeliac disease are to a large extent helped by the fact that the key environmental trigger, gluten, is known and can be avoided in the diet. What triggers and sustains IBD remains to be defined, but in both these common disorders, inheritable factors interacting with intestinal bacteria may play a crucial role.[15, 16]

Technological advances in the field of genetics have paved the way not only for detection of the inheritable basis of diarrheal disorders, but also allowed snapshots of intestinal biological processes, such as tissue gene expression, and the composition of the gut microbiota.[17, 18, 19, 20] To what extent these insights may help to delineate pathogenic mechanisms, may however rely on how well knowledge about genetic factors associated with disease can be combined with knowledge about the relevant biological processes which they involve.[21]

One important challenge herein relate to the fact that the outcome in common disorders may depend on a complex and subtle interplay of numerous genetic and

non-genetic factors, where the individual effects may be hard to discern. Rare familial diarrheal disorders may make an exception since they are usually governed by high impact changes in a single gene.

1.1.1 Monogenic diarrheal disorders

Cystic fibrosis (CF) is the most common monogenic disorder causing diarrhoea in the Caucasian population, with a birth incidence in the range of 1/1,600- 1/5,000. That maldigestive diarrhoea in this disorder results from pancreatic enzyme deficiency has since long been appreciated and can be remedied by enzyme replacement, but CF still remains a serious disorder leading to lung destruction and a median survival ~ 40 years.[22] The discovery and mapping of the genetic changes underlying CF have been crucial to understand disease mechanisms including perturbed chloride transport across epithelia, mucus changes, obstruction of pancreatic ducts and chronic airways disease due to functional changes in the ion channel CFTR.[23] The importance of genetics can hardly be overemphasised in CF and is now expanding from basic science and diagnostics, to tailoring treatment with drugs that specifically target different CFTR defects, providing a realistic hope of substantially improving the outcome in this disorder.[24, 25, 26, 27]

Three brothers in the family described in this thesis were referred to genetic evaluation almost 20 years ago. The geneticist noted that their early onset diarrhoea with dehydration and increased sodium in the stools resembled congenital sodium diarrhoea (CSD), although showing a milder presentation and a different pattern of inheritance than CSD. At that time the tools for studying genetics were fewer, and the molecular basis for CSD as well as many other inheritable diarrheal disorders was still not known.[28] In the following years genetics have helped to reveal the molecular basis for many diarrheal disorders, to diagnose and to inform about prognosis and therapeutic opportunities.[29, 30, 31, 32, 33] The evolving spectrum of monogenic diarrheal disorders show that the aetiology of diarrhoea is variable, usually belonging to at least one of 5 categories:

1) Disorders of epithelial transport (e.g. CF, CSD, congenital chloride diarrhoea)

2) Disorders of epithelial digestive enzymes (e.g. sucrose-isomaltase deficiency)
3) Disorders of lipid transport (e.g. chylomicron retention disease)
4) Disorders of epithelial architecture (e.g. tufting enteropathy)
5) Disorders of immune function (e.g. very early onset IBD)

About half of the roughly 7000 monogenic disorders listed in the OMIM database had the genetic cause resolved by 2013, partially owing to the decreasing cost and increasing availability of high-throughput genetic sequencing.[34] Predictions that the genetic cause of all these disorders will be resolved by 2020 may overly optimistic, e.g. given the challenges relating to discovery and verification of variants that reside outside the coding parts of the genome.[34]

1.2 The human genome and genetic variation

Every human somatic cell contains a full genome, which consists of 3 billion pairs of nucleotide bases arranged in 23 pairs of chromosomes. The information held by the genome, depends on the order of 4 different bases, adenine (A), cytosine (C), guanine (G), and thymine (T). The chain of bases is covalently linked to a complimentary strand where A combines with T and G with C (**Figure 1**). The dynamic properties of this double-strand structure, the deoxyribonucleotide acid (DNA) double helix allows the genome to be copied, translated into proteins and to be repaired.

Only 1.5 percent of the genome encodes proteins. This fraction, called the exome, is divided in 180,000 coding sequences (exons) and roughly 25,000 genes.[35] The remaining part of the genome consists of intergenic regions and intragenic non-coding regions (introns). Except for the male Y and X chromosomes there are 2 copies of each gene and chromosome. Each nucleotide has thus a corresponding partner on another chromosome. Human genetic diversity is due to variations in single nucleotides (SNPs, **Figure 1**), such as substitution of one nucleotide with another, deletions or insertions or due to variations in larger segments of the genome including copy number variation and translocations. The specific variant at a given

position is called the allele of a gene. The combination of two alleles on a chromosome pair then defines the genotype at that position.

When the alleles at a certain position on the chromosome pairs share the same nucleotide, they are defined as homozygous. In the case of nucleotide discrepancy they are referred to as heterozygous. Allelic variation may or may not affect gene function. Genetic variants resulting in reduced functions of the protein (loss-of-function) may have no or minimal impact on the organism if affecting only one allele, but tend to be more harmful if homozygous. On the other hand heterozygous allelic variants are more likely to cause observable traits or disease when conferring a gain-of-gene-function. Across generations, the frequency of these alterations may be subject to selective pressure leading to evolution of the trait.

One frequent aim of genetic studies is to resolve whether allelic variants impact on susceptibility for disease. In genetic terminology, “phenotype” is the term used for the disease or clinical features of the disease / trait that is under study. Single gene disorders display a strong genotype-phenotype correlation, and may be inherited in an autosomal dominant (requiring only heterozygosity for the less common allele) or recessive (requiring homozygosity for the minor allele) inheritance pattern.

The genetic constitution of an individual is inherited from the parents, who both contribute with a single version of each chromosome carried by the reproductive cell (in contrast to the somatic cell that carries a pair of each chromosome). The genetic code stored in the reproductive cell is again the result of an exchange (recombination) of material that took place before conception between the parent chromosome pairs. Despite this recombination, sequences that are spatially close tend to be inherited together, i.e. display linkage. Within families linkage between neighbouring genetic variants helps to track the segregation of disease-causing genetic variants using linkage analysis.

At the population level, the degree to which genetic variants display linkage is expressed statistically as linkage disequilibrium (LD). Knowledge about LD structure

and of the normal genetic variation is vital in case control-studies studies aiming at identifying disease associated genes.[36]

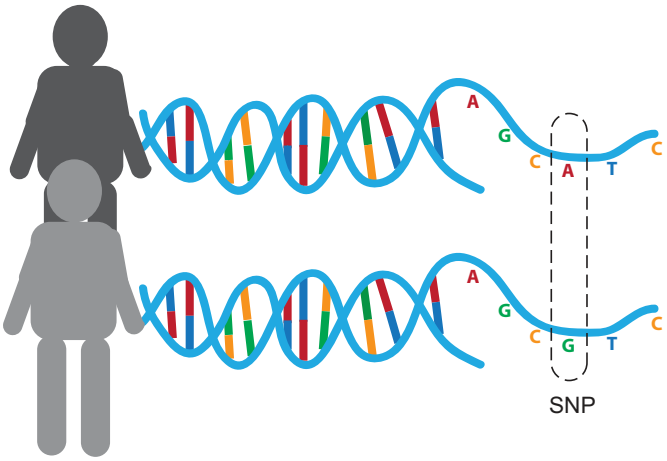


Figure 1. The human genome is organised in double strand (DNA) helices containing a sequence of 4 types of bases (A, T, C, G) that make up the genetic code. Each strand is covalently linked to the complimentary opposite strand where A links with T and G links with C. Variation of base type at a specific position is termed a single nucleotide polymorphism (SNP)

Table 1. Vocabulary	
Allele	Each of two or more alternative variants (of a gene) at a given chromosomal position.
Alpha diversity	Ecological diversity (e.g. richness of species) within a sample.
Amplification	Increasing the amount of a specified DNA sequence e.g. by PCR.
Amplicon	The product of amplification / replication.
Beta diversity	Ecological diversity between samples.
Endogenous	Originating within the organism.
Genotype	The particular combination of alleles at a specific genetic locus.
Genus (pl. genera)	The taxonomic level below family and above species.
Genome	The totality of genetic information of a cell or an organism.
Genomics	Study of DNA sequences and properties of entire genomes.
Haplotype	A specific combination of 2 or more alleles on a single chromosome.
HEK293	Human Embryonic Kidney 293 cells, an epithelial cell line used as a cell-model.
Heterozygous	Having different alleles on the two chromosomes at a particular locus.
Homozygous	Having identical alleles on the two chromosomes at a particular locus.
Horizontal gene transfer	Passage of genes from one organism to another independent of cell division. E.g. virulence factors or antibiotic resistance.
Hybridization	Joining together two complementary strands of DNA from different sources.
Ileocolonoscopy	Inspection of the large bowel and distal small bowel with a flexible endoscope.
Ileum	The distal part of the small bowel.
Library	Collection of DNA fragments originating from an organism or a sample (e.g. stools).
Linkage disequilibrium (LD)	Non-random concurrence in the population of specific genetic variants at different positions mainly accounted for by their physical proximity on the chromosome.
Locus	Position on a chromosome.

Nominal significance	A significance level / p-value not adjusted for multiple testing.
Penetrance	The degree to which a genetic change leads to a specific trait.
Permutation analysis	Statistical analysis using a 0-distribution created by repetitive shuffling of labels.
Phenotype	The observable (e.g. physical, biochemical) characteristics of an individual. Genotype is the genetic contribution to phenotype.
Phylum (pl. phyla)	The taxonomic level below kingdom and above class.
Plasmid	Small circular DNA molecule that replicates independently of the genome.
Recombination	Process in which DNA molecules are broken and re-joined in new combinations. E.g. cross over during meiosis.
Single nucleotide polymorphism (SNP)	Variation in a single nucleotide at a specific position that occurs to an appreciable degree (>1%) in the population.
Taxon (pl. taxa)	A group of organisms that are evolutionary related. In taxonomy low rank taxa are aggregated to form higher rank (more generic) taxa.
Taxonomy	Defining and naming groups of organisms according to evolutionary relationships. Organisms are grouped according to a hierarchy of taxonomic levels with the highest level being the most generic (kingdom) and the lowest level the most specific (species).
Transcription	Copying of one strand of DNA into a complementary RNA sequence.
Type I error	False positive test result.
Type II error	False negative test result.
UniFrac	A distance metric that incorporates information about relative phylogenetic relatedness, used to compare biological communities.
Volvulus	Rotation of mesenterium and intestine leading to obstruction.
Wild type	Natural occurring or typical genetic variant.

1.2.1 Monogenic versus complex diseases

In medical genetics we are concerned with how human diseases and related traits (phenotypes) are influenced by genetic variation. High blood pressure and tall stature are examples of phenotypes that are partially inheritable. Blood pressure may be determined by a combination of several genes, environmental exposures and life style such as smoking, diet and exercise, and is thus regarded as a complex or polygenic trait. In contrast, two copies of the F508del variant in the *CFTR* gene inherited from healthy, heterozygous carriers inevitably leads to the serious disorder CF. Since CF is caused by variants in only one gene (*CFTR*), it is called a monogenic disorder. However other genes and environmental exposures do affect the phenotype in monogenic disorders. For instance, approximately 15 % of CF neonates presents with severe intestinal obstruction due to impaction of sticky meconium. A comparison of CF patients with and without this complication (called meconium ileus) has revealed that the added burden of common genetic variation (SNPs) in ion channels other than CFTR may determine this severe phenotype.[37] Similarly other genetic variants (in different pathways) may aggravate liver injury in CF.[38]

Finding genetic variants (modifier genes) that impact on the phenotype may help to further understand disease mechanisms in monogenic diseases.[39]

The distinction between monogenic and complex disorders is conceptually useful, but in reality genetic disorders exist along a spectrum, where alterations in one or a few genes can have high or low impact on development of a phenotype, with the number of genetic and non-genetic factors interacting in causing disease being higher in complex disorders **(Figure 2)**. Monogenic and complex disorders that are phenotypically similar may display genetic overlap, suggesting that shared mechanisms are involved. Studies of IBD provide a relevant example, where rare variants in *IL10RA* encoding the interleukin 10 receptor, impair development of tolerance to gut bacteria and give rise to severe "Very early onset IBD", while common variants in the *IL10* gene contribute to the polygenic risk for conventional IBD. [40, 41]

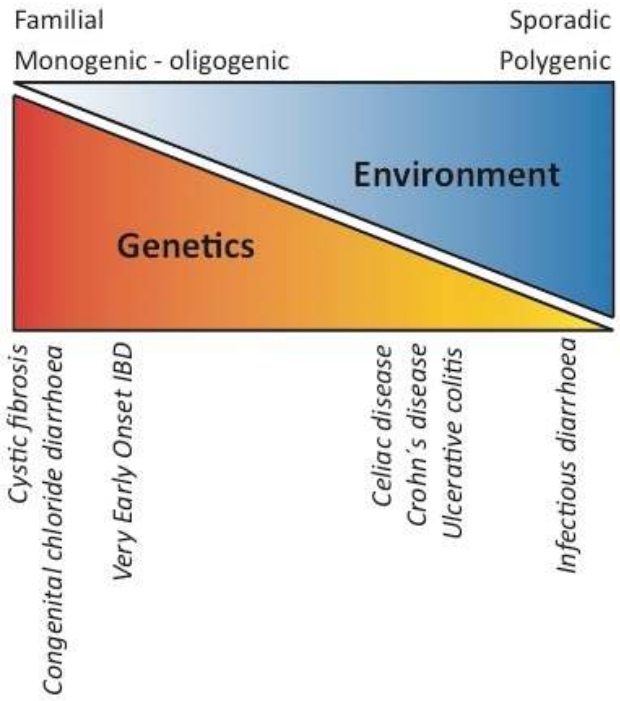


Figure 2. The rectangle schematically illustrates the relationship between genetic (orange) and environmental (blue) contribution to selected diarrheal disorders, and suggests a concept where there is a continuum between monogenic and complex (polygenic) disorders. Single genes play a dominant role in monogenic disorders (e.g. CFTR in cystic fibrosis, IL10RA in very early onset IBD) on the left, while the complex disorders towards the right side are thought to arise predominantly from environmental exposures interacting with several genetic factors.

1.3 Inflammatory bowel disease

IBD are disorders characterized by chronic/ relapsing intestinal inflammation thought to arise as an inappropriate immunological response to microbial or other environmental exposures in a genetically susceptible host.[21] Ulcerative colitis (UC), which is restricted to the colon and CD, which can affect any segment of the gastrointestinal tract are the two main subtypes of IBD.

IBD typically presents in early adulthood with prospects of considerable lifetime morbidity.[42, 43] Diarrhoea, malabsorption, anaemia and abdominal pain are typical symptoms of IBD. The course of the disease, location, extent of inflammation and complications are still highly variable suggesting underlying differences in pathophysiological mechanisms. Patients with CD are particularly prone to acquire intestinal fibrotic strictures, fistulas or abscess formation with the need for surgery. Development of colorectal carcinoma, extraintestinal autoimmune phenomena such as arthritis, uveitis, liver disease and psoriasis are associated with both types of IBD.[44, 45, 46]

Management of IBD relies on anti-inflammatory drugs, nutritional support and sometimes surgery. Biological treatments, e.g. antibodies directed towards tumor necrosis factor alpha (TNFα) or integrin α4β7, have revolutionized the treatment of IBD, but still lack of response, loss of response and concern about long-term side effects call for better understanding of disease mechanisms and novel treatment strategies.[47, 48]

A rising incidence of IBD in industrialised countries during the last decades suggests that life style changes and environmental exposures are important for disease initiation.[14] A role of dietary factors is also indicated by the therapeutic response found in many paediatric CD patients treated with exclusive enteral nutrition.[49] Long-term exclusive intake of a nutritional formula is poorly tolerated and there is an increasing interest as to whether other means of dietary modification (e.g. removal of causal factors) or oral therapies that modulate the intestinal microbiota hold clinical utility.[50, 51]

1.3.1 Genetics of IBD

Familial aggregation of IBD is evident in the clinic as well as in epidemiological studies indicating that IBD is partially inheritable.[42] Twin studies have estimated that heritability contributes to > 50% of the liability to CD and somewhat less to UC.[52] Many families affected by IBD participated in the earliest efforts to map the genetics of these disorders using linkage analysis. The first risk locus discovered by this

method was dubbed IBD1.[53] Further dissection of this locus revealed that the linkage is specific for CD and is explained by SNPs in the gene *NOD2*.[54 55]

This landmark finding demonstrated that it is possible to detect genetic factors influencing the risk for development of IBD and other complex disorders. However the linkage study design is limited by the availability of families with a sufficient number of affected members and has proven inadequate when searching for genetic variants with low effect size.

The focus thus switched to case-control studies, in which the genetic constitution of patients with IBD was compared to that in healthy unrelated controls. This design called genome-wide association study (GWAS) was facilitated by the microarray technology that allowed rapid and affordable methods to capture SNPs throughout the whole genome in a large number of participants.

At the start of the present project the number of IBD associated genetic loci detected by GWAS approached 100, and with subsequent studies recruiting increasing number of participants the number now exceeds 200 with approximately half of the risk genes being shared between UC and CD.[17, 56, 57] *NOD2* remain the strongest association signal in CD. This gene encodes nucleotide oligomerization domain containing 2 (NOD2), an intracellular receptor for bacterial wall patterns.[54, 55] This finding, followed by functional studies of NOD2 supported the suspicion that CD is linked to sensing and response to intestinal bacteria.[58]

Both IBD subtypes are associated with genes implicated in the innate and adaptive immune system as well as the intestinal mucosal barrier. In UC the human leucocyte antigen (HLA) complex on chromosome 6 represents the strongest genetic association,[59] which could imply that the colitis is an autoimmune response to a yet unknown antigen. This concept is supported by the continuous and relatively superficial distribution of UC.[60] CD on the other hand tends to be discontinuous, more invasive and CD specific genetic risk is linked to impairment of bacterial sensing and autophagy, i.e. important factors in the maintenance of a gut barrier function to commensal flora and pathogenic microbes.[61, 62, 63] Even if vira and

bacteria are found in mucosal specimens of CD patients, the concept of CD as a chronic infectious disorder is likely overly simplistic.[64, 65] Furthermore, genetic studies have challenged the diagnostic distinctions under the IBD umbrella, suggesting that both CD and UC comprise several distinct disease entities representing a spectrum of underlying pathology.[66, 67]

Key questions regarding the initiating factors in IBD and the role of genetics remain unanswered. The majority of the risk markers reside outside the coding parts of the genome and their implications for individual gene function remain to be clarified.[68] While increasing the number of participants in ever larger studies have provided power to discover new associations, a large gap still remains between the heritability estimated from epidemiological studies and the heritability explained by GWAS.[68] The "missing heritability" could partially be explained by rare genetic variants that are not captured by the association study design, but typically found in patients with severe phenotypes or with familial clustering of disease.[21]

An association between familial diarrhoea and IBD had been reported in patients with congenital chloride diarrhoea, which is caused by loss-of-function variants in the anion exchanger gene *SLC26A3*.[32] Interestingly common genetic variants in the same gene have been associated with risk for development of UC, suggesting that aberrations in this ion transporter might act in similar ways to increase susceptibility to IBD in the general population as in patients with inheritable diarrhoea.[69, 70] That alterations of ion transport could shape the intestinal bacterial communities had been reported at the time of publication of paper I, but to our knowledge no studies had hypothesized this could be a potential driver of intestinal inflammation.[71]

1.3.2 The intestinal microbiota

Large communities of bacteria, archaea, vira, fungi and protozoa, collectively referred to as microbiota, inhabit the human gut. Traditionally medicine has regarded microbes largely as potential troublemakers (e.g. *Vibrio cholera*, *enterotoxigenic E. coli, Salmonella, Clostridium difficile* etc.). Diarrhoea caused by *Clostridium difficile* usually follows use of antibiotics, but are effectively treated by rectal infusion of

microbiota, showing that our microbiota can also protect against illness.[72] In fact the intestinal microbiota is now emerging as a new "organ" that serves its host by contributing to metabolism, immunity and nutrition.[73]

The study of the interplay between the host and the microbiota has been greatly facilitated by the application of genetic methodology.[74] New sequencing technology has made it possible to capture the totality of microbial genes (microbiome) of the gut, without depending on cultivating these organisms in the lab. The genes carried by the intestinal microbiota can tell us not only what species are present, but also what these microbes can do. The intestinal microbiome encode proteins involved in a large number of metabolic and other functions.[75] This enables the microbiota for example to generate nutrients and immunologically active molecules (e.g. butyrate) from otherwise indigestible dietary fibers.[76]

We use two basic terms regarding microbiota diversity. Alpha diversity denotes the richness of species within an individual (or sample). Beta-diversity is a measure of diversity between individuals/samples. While the composition of bacterial species is quite variable between healthy individuals the metabolic capacity of the intestinal microbiome is fairly similair.[75]

Classification of the microbiota adopts the taxonomic principle, which places the organism in a hierarchic system (**Figure 3**). Today genetic overlap and discrepancy between microbes define their taxonomy, i.e. their relatedness, conceptualised by Darwin as branches in the "tree of life", as a function of microbial evolution. Horizontal gene transfer between microorganisms defies the evolutionary tree of life concept and microbial taxonomy is therefore a knotty and evolving discipline.[77]

The intestinal microbiota undergoes compositional changes from birth and onwards, partially depending on microbial exposure during delivery (caesarean section versus vaginal) and dietary factors (breastmilk versus formula, moment of weaning).[78] From later childhood the intestinal microbiota remains fairly stable throughout life, but is responsive to dietary changes and use of antibiotics.[73, 78, 79] Cell-wall polysaccharides from bacteria encountered in early life such as *Bifidobacterium species* and *E. coli*

may play a decisive role in immunological tuning of the gut and affect the fine balance between tolerance and responsiveness to commensal bacteria and thereby susceptibility to a range of diseases.[80, 81, 82]

Taxonomy		
Kingdom	Bacteria	Animalia
Phylum	Proteobacteria	Chordata (vertebrates)
Class	Gammaproteobacteria	Mammalia
Order	Enterobacteriales	Rodentia
Family	Enterobacteriaceae	Muridae
Genus	*Escherichia*	*Mus*
Species	*Escherichia coli*	*Mus musculus (house mouse)*

Figure 3: Taxonomy of ***E. coli*** and the common house mouse

Apart from age and diet, the composition of the intestinal microbiota varies across geography and disease state. Loss of alpha-diversity is a prominent feature of several diseases, including morbid obesity, diabetes and IBD.[16, 83, 84, 85] Interestingly this phenomenon is associated with the transition from a traditional, mainly plant based diet to a modern diet, dominated by a higher intake of simple sugars, fat and animal proteins as well as the increasing use of antibiotics.[86, 87] Alterations in the intestinal microbiota have thus been proposed as an explanatory causal link between the modern life style and the rising prevalence of disorders such as IBD and diabetes.[51, 85] This hypothesis has been supported by animal and human studies including one showing that obese patients with the metabolic syndrome restored insulin-sensitivity after rectal infusion of faeces from a healthy donor.[88]

Faecalibacterium prausnitzii is a prominent beneficial and usually highly abundant bacterium of the gut. This bacterium contributes to the intestinal homeostasis by producing butyrate, a short chain fatty acid that nourishes the intestinal epithelium and regulates ion transport and inflammation.[89, 90] Depletion of *F. prausnitzii* has consistently been demonstrated in the stools and mucosa of IBD patients,[16] and is a predictor for relapse following surgery in CD patients.[91] Various experiments support its role in protection against inflammation.[92, 93] Still, this notion has been challenged by a trial showing persisting loss of *F. prausnitzii* and clinical improvement in paediatric CD patients treated with exclusive enteral nutrition.[94]

Enrichment of the family Enterobacteriaceae is another characteristic of the compositional microbiota imbalances (dysbiosis) found in IBD.[16] One important question is whether the dysbiosis contributes to initiate and sustain inflammation or whether it is only a result of inflammation.[95] To this end it is of great interest to study the microbiota in individuals that are at genetic risk, but have not (yet) developed IBD.[96]

Given its importance in health and disease and the complexity of the interaction between man and his/her intestinal microbiota, it is vital to further understand the factors that shape gut ecology.

2. Aim of project

The project had 2 main aims,

1. To identify genetic and clinical characteristics of a novel inheritable disorder. This comprised 3 subprojects described in paper 1.
   a. To find the chromosomal region where a causal genetic variant was residing, and by sequencing of this region to discover the variant.
   b. To study the biological effects of the identified variant in a cell-model.
   c. To gather and systematise clinical information about the patients.

2. To look at possible biological underpinnings of the disease manifestations, focusing on the intestinal inflammation. This aim included 4 subprojects:
   a. To test whether the gene discovered in paper I and other genes implicated in the pertinent biological pathway are enriched in GWAS of IBD.
   b. To interrogate whether known genetic IBD risk variants relate to clinical manifestations of the disorder.
   c. To characterise the global gene expression in distal small bowel of patients compared to healthy controls and to unrelated patients with IBD.
   d. To characterise the intestinal microbiota in patients, compared to related and unrelated healthy controls and to unrelated disease controls with IBD.

3. Methodological considerations

3.1 Overview and study design

Paper I

We used genetic linkage analysis to search for the genetic cause of a presumably novel inheritable diarrheal disorder. Following identification of a genetic region linked to disease status, we searched for a candidate genetic variant using DNA sequencing. Any rare genetic variants discovered after sequencing were evaluated as to their pathogenic potential and lastly functionally tested in a cell-model using site-directed mutagenesis.

The concept of genetic linkage analysis is that disease-causing variants tend to be inherited (linked) together with other genetic markers residing nearby on the same chromosome. After typing markers across the whole genome in each participant, biostatistics tools aggregate information from these markers, how participants are related and disease status to identify the genomic region most likely containing the causative genetic change. For detection of previously unknown rare variants and their pathogenic potential, the sequence data were compared with reference DNA sequences in humans as well as in other species using different bioinformatics tools (see **Table 4**).

The choice of using linkage analysis seemed reasonable because the high number of affected individuals in the pedigree provided robust statistical power, and also due to the relative low cost of microarray SNP mapping.[97] An alternative approach, is to sequence the whole genome (WGS) or only its protein encoding part (whole exome sequencing/WES) and look for rare variants segregating with disease status. Here sequencing and evaluation of rare variants however easily become more extensive, and the overall project potentially more costly and time consuming.[97, 98]

Paper II

This paper reports 3 different experiments aiming at elucidating a molecular link between GC-C activation and development of IBD.

1) We hypothesised that common variants in genes participating in the GC-C pathway may contribute to development of IBD. Therefore, we used gene set enrichment analysis (GSEA) to test whether a set of genes in this pathway was over-represented (enriched) in the top end of the association statistics from publicly available IBD GWAS.[99] Although several limitations apply regarding the interpretation of GSEA in GWAS data, the method is valuable to detect the accumulated effect of genetic variants that individually might be too weak to pass the threshold for genome-wide significance (defined as $P<10^{-8}$).[100 101, 102]

2) We then evaluated whether IBD genetic risk variants (as previously identified in GWAS) impact on development of IBD in FGDS patients. This was an exploratory study, where genotypes of IBD affected and non-affected FGDS patients were compared taking into account various possible modes of inheritance.[39]

3) Finally, we performed a hypothesis-free experiment of gene expression in the distal small bowel from biopsies of FGDS patients compared to healthy controls and to disease controls with CD. The aim was to identify molecules that could point to biological pathways potentially involved in the pathogenesis. Biopsies from the distal small bowel were chosen since this is the predilection site for inflammation in FGDS patients with concomitant IBD. Inference about cause and effect is limited in a cross-sectional study design compared to prospective studies (e.g. studying patients before and after development of a certain complication). However, using microarrays we could evaluate the difference in transcription of > 25,000 genes between the different groups. We used quantitative polymerase chain reaction (qPCR) as a complimentary method to more accurately quantify gene expression and validate key findings. Ultimately we examined biopsies with

immunohistochemistry to evaluate qualitatively whether transcriptional changes also were evident at the protein level.

Paper III

This project was a cross-sectional study of the intestinal microbiota in FGDS patients compared to unrelated IBD patients, as well as related and unrelated healthy controls. The aim was to characterise the impact of GC-C activation on the intestinal microbiota. We chose to use 16S rRNA gene sequencing of faecal samples because this is a well-established method, which provides good coverage of different types of bacteria. By targeting only the 16S rRNA marker gene, it limits the sequencing load and costs.[103] The more expensive, time- and computer-intensive alternative, "shotgun" sequencing can provide data on the full microbial genome (microbiome), add information about its functional properties and give a better resolution of particular bacterial species, but none of these advantages were considered crucial as a first step to clarify whether changes could be detected at all.

We primarily tested the hypothesis that *GUCY2C* mutations constitute an inheritable factor influencing the intestinal microbiota. Secondarily, we examined whether findings in FGDS patients aligned with published reports on gut microbiota involved in intestinal inflammation and whether findings in FGDS with IBD were similar to those found in conventional IBD patients

.

Most samples were sequenced three times (according to two different amplification protocols and in two institutions) allowing us to control for technical and methodological biases. Stool samples from healthy familial controls (without FGDS) were collected to assess potential confounding effects of geographics, other genetic factors and nutritional habits of the family.[79, 104]

3.2 Selection of participants and clinical data

The papers constituting this thesis report clinical data that are predominantly cross-sectional or retrospective. Only partially the papers reflect the point that most of the

FGDS patients have been evaluated regularly from 2012. This follow up has helped us to update clinical data as needed and ensured real-time observations of FGDS patients experiencing complications, e.g. intestinal obstructions needing surgery.

Paper I

Accurate assignment (affected or not affected) of participants is important when searching for disease causing variants in genetic linkage analysis. Even if the familial diarrhoea appeared to be inherited in an autosomal dominant pattern, we could not know in advance whether the segregation of the trait (diarrhoea) perfectly aligned with the underlying causative genetic variant we were searching for. The trait could have appeared at later age in the apparently healthy relatives (late penetrance) or only in a subset of mutation carriers (partial penetrance). Accordingly in the linkage analysis only adult family members were included in the group classified as unaffected.

The patients with familial diarrhoea as included in paper I were members of three branches (A, B and C) of a large family (**Figure 4**), and two of these branches (A and B) were studied independently before we realised that they belong to the same kindred and carry the same mutation. Since branches B and C were discovered later (by the author of this thesis), only patients in branch A were included in the initial genetic linkage analysis.

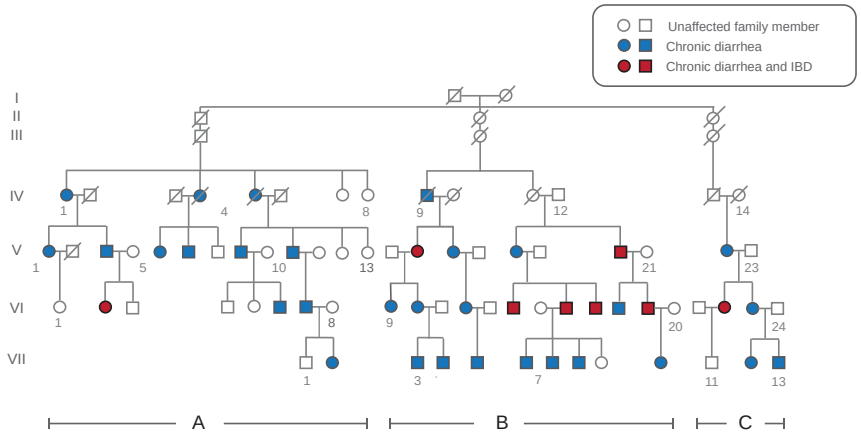


Figure 4: Pedigree of the large family studied in this project. Filled circles (females) and boxes (males) mark individuals affected by diarrhoea, and red filling marks individuals that are also diagnosed with inflammatory bowel disease (IBD). There are three main branches (A, B and C) in the family.

All affected family members contributed with clinical information of help to outline the variable complications associated with carrier status of the disease causing genetic variant. This information was obtained during systematic clinical consultations, questionnaires, review of clinical records and lab results. We performed further investigations such as radiology and hydrogen-breath tests in some of the patients.

We used the Rome 2 questionnaire (**Table 2** and **Appendix 1)** validated for functional abdominal disorders from patients and healthy relatives alike. This choice was based on the experience that symptoms (e.g. intestinal gas, loose stools, abdominal pain) are similar in functional abdominal disorders and FGDS.

Table 2.	Questionnaires and scoring systems used in the thesis	
	What it is:	Where used:
Rome 2	Self reported symptom scoring used to diagnose and classify functional abdominal disorders.	Adult family members, paper I
GSRS	Self reported abdominal symptom score, shorter than Rome 2/3	All participants paper III
Food frequency questionnaire (FFQ)	Self reported habitual intake of >200 food items	FGDS and healthy relatives, paper III
Harvey-Bradshaw Index (HBI)	Disease activity score in CD based on objective findings and reported symptoms	FGDS and CD, paper III
Simple Clinical Colitis Activity Index (SCCAI)	Disease activity score in UC based on objective findings and reported symptoms	UC patients in paper III

Paper II

Publicly available IBD genetic case-control association (GWAS) data

The first experiment in paper II used publicly available datasets obtained in GWAS performed by a consortium of groups, representing in total 12,882 IBD cases and 21,770 healthy controls.[17] Our project did not contribute to collection of these data. We used the summary statistics for association analysis of the genetic markers. All participants were of European descent, and since IBD is more prevalent in certain populations and tend to run in families, the data had undergone some degree of control for population stratification and close relatives had been removed. IBD was represented by CD, UC, as well as the combined data from both entities, there were thus three datasets available for our analysis. Since FGDS patients had been diagnosed with CD, and the GC-C animal models colitis we chose to use all three datasets in the analysis.

IBD-risk variants in FGDS

In paper II we assessed known IBD genetic risk variants in FGDS, hypothesising that development of CD in FDGS patients could be explained by a higher burden of IBD-genetic risk variants. The median age of onset of CD in the population as well as in FGDS is ~30 years and we chose this as the lower age-limit for inclusion in the control group to represent FGDS patients without IBD.[42] FGDS patients were grouped as CD affected when fulfilling both of the following criteria:

1) A history of clinical suspicion of CD, supported by endoscopy findings.
2) A history of treatment for CD or ileitis

This classification differs from standard IBD diagnostic criteria by not requiring histopathological examination obtained by biopsies.[105] On the other hand using standard criteria would have resulted in exclusion of two patients who had been treated for IBD, but for whom histopathology specimens were lacking.

Gene expression in ileal mucosa

All FGDS patients in branch A (**Figure 4**) were invited to participate in the study of gene expression in the distal small bowel. This involved investigation with ileocolonoscopy, which is an invasive and sometimes unpleasant endoscopic procedure. Several of the FGDS patients had already undergone this investigation at least one time before the study. After recruiting the first eight patients we decided that further inclusion should be restricted to situations where ilecolonoscopy was considered helpful for clinical decision-making, e.g. to rule out or monitor IBD. Additional four mutation carriers were investigated with ileocolonoscopy in the study period, and three of these were included in the study.

Healthy controls had been recruited for an unrelated project by advertisement and a financial compensation to perform the procedure, and absence of gastrointestinal symptoms was the main inclusion criterion. Individuals that volunteer for such studies may differ systematically from cases e.g. regarding lifestyles that could impact on intestinal gene expression. In our study smoking was indeed more common

in healthy controls than in FGDS patients.[106] For the sake of our own reassurance, we therefore performed an exploratory analysis (not commented or shown in paper II) excluding smokers and validated the main results shown in paper II. Healthy family members might generally be motivated to contribute to increase knowledge about health problems affecting close relatives and would arguably have constituted the ideal control group. However, these advantages were outweighed by the availability of the above mentioned biopsies, and we found it not to be ethically compliant to expose these healthy relatives to endoscopy.

The unrelated CD patients were recruited from the outpatient clinic at Ullevål University Hospital, which provide tertiary- and secondary-level health services. In order to minimize the contribution of inflammation to the gene expression profile, normal histology, was made a selection criterion for the unrelated CD patients.

Paper III

Collection of microbiota samples in patients and controls

We invited all FGDS patients regardless of age to participate in the study of the intestinal microbiota. Since the number of children was low and age is associated with the gut microbiota composition, participants below 15 years of age were excluded.[87] All adult healthy family members were invited to participate as one of the control groups in the study, intending to take advantage of their quality of matching FGDS in many aspects relevant for gut ecology such as genetics (apart from the disease causing variant), household factors, nutrition and geography.[79, 104, 107]
We also included a larger group of unrelated healthy controls recruited from the Norwegian Bone Marrow registry.

Using individuals that have volunteered as bone-marrow donors as a healthy control group holds the advantage that they had been screened for diseases and are generally considered healthy. Still, health issues that can influence the intestinal microbiota, could theoretically have arisen for certain individuals after inclusion in the registry and cannot be accounted for.

Antibiotic use and nutrition influence the intestinal microbiota, so we excluded all samples taken within 4 weeks of antibiotic treatment and participants reporting special diets (e.g. cow milk-free, gluten-free, vegan).

Collection of other variables in participants

In order to assess whether FGDS patients share food preferences with their relatives and the potential impact of nutritional habits on observed changes in the gut microbiota, we obtained so-called "food frequency questionnaires" from these two groups. The food frequency questionnaire, validated for use in the Norwegian population, inquires about the habitual intake of a large list of commonly used food items.[108] This information is extracted and converted to estimates of intakes of various macro- and micro-nutrients (e.g. energy, fat, carbohydrates, proteins, fibre, added sugar, alcohol). The reporting is liable to systematic biases (e.g. underreporting the total amount of food and/or specific items perceived to be "unhealthy"), and in an attempt to minimize such bias we normalized measurements to total energy intake.[109]

While 24-hours food recall or 4-days food registration may outperform food frequency questionnaires when assessing the food intake on short term, the latter is considered better suited to define food habits and thus nutrients that could influence intestinal microbiota over time, e.g fibre.[110, 111] We calculated the basal metabolic rate (BMR) according to weight, age, and gender,[112] in order to estimate the rate of obvious underreporting.[108]

We used the GSRS (gastrointestinal symptom rate scale, see **Table 2** and **Appendix 2**) questionnaire to record symptoms in paper III.[113]

FGDS patients were stratified for IBD status using the same criteria as in the study of IBD risk variants in paper II. Furthermore, we included stool samples from unrelated IBD patients from the outpatient clinic of Oslo University Hospital Ullevål. The same exclusion criteria were applied as in healthy controls. In order to detect active

intestinal inflammation, we used clinical disease activity scores (Harvey Bradshaw Index (HBI) and the Simple Clinical Colitis Activity Index (SCCAI) in CD and UC respectively) and calprotectin as a stool marker of intestinal inflammation. [114, 115, 116]

3.3 Sample collection, preparation and storage

DNA isolation

Human DNA used for mapping SNPs, sequencing and IBD-risk variant genotyping in paper I and paper II was isolated from whole blood using a standard procedure that combined mechanical, chemical and electromagnetic extraction.[117]

RNA isolation from small bowel biopsies

The transcript of genes, RNA, shows variation throughout the day, across different tissues and cells, in response to environmental exposures and disease states, and can easily be altered during the sampling procedures due to degradation (e.g. by RNases). [106, 118, 119] All tissue samples were obtained during endoscopy after an overnight fast. In order to prevent degradation, we preserved the biopsies immediately in fluid nitrogen at -80˚C. We extracted total RNA was in presence of an RNase inhibitor after thawing and physical disruption of biopsies. RNA extraction followed a standard method and was performed simultaneously in all patient and control materials.[117, 120] The RNA was amplified and reversely transcribed to complimentary RNA (cRNA), and then biotin-labelled. We checked the amount and purity of RNA by electrophoresis and photospectrometry. Three samples (two healthy controls and one CD) were excluded due to low RNA-integrity.

Changes in cell composition following proliferation, differentiation or recruitment of certain cells e.g. due to inflammation can contribute to gene transcript differences when comparing two disease states.[19] A standard approach to this issue, which we adopted, is to obtain and examine qualitatively an adjacent biopsy in light microscopy.[19, 121] These biopsies were formalin fixated and embedded in paraffin,

stained and evaluated according to standards used in clinical practice. The histology samples were then used to examine whether particular changes of gene expression were also observed at the protein level as assessed by immunohistochemistry (see below).

DNA isolation from stool samples

The composition of the cell-wall differs substantially among prokaryotic organisms, e.g. thick in gram-positive bacteria such as Actinobacteria and Firmicutes, thin in gram-negative bacteria such as Enterobacteriaceae. The choice of extraction method may thus impact on retrieval of DNA of the respective taxa.[117, 122]

Different taxa may show variable degree of survival and proliferation following sampling, and immediate analysis of fresh stool samples may be considered ideal in microbiota studies,[103] but with most participants living distant from the laboratory this was not practicable. We provided the participants with a collection device and stool collection tubes (Stratec Molecular GmbH, Berlin, Germany) containing reagents specifically designed for preservation and subsequent extraction of DNA for sampling at home and instant expedition by mail. [123, 124] In accordance with the manufacturer instructions, specimens arriving the laboratory more than 72 hours after sampling were excluded. Specimens were frozen at a maximum temperature of -20˚C upon arrival.

All subsequent DNA extraction was done at a single laboratory bench with careful attention to hygiene in order to avoid contamination and degradation of samples. DNA extraction protocols used for prokaryotic DNA usually employ chemical, mechanical and thermal methods.[122] In our study, we used the PSP Spin Stool DNA Kit method (Stratec Molecular GmbH, Berlin, Germany), which combine thorough homogenisation, enzymatic protein digestion at 80º C and zirconium bead beating followed by removal of substances that could inhibit subsequent PCR-amplification of DNA.

3.4 Sequence analyses of DNA and RNA

Availability of robust methods to determine the order (sequence) of DNA and RNA was a prerequisite for the current project, and several have been employed, as outlined in **Table 3**.

Table 3. Methods used for determination of DNA or RNA sequences in this project

Application:	Microarray	qPCR	Sanger sequencing	Massively parallel sequencing
Linkage analysis	(1)			
Variant-detection				(2)
IBD genetic risk variants in FGDS	(3)			
Gene expression	(4)			
Microbiota				(5)
Throughput	Intermediate	Low	Low	High
Detect new variants	No	No	Yes	Yes
Cost	Low	Low	Intermediate	High

Equipment:

1) Human Mapping 250K NspI Array (Affymetrix, Santa Clara, USA)
2) Illumina HiSeq (Illumina Inc., San Diego, CA, USA)
3) The "Immunochip", Illumina iSelect HD custom genotyping array
4) Illumina HT12 v4 Bead Chip
5) Illumina MiSeq (performed with 2 alternative library protocols, at 2 independent institutions)

Table 3 We used 4 in principle different methods (columns 2-5) to obtain nucleotide sequence data in this project. The chosen methodology is marked with a shaded area for 5 subprojects (rows 3-7). Apart from microarray SNP mapping employed in linkage analysis, technical validation using complimentary methods was performed in all subprojects.

qPCR; quantitative polymerase chain reaction

3.4.1 Microarrays

A common feature of the microarrays used in this project is that they provide a spatial arrangement (array) of probes designed to capture sample DNA/RNA by hybridization. The annealing of sample to probe is registered as a fluorescent signal emitted from the respective positions on the microarray upon excitation with laser.

Mapping genetic markers for linkage analysis

Genetic markers used for tracking monogenic inheritance with linkage analysis should show appreciable variation between individuals and be distributed throughout the whole genome with intervals short enough to minimize recombination between the unknown mutation and the nearest markers. SNPs fulfil these criteria, and have replaced other types of markers previously used in linkage analysis.[53 125] We chose to map SNPs for linkage analysis in paper I using microarrays covering a modest number of SNPs (~250,000).

The SNP microarray is designed with clusters of oligonucleotide sequences (probes) selected for being complimentary to sequences containing SNPs, imprinted in defined positions on a glass surface. Sample DNA is applied after it has been fractioned into shorter sequences and labelled with a substance able to emit a fluorescent signal. Hybridizing to respective probes subsequently captures sample sequences. A scanner reads the pattern and intensity of fluorescent signals from the sample sequences on the microarray, and using bioinformatics tools (see **Table 4**) these are converted to information about the genotype, (e.g. C/C, C/T or T/T) for each SNP position.

Genotyping IBD-risk variants in patients with FGDS

We chose to genotype the current (2012) top IBD-associations, representing 163 genetic loci, and to perform full sequencing of the *NOD2* gene.[68] Genotyping was performed using the Immunochip (Illumina Inc., San Diego, USA). The probe design of this microarray provides high-density SNP coverage in genetic regions implicated by the earliest GWAS in autoimmune and inflammatory disorders, including IBD.[68,

[126] The Immunochip is conceptually similar to the microarray used in genetic linkage analysis, but allows much more narrow spacing of probes.[127]

Gene expression by microarrays

Gene expression microarrays can be used to detect and quantify a comprehensive selection of gene transcripts (RNA sequences) in a single experiment.[119] Massively parallel RNA-sequencing provides deeper coverage of transcripts (e.g. better ability to capture sequence variation and a broader dynamic range for quantifying gene expression), but this alternative was discarded partly due to the higher costs.[119] Subtle signal differences between microarrays might introduce a technical bias in gene expression studies. This was one important reason that we chose to use the bead chip technology described above (Illumina HT12 v4 bead chip, Illumina Inc., San Diego, USA), which allow 12 samples to be analysed on one chip. This microarray capture 47,000 different RNA sequences, covering most protein encoding as well as a large number of non-coding transcripts. In order to minimize batch effects we assured that samples from FGDS, healthy controls and unrelated CD patients were all evenly represented on each chip.

Processing of microarray data

Control for systematic variations between chips regarding signal intensities is recommended, and this possible source of bias was addressed by normalising expression values according to mean within the respective inter-array quantile.[128] Correspondence analysis can reduce complex data to a minimum of two dimensions, of help to visualize trends in the data.[129] Our correspondence analysis of the data did not show deviating samples or batch effects related to the equipment or the RNA-amplification procedure. During the first steps of microarray data processing, we used the GenomeStudio Data Analysis Software (Illumina, Inc., San Diego, USA). We used the freely available Jexpress2012-tool (Jexpress.bioinfo.no) developed at Bergen University for further processing and statistical analyses of gene expression data.[130]

3.4.2 Quantitative PCR and immunohistochemistry

Polymerase Chain Reaction (PCR)

In paper I we used PCR to selectively increase the amount of the *GUCY2C* gene, and in paper III to amplify the prokaryotic 16S rRNA gene (**Figure 3)**. Amplification of a target nucleotide sequence by PCR requires a polymerase enzyme, a template (single stranded DNA), free nucleotides, and a primer pair (flanking each end of the target nucleotide sequence).

Quantitative PCR

In paper II we used quantitative PCR (qPCR) to measure and verify relative changes of genes highlighted by the gene expression microarrays. In qPCR the yield of each PCR cycle is measured as the intensity of fluorescence emitted by labelled probes when digested by the polymerase. The number of cycles necessary to reach a certain signal threshold thus gives information about the target gene quantity in each sample. Target gene quantities can be normalised against genes that are constitutively and thus evenly expressed throughout the samples in order to control for differences across samples regarding the yield of RNA extraction, reverse transcription and amplification.[131, 132] As a control gene we used *ACTB*, which displayed an even expression profile across the groups as evaluated visually and statistically in the microarray data. Subtle variation in control gene expression could still impact on the results and we therefore validated all findings using a second control gene (*B2M*).

Immunohistochemistry

We used hmmunhistochemistry (IHC) to assess whether differentially expressed genes also translated to altered levels of the respective proteins in small bowel biopsies.[133] Slides embedded in paraffin maintain tissue architecture and can thereby show where changes are taking place and which cell types are affected. During the IHC procedure, histology slides were exposed to monoclonal mouse anti-metallothionein (the primary antibody), and stained using the ultraView Universal DAB Detection Kit and the BenchMark Ultra immunohistochemistry/*in situ* hybridization staining module (Ventana Medical Systems, Basel, Switzerland). We

analysed tissue architecture as well as immunostaining of the tissues using qualitative assessments.[133]

3.4.3 DNA sequencing of *GUCY2C* and *NOD2*

Sanger sequencing

Genetic linkage analysis performed in paper I identified a candidate region spanning 28 protein-encoding genes. Our suspicion was early on targeted at *GUCY2C*, due to its role in infectious diarrhoea. In sequencing the exomes of *GUCY2C*, we chose to use a current version of the chain termination principle, also known as Sanger sequencing.[134] Amplification of target sequences (here *GUCY2C* exomes) by PCR was performed in a mix of normal nucleotides (deoxyribonucleotides) and nucleotides chemically modulated to terminate the nucleotide chain (dideoxyribonucleotides). Each of the four different dideoxyribonucleotides were provided with a specific fluorescent dye. The target sequence was resolved by aligning the dye-tags according to information about the length of the respective copies, as determined by capillary electrophoresis and fluorescence detection. This method was also used to sequence the *NOD2* exomes in paper II.

Sequence variant analysis tools-bioinformatics

The raw (unprocessed) data generated from microarrays and other types of sequencing need extensive processing before they can give us biological knowledge. Computer technology allows fast analysis of large datasets, aided by software able to integrate existent knowledge stored in databases, with mathematical and statistical procedures. This field of science called bioinformatics, has become a prerequisite for genomic research. During this project we used a number of bioinformatics tools as outlined in **Table 4**. For example we used CASAVA (Illumina Inc., San Diego, USA) and GenomeStudio (Illumina) to perform data quality-control in paper I and II.[135] We used various tools to align sequence data (reads) and map these to the human reference genome,[136, 137] and to discover and evaluate (annotate) variants (see **Table 4**).[138, 139] The Genome analysis tool kit (Broad Institute, MA, USA) accesses the

knowledge database NCBI human db SNP database, and was used to evaluate whether variants detected had previously been reported.

Whether any genetic variant would change the amino-acid sequence in the respective protein is another important clue as to its potential to alter gene function and cause disease. For the latter assessment, we used a tool that predicts protein-structure from amino-acid sequence and compares the amino-acid sequence of a particular protein across many species (Annovar, www.annovar.openbioinformatics.org).[138] Genetic variants that lead to changes in amino acid are more likely to be detrimental if the amino acid for the given position has remained unchanged throughout evolution, in other words the same in many species. To prove how and whether a genetic variant truly affects protein function may however require testing in a genetically modified organism (e.g. mouse) or cells.

Table 4. Bioinformatics tools and databases used in the project	
Allegro deCode Genetics, Reykjavik, Iceland	Software for genetic linkage analysis.
Annovar www.annovar.openbioinformatics.org	Software and knowledge database for evaluation of genetic variants.
Burrows-Wheeler transform	Data compression algorithm e.g. used when aligning sequence data against a reference genome.
CASAVA Illumina Inc., San Diego, USA	Tool for processing raw DNA and RNA sequence data.
DAVID National Cancer Institute, Frederick, USA	Software for statistical analysis of microarray data on various knowledge databases.
FLASH *Johns Hopkins Medicine, Baltimore, USA*	Software for aligning DNA/ RNA sequences generated by the paired-end procedure.
GeneCards *Weizmann Institute of Science, Rehovot, Israel*	Knowledge database of genes with information e.g. about their function.
GeneOntology www.geneontology.org	Knowledge database of genes and their functions, products and interactions.
Genome analysis tool kit Broad Institute, Boston, USA	Algorithms for analysing DNA and RNA sequence data, e.g. variant detection.
GenomeStudio Illumina Inc., San Diego, USA	Tool for processing and quality control raw sequence data e.g. from microarrays.
GreenGenes www.greengenes.secondgenome.com	Database of 16S rRNA sequence variation in bacteria.
Gene set enrichment analysis/GSEA Broad Institute, Boston, USA	A variant of statistical enrichment analysis in genomic data.
GraphPad Prism GraphPad Software, San Diego, USA	Software for basic statistics.
Ingenuity Pathway Analysis Qiagen, Hilden, Germany	Software and knowledge database for pathway analysis in genetics.
J-EXPRESS 2012 University of Bergen, Norway	Software for processing and statistical analyses of gene expression data.
LDsnpR University of Bergen, Norway	Algorithm for assigning SNPs to genes based on location and linkage disequilibrium.
LEfSE Harvard University, Boston, USA	Statistical tool for analysis and visualisation of microbiota data.
MaAsLin Harvard University, Boston, USA	Statistical tool for controlling co-variables (linear regression) in microbiota analyses.
NCBI human db SNP database National Center for Biotechnology Information, Bethesda, USA	Database of known human genetic variants (SNPs).
Oligo National Bioscience, Plymouth, USA	Software for designing primers e.g. for amplification of exomes.
Picard www.sourceforge.net	Algorithm for manipulating high-throughput sequence data e.g. remove PCR duplicates.
QIIME www.qiime.org	Software package that integrate several bioinformatic tools in a pipeline, developed for analysing and visualising microbiota data.
R www.r-project.org	Programming language developed for statistical analysis.
RefSeq National Center for Biotechnology Information, Bethesda, USA	Reference genome database.
SAMtools www.sourceforge.net	Algorithms for transforming and storing sequence data e.g. into appropriate file formats.
Seqscape ABI, Foster City, USA	Software for analysing sequence data, against reference sequences.
SPSS IBM, New York, USA	Software for basic statistics.

3.4.4 DNA sequencing of the intestinal microbiota

Our study of the intestinal microbiota in FGDS was limited to focus on bacteria and archaea, in terms of composition and diversity. For this purpose, we chose to perform targeted sequencing of the prokaryotic 16S rRNA gene. This gene has nine hyper-variable regions (V1-V9) (**Figure 3, panel A**) that show considerable sequence variation between different types of bacteria and archaea, and is thus a versatile marker of bacterial taxonomy. Considering that the length of each sequence (read) generated by current technologies is restricted, one must select which 16S rRNA hyper-variable region to sequence. This choice may introduce a bias, since the selected region may not show the necessary degree of variation to allow detection of certain taxa.[103] Addressing this issue we analysed our samples by two library preparation procedures, differing in regard to the 16S rRNA coverage.

Library preparation

In order to selectively increase sample 16S rRNA DNA for sequencing, PCR was used with primers flanking the relevant hyper-variable region. We created two collections (libraries) of sample DNA using primers flanking V3 –V4 (**Figure 5**) and V4 respectively.

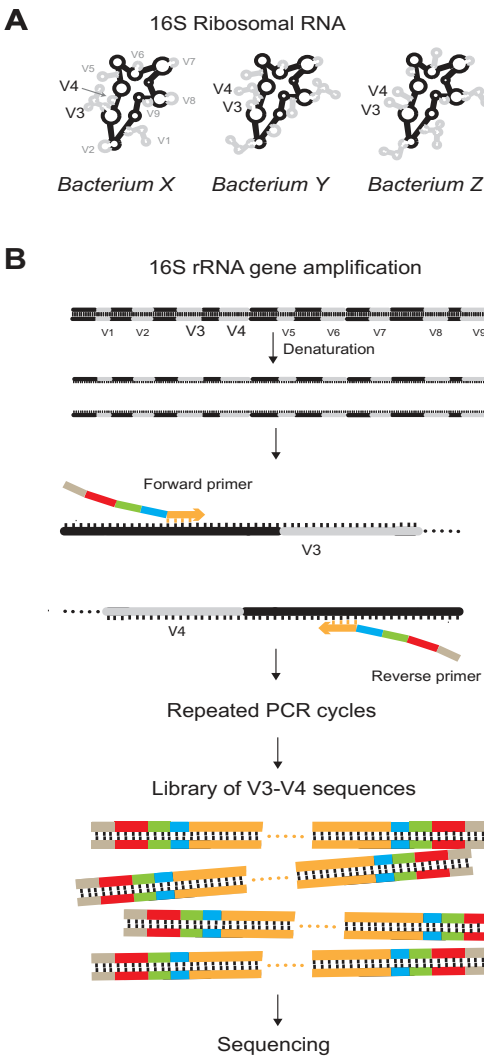


Figure 5

A: Conceptual illustration of 16S rRNA, with regions that have same (black) nucleotide sequence across different types of bacteria, but variable (grey) in 9 regions (V1-V9) of use as a marker of bacterial taxonomy.

B: The primer design used to amplify the 16S rRNA V3-V4 in the microbiota samples. Adapted from and with permission from Martin Kummen, thesis ("Primary Sclerosing Cholangitis and the intestinal microbiota" 2016)

Upon denaturation of DNA into 2 strands, the target sequence (V3-V4) is amplified using PCR with specifically designed primers.The primers used each contain 5 elements:

Genetic primers: targeting the conserved flanking regions of the hypervariable region of interest, for the forward and reverse primer, respectively. Linkers: selected to avoid homology with 16S sequences in a reference database. Pads: used to adjust the estimated melting temperature of the total sequence. Index regions: each combination in the forward and reverse primer is unique for each, serving as an identifier showing what sample the read originates from.Adapter sequences: these are designed to ensure that the fragments adhere to the flow-cell during sequencing.

Sequencing-by-synthesis (reversible-chain-termination) sequencing

We used sequencing-by-synthesis to sequence the 16S rRNA gene. This is one of the most popular principles of massively parallel sequencing and was also used for validation of the *GUCY2C* exome sequencing in paper I during the review process. Fragments of the target DNA are hybridised to oligonucleotides on a flow-cell and then amplified into clusters each containing approximately 1,000 copies of each fragment. To the flow cell are then added a mixture of single nucleotides chemically modified to block the 3´end and to emit a base-specific fluorescent signal upon ligation.[140, 141] A photo of the chip is taken and read by a scanner detecting which bases have been added at given positions. Next the unincorporated nucleotides are removed, 3´-ends chemically de-blocked and another cycle can start. While comparatively accurate, sequencing errors are prone to occur especially at the beginning and the end of the reads. However correct alignment of reads into longer sequences (contigs) is helped by many (parallel) reads generated for any nucleotide sequence, and by paired-end reads.[142]

Samples were sequenced on the Illumina MiSeq platform at the Norwegian Sequencing Center, Oslo. In order to validate our methods and assess the impact of primer choice we generated and sequenced V4 specific libraries in all FGDS, healthy relatives and a large subset of the unrelated healthy and IBD controls. DNA extracts from the same samples were also sent for library construction and sequencing according to the same protocol at an external institution (Broad Institute, Boston, USA). While we as expected observed an influence of primer choice on retrieval of certain taxa, we were able to validate V3-V4 findings in V4 data from both institutions.

Furthermore we compared duplicate stool samples provided by FGDS and healthy relatives with at least one week interval showing high reproducibility and stability (**Figure 6**).

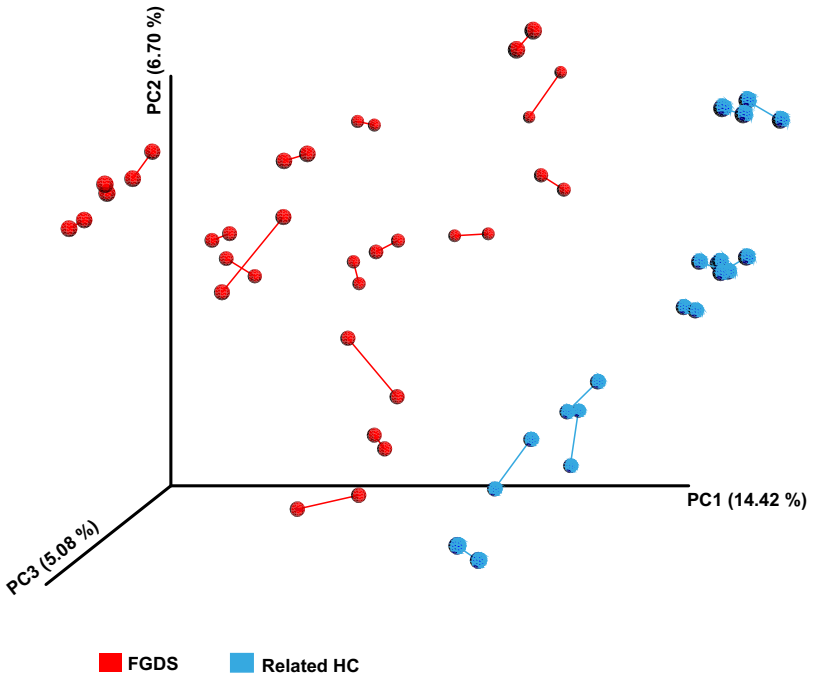


Figure 6. Consistency of data from duplicate samples. Illustration of beta diversity based on unweighted UNIFRAC- distances showing highly similar bacterial profiles (r^2=0.92, P<0.001) in the same individual across 2 samples (provided with at least one week interval). Samples from one individual are connected with a straight line. The illustration also shows that FGDS (in red) differ from healthy relatives (in blue).

Microbiota data processing with bioinformatics

While the human reference genome serves as a scaffold of help to assemble human sequence data, 16S rRNA sequence data in microbiota studies typically originate from a large number of species, and sequence data are evaluated against reference data that may be incomplete or still even erroneous for some taxa, e.g. by misalignment of sequences that originated from two different species into a hybrid (chimeric) sequence.[143] We used the Quantitative Insights Into Microbial Ecology (QIIME) bioinformatics pipeline, which integrates tools for the analytical steps, including processing raw data, quality control, sequence assembly, taxonomic assignment, visualisation and statistical analysis.[144, 145, 146] In order to assign sequence

data to microbial taxonomy and minimize errors e.g. due to chimera, we mapped reads against the16S rRNA reference sequence database GreenGenes.[143] Here we assigned sequence data to operational taxonomic units (OTUs), which represent clusters with 97 % sequence similarity retrieved from the GreenGenes database and are thus proxies of actual microbial taxa. The OTUs are the principal units for further analyses.

The number of mapped reads may vary across samples and if critically low, the within-sample diversity (alpha-diversity) may be poorly captured, and the sample might need to be discarded (cut-off at 8,750 in our data).

3.4.5 Functional analysis of the *GUCY2C* mutation

Site-directed mutagenesis

Based on the symptoms of the patients, we suspected a gain-of-function effect, and we settled cooperation with a scientific group led by prof. Visweswariah at the Indian Institute of Science, Bangalore, India that had established methods to study mutated *GUCY2C* in a cell-model.[147, 148] Complementary DNA of the *GUCY2C* variant was cloned into the mammalian expression vector pcDNA3, and together with non-mutant DNA transfected into HEK293 cells in order to study and compare the function of the mutant and the normal (wild-type) GC-C protein.

3.4.6 Manual curation of GC-C pathway gene set

For our analysis of IBD GWAS data in paper II we curated a gene set, aiming to capture the most important regulators of intestinal fluid shifts related to GC-C and CFTR signalling. We identified relevant genes using the search terms ”guanylate cyclase C” and ”CFTR” in Pubmed and GeneCards. The pathway is illustrated in **Figure 7**. While gene selection was validated using the Ingenuity database (https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/), the exact selection is to some degree subjective and the final number of genes arbitrary

and dependent on prior knowledge on gene function. This is an inherent limitation pertaining to grouping genes according to their function, and must be kept in mind when interpreting the results in paper II.[149]

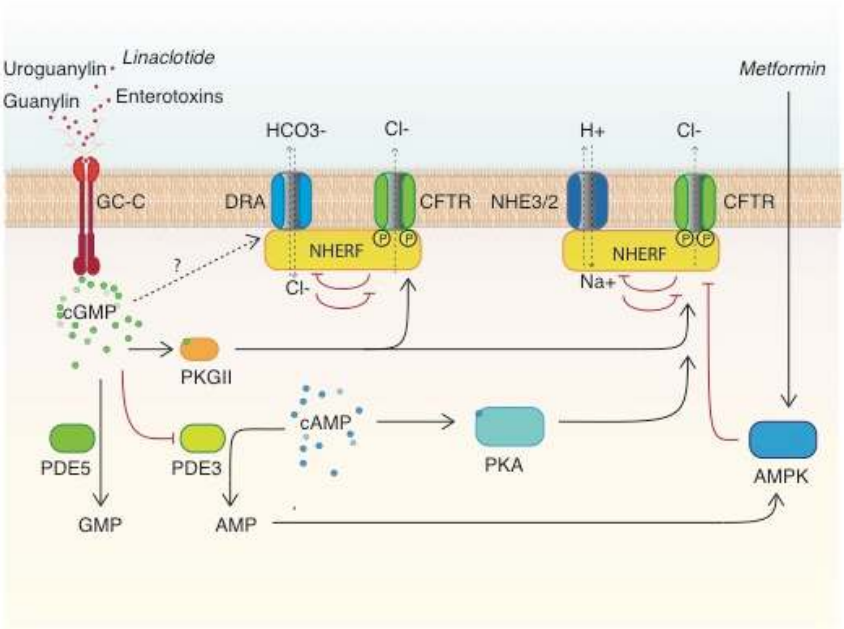


Figure 7 The GC-C —CFTR pathway

Illustration of guanylate cyclase C, cystic fibrosis transmembrane conductance regulator and interacting molecules that influence intestinal efflux of chloride (Cl-), sodium (Na+), bicarbonate (HCO3-) and water. (Adapted from figure 1, paper I [150] and Kato A et al. [151])

(c)AMP; (cyclic) adenosine monophosphate, AMPK; AMP-activated protein kinase, CFTR; cystic fibrosis transmembrane conductance regulator, GC-C; Guanylate cyclase C, DRA; Down-Regulated-in-Adenoma (encoded by SLC26A3), (c)GMP; (cyclic) guanosine monophosphate, NHE;sodium-hydrogen exchanger, NHERF; NHE Regulatory Factor, PDE; phosphodiesterase, PKA; protein kinase A, PKGII; cGMP-dependent protein kinase

3.5 Statistics

3.5.1 Basic statistical tests

For comparisons of continuous variables with normally distributed values we used the independent sample t-test. For non-normally distributed values comparisons were performed based on rank, i.e. the non-parametric Mann-Whitney U test. Comparisons of categorical variables were performed with the Chi-squared test (or Fisher's exact test as required). Depending on the data distribution, correlation analyses were performed either with Pearson correlation test (normally distributed data) or Spearman's rank correlation test (non-normally distributed).

3.5.2 Linkage analysis

For each typed SNP we calculated the Logarithm of Odds (LOD) for linkage to the unknown disease causing mutation taking into account the phenotype (here presence or absence of the diarrhoea) and the pre-test probability for linkage (around 2%). A LOD score >3 corresponds to a post-test probability of linkage >95%. Since a large number of markers were tested, we carried out this procedure (multi-point linkage analysis) using the Allegro version 2 software package (Decode Genetics, Reykjavik, Iceland) that also offer analytical solutions for common pitfalls, such as incomplete penetrance and phenocopies.[152]

3.5.3 Gene expression data analyses

Gene expression data generated by microarrays consist of a large number of variables representing differing overall gene expression levels. In paper II we compared gene expression in FGDS using a modified version of the t-test statistics, Significance analysis of Microarrays (SAM), which penalize small variance by adding a constant factor in the denominator.[153] Each gene is given a score (a proxy for the t-score), with significance determined according to how much the actual score deviates from what is expected given the 0-hypothesis (that the group assignment of participants does not affect the distribution of gene scores).

When comparing two distributions we need to take into account that gene expression data generally do not show normal distribution. We established the 0-distribution, i.e. the distribution of all possible gene scores given the actual data, by shuffling (permuting) the group labels 1000 times. We used this principle, called permutation analysis to assess significance also in the GSEA analyses as illustrated in **Figure 8, panel C**.

Key findings were validated ranking differentially expressed genes according to the "rank product" procedure.[154] We used the bioinformatics software package J-Express2012 for these analyses.[130, 155]

3.5.4 Enrichment analyses

The idea of enrichment analysis is to test whether groups of genes that may share some functionally defined role (e.g. as members of a particular pathway) are overrepresented (enriched) in experimental data. Enrichment analyses require the input of experimental data as a list of genes (e.g. from GWAS or expression analyses) upon which to query at least one pre-defined set of biologically related genes. There are several different statistical approaches out of which we used two; modular enrichment analysis and gene set enrichment analysis.

Modular enrichment analysis

Modular enrichment analysis allows experimental data to be analysed against pre-defined gene sets from various databases. We analysed gene expression data with this option in the **D**atabase for **A**nnotation, **V**isualization and **I**ntegrated **D**iscovery (https://david-d.ncifcrf.gov/home).[156] This freely available bioinformatics tool offers an online interface for analyses and a clustering algorithm that organizes gene sets originating from different biological databases into functionally related modules.[157] A modified Fishers Test statistic (EASE score) is used to find gene sets that are overrepresented in the datasets. The enrichment score for a module ("functional cluster") is then calculated according to *P*-values for gene sets within the cluster. For each gene set multiple testing adjusted *P*-value is generated using control for false

discoveries. A drawback with this method is that the input in the analyses is not ranked according to effect and is limited to an arbitrary significance threshold.[156]

Gene set enrichment analysis

Gene set enrichment analysis (GSEA) is an alternative approach, which employs ranking and avoids the threshold-restraint.[158] The gene expression data were ranked according to differences between FGDS and healthy controls, using signal-to-noise ratio.[159] Gene sets used in the analysis were retrieved from the GeneOntology database. (http://www.geneontology.org/GO.downloads.ontology.shtml).We used J-Express2012 software (Jexpress.bioinfo.no) to perform the analysis.[130, 155] The enrichment score (ES) obtained for each gene set is explained in **Figure 8.**

Gene set enrichment analysis in IBD GWAS data

In paper II we tested a gene set representing the GC-C pathway for enrichment in three publicly available GWAS datasets. The principle is illustrated in **Figure 8**, using our analysis of the IBD dataset as an example. Here**,** the order of genes on the ranked list was determined by scoring each gene according to GWAS *P*-values for its respective SNPs. The gene scores were corrected for the overall number of SNPs in the gene and their LD using a Brown score (Fisher based correction).[160, 161, 162]
The datasets were downloaded and analysed on the online GSEA-software tool provided by the Broad Institute (Boston, MA, USA) (www.broadinstitute.org/GSEA/index.jsp).

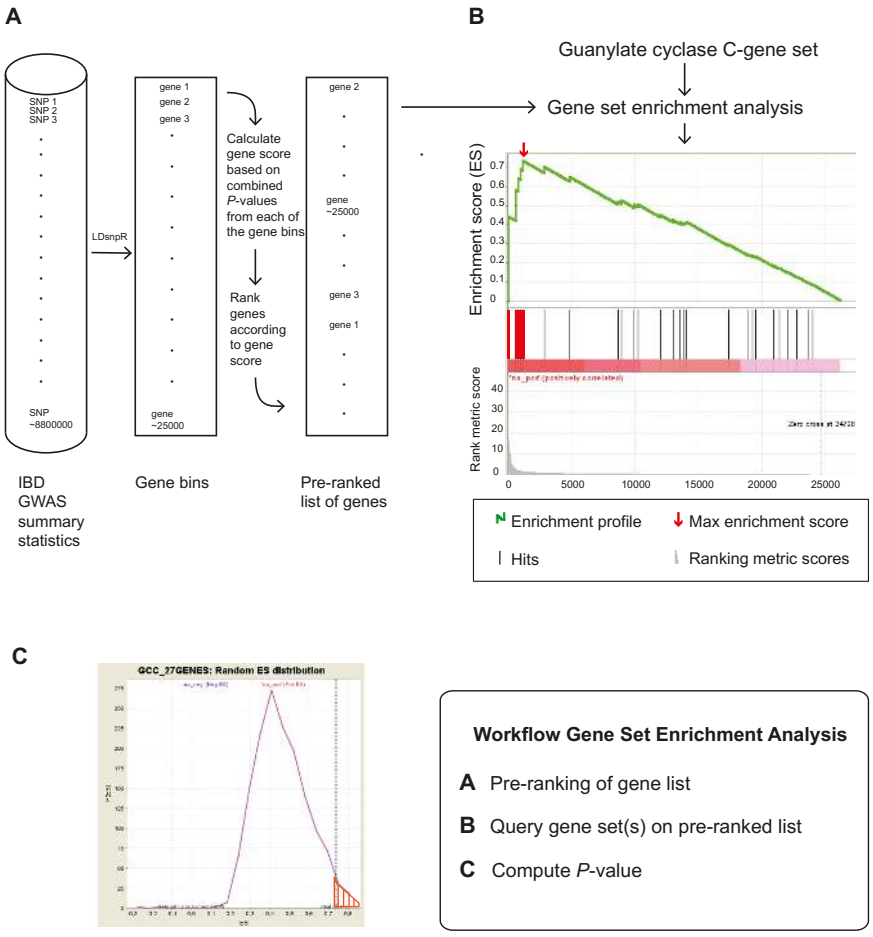


Figure 8. GSEA in GWAS data (inspired by Ersland KM et al. 2012 [162])

Panel A: A pre-ranked gene list was created using summary statistics from GWAS in IBD. Genes were scored and ranked according to P-values for their respective SNPs.

Panel B: The GC-C gene set was downloaded and queried against the pre-ranked list. A running enrichment score (ES) is computed, starting at the top and moving down the pre-ranked gene list. The ES increase according to the respective gene score every time a member of the candidate gene set coincides with a pre-ranked gene (shown as vertical bars "hits"), otherwise it decrease as shown by the enrichment profile (green). The peak of this profile marks the maximum ES, in this case 0.74.

Panel C: The permutation/0- distribution (purple line) of maximium enrichment scores was obtained running GSEA 1000 times shuffling the genes of the pre-ranked list. The P-value for the experiment is calculated as the area under the curve (red hatched) at and above its maxium ES (0.74, vertical dotted line). Here P = 0.024.

3.5.5 Microbiota analyses

We decided that genetic relatedness/distance between taxa should be accounted for when testing statistically whether the composition of bacteria in one group differed statistically from the composition in another. This was done by calculating the number of branches in the "tree of life" (UNIFRAC distance) that separated taxa only present in sample A from taxa only present in sample B etc.[163] The UNIFRAC distances were then used for visualisation of data, as shown in **Figure 6**, and for comparing statistically the microbiota composition (beta-diversity) between the groups. The structure of the data precludes the use of parametric tests, and we thus employed permutation based analysis of variance (PERMANOVA).[164] Permutation used in statistical significance testing is explained above (see 3.5.3).

For analysis of alpha-diversity (within sample diversity) we used the Shannon index, which accounts for richness (total number of different taxa) as well as the evenness of taxa. [165] We verified the results using other alpha-diversity measures.

In the statistical analyses regarding specific taxa associated with FGDS, we used non-parametric tests (Mann-Whitney) and corrected for multiple testing using the False Discovery Ratio (explained in section 3.5.6).[103]

IBD status in FGDS was an important co-variable that we accounted for in the subsequent analyses. For evaluation of the impact of co-variables we used two freely accessible bioinformatics tools, LEfSe and MaAsLin (http://huttenhower.sph.harvard.edu/galaxy/) that have been developed specifically for analyses of the microbiota. The former employs a combination of non-parametric tests and the Linear Discriminant Analysis principle to identify the features in the data that best explain differences between groups with consistency (in this case regardless of IBD status) and assign an estimated effect size.[166] We used the graphic options offered by LEfSe to illustrate our findings in paper III.

Next IBD status and other factors associated with the microbiota (age, number of antibiotics courses last year and BMI) were included as co-variables when comparing

the groups using MaAsLin, which is a version of linear regression.[167] The taxa that remained significant throughout this sequence of tests were ultimately validated comparing FGDS without IBD to healthy controls using Mann Whitney test.

3.5.6 Statistical significance and multiple testing

This thesis reports the results of many statistic tests, with a high risk of false positive observations (type I errors). In a single experiment reported in paper II, more than 25,000 variables were compared between two classes represented by a total of 27 samples. As many as 1250 (5%) observations might here be expected to pass the nominal significance threshold at alpha=0.05 by coincidence. One common way to adjust significance thresholds according to number of observations is the Bonferroni correction, using the formula alpha = 0.05/number of tests. However, the Bonferroni method has been criticised for being too conservative and may increase the number of false negative findings (type II errors).

One popular approach to this problem, which we adopted, is to rank all observations (discoveries) according to their nominal *P*-values, and for each estimate the false discovery ratio (FDR /Q). The FDR is computed as the proportion between the number of actual discoveries for a given nominal *P*-value threshold and the number of the expected discoveries given the 0-hypothesis.[168] In the present thesis, $P<0.05$ is considered statistically significant unless otherwise indicated by the FDR correction.

3.5.7 Power estimates

Given the scarcity of FGDS patients, considerations about the expected magnitudes of effects to study and the number of samples needed to detect these effects were an important part of planning the current projects. Power estimates help to predict the level of confidence that can be obtained to avoid pointless analyses. However, when studying a novel and rare disorder, available knowledge on expected effects as well as individuals eligible for the study are necessarily limited.
When we decided to carry out linkage analysis, eleven family members had been categorized as affected and 14 as unaffected. Power calculations showed that

genotyping only the affected might suffice to identify significant disease linkage with maximum LOD score >3.

For the gene expression study (paper II) and the microbiota study (paper III) we estimated a minimal sample size of 10 participants per group. We presumed that biological effects of the *GUCY2C* mutation are evident in a relatively small dataset, possibly intermediary between what can be observed in genetically modified animals (where sample size of <6 may suffice) and in complex human disorders where larger numbers are needed.[169]

Genetic risk variants generally show low effect size (odds ratio typically 1.5 and below), making it unlikely that statistically significant difference could be detected.[170, 171] On the other hand the impact of common genetic variants could be stronger when present in individuals already at high risk, as for the FGDS patients. We therefore reasoned that a targeted and descriptive assessment of IBD-risk variants in FGDS patients appeared biologically well funded, meriting the exploratory study performed.[37]

In our assessment of dietary habits in FGDS and healthy relatives, we calculated 80 % power to detect 10-15 % difference in intake of a single macronutrient recorded with food frequency questionnaires if provided by all eligible family members. With the many nutrients under assessment, this study was clearly not powered to detect subtle differences in food preferences after corrections for multiple comparisons. Obtaining the dietary data was still prioritised since it could help to detect overt group differences for this important confounder.

3.6 Ethical considerations

The Regional Committee for Medical and Health Research Ethics of Western Norway (reference 2009/902), or the Regional Committee for Medical and Health

Research Ethics in South-Eastern Norway (reference number 2012/286b) approved the studies conducted throughout this project.

During genetic testing, we could incidentally discover genetic changes predicting future unpreventable and untreatable ailments and unrelated to the disorder of interest. Genetic analyses were however targeted either for common variants with presumably little impact, or towards small regions of the genome, and taken together this minimized the potential risk for incidental findings.

Pedigrees displayed in paper I and II constitute a step away from an ideal of full participant anonymity but were considered crucial for illustrating the nature of our study and its findings. The participants had been informed in advance and approved the use of pedigrees in the scientific publications.

All participants or their parents were thoroughly informed about the investigations and had given their spoken and written consent.

4. Summary of findings

4.1 Paper I

4.1.1 Patient demographic and clinical characteristics

From the pedigree shown in **Figure 4**, we studied a total of 32 family members. The clinical characteristics of the patients define the diagnostic features of FGDS. The patients typically had onset of diarrhoea in early infancy, and eight family members had been hospitalized within one month from birth due to dehydration and metabolic acidosis.

Increased intestinal gas was reported in most of the patients, and several patients also reported recurrent abdominal pain. These symptoms resemble IBS, and in fact four family members had previously received this diagnosis. While these individuals did not fulfil IBS criteria as assessed by the current symptom questionnaire, five other cases did fulfil these criteria.

Ten family members underwent surgery because of suspected bowel obstruction and in eight of them this was confirmed. The aetiologies of intestinal obstructions as evaluated by the surgeon were classified as either or a combination of: coecal volvulus, internal hernation, ileal inflammation, and adhesional bands. All the confirmed cases of bowel obstruction involved the distal small bowel and/or the proximal large bowel. Intensive care due to multi-organ failure was needed in three patients following bowel obstruction. One relative had bowel problems and died in infancy. Another relative with chronic diarrhoea died at age 50 following surgery for bowel obstruction. There is no other information indicating premature death.

Seven patients had been treated for suspected IBD, in all cases affecting the terminal ileum +/- the colon. Median age at CD presentation was 36 years.

We found reports on dilatation of the small bowel in absence of obstruction in several patients and long segments of gross dilatation in three patients, consistent with intestinal dysmotility.
Deficiency of vitamin B12, esophagitis and urolithiasis had been reported in six, five and four patients respectively.

4.1.2 Genetic characteristics of FDGS

We found linkage to FGDS on the short arm of chromosome 12 with a LOD-score of 5.1. This region comprised 28 putative protein-coding genes, including *GUCY2C*, known for its role in bacterial toxin induced diarrhoea. By DNA sequencing, we discovered a novel missense variant, c.2519G→T in exon 22 of *GUCY2C*, representing an amino acid substitution from serine to isoleucine in position 840 in the catalytic domain of GC-C. We found that baseline activity was unaffected by the mutation, but the variant GC-C responded with higher activity than the wild-type GC-C upon ligand stimulation. We also found that the amount of uroguanylin needed to stimulate the mutant receptor is lower than in the normal receptor. This suggests that the concentrations of uroguanylin present in the intestine could result in abnormally elevated levels of cellular cGMP in intestinal cells harboring the mutant receptor (gain-of-function).

4.2 Paper II

4.2.1 Assessments of IBD risk loci in FDGS patients

We found that the candidate set of genes representing members of the GC-C - CFTR pathway displayed significant enrichment of association in IBD case-control genetic data but not in CD or UC separately.

The ratio of particular risk allele carriage between FGDS patients with and without CD showed minimal deviation from 1.0 (i.e. no difference between the two groups) except for rs5743289 in the *NOD2* gene (**Figure 9**).

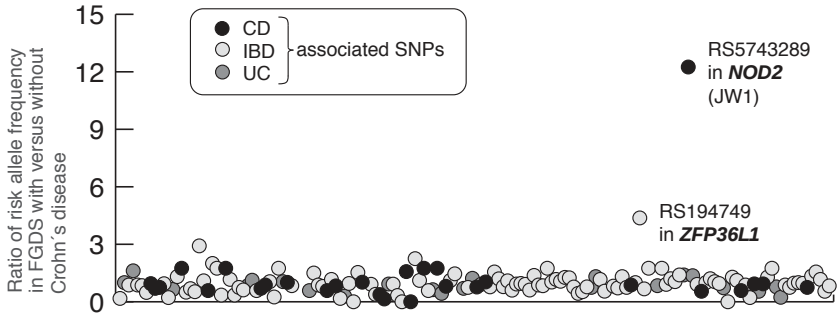


Figure 9. The risk allele ratio in FGDS with (n=8) versus without IBD (n=14). Each spot represents one of 163 IBD risk alleles. RS5743289 in NOD2 is a substantial outlier with 12.3 fold higher frequency in FGDS patients with IBD than in those without IBD.

4.2.2 Global gene expression changes in the distal small bowel mucosa of FGDS patients

We investigated global gene expression in biopsies from the terminal small bowel in FGDS patients, unrelated CD patients without ileal inflammation and unrelated healthy controls.

Twenty-one genes were significantly differentially expressed when comparing FGDS to healthy controls. Six genes encoding metallothioneins were down regulated in FGDS patients, and four of these were verified using qPCR. Qualitative assessment of biopsies using immunohistochemistry suggested that transcriptional down regulation of metallothioneins in FGDS was translated also to the protein level. Metallothionein related terms prevailed in pathway analyses of the gene expression data. We considered tissue zinc depletion as a possible explanation for these findings, but no differences were observed in plasma zinc between FGDS and healthy controls. Metallothionein down regulation in intestinal biopsies from unrelated cases with IBD has been reported, but we were not able to reproduce this observation in our material.[172, 173]

4.3 Paper III

4.3.1 Stool calprotectin and bacterial characteristics of FGDS patients

Of the 20 FGDS patients in this study, five had concomitant CD. In this group, disease activity as assessed clinically was comparable to unrelated CD, while stool calprotectin was significantly lower.

We found that the faecal bacterial composition differed in FGDS patients as compared with unrelated as well as related healthy controls and unrelated patients with IBD. The gut microbiota of healthy relatives of FGDS patients was similar to that of unrelated healthy controls. Alpha diversity in FGDS patients was similar to healthy controls, whilst significantly reduced in IBD as compared with FGDS patients and healthy controls.

We found that 28 microbial taxa were changed in FGDS patients compared to healthy controls when stratifying for concomitant CD and controlling against healthy relatives. The quantitatively largest differences were seen within the phylum Firmicutes, including a marked depletion of the genus *Faecalibacterium*. Loss of *Bifidobacterium* in the FGDS patients was near to total. On the other hand, Enterobacteriaceae and methane producing archaea were increased in FGDS.

Intake of fat, protein, carbohydrates, fibre, added sugars and alcohol as reported in food frequency questionnaires was similar in FGDS patients and healthy relatives.

4.3.2 Microbiota associations with IBD in FGDS

We found that *Faecalibacterium* was significantly lower in FGDS patients affected by CD than in those unaffected by CD, but still lower in FGDS patients without CD as compared with healthy relatives. There was a tendency towards loss of alpha-diversity in CD affected versus non-CD affected FGDS patients, but the difference was not statistically significant. We found that diarrhoea was associated with CD status and correlated with loss of alpha diversity and loss of the archaea genus *Methanobrevibacter*.

5. DISCUSSION

5.1 FGDS - a novel inheritable disorder

The experiments performed in paper I leave little doubt that the c.2519G→T variant in *GUCY2C* is the cause of familial diarrhoea in the Norwegian family we studied. Deletion of *Gucy2c* by genetic engineering had previously been performed in mice as part of general studies of the functions of this gene, but our work was the first to report of a human disorder caused by a mutation in *GUCY2C*.[2] The inheritance pattern and genetic findings strongly suggested this was a novel monogenic disease, and the medical literature provided no similar entity with the overall constellation of symptoms and complications that we observed in our patients. We named the disorder familial *GUCY2C* diarrhoea syndrome - FGDS. In the following, I will discuss opportunities, interpretations and limitations pertaining to our work along four principal directions:

1) How to diagnose FGDS in other patients?
2) What can FGDS teach about the function of *GUCY2C*?
3) Can FGDS provide insights of general relevance and relevant to other diseases?
4) How should we manage FGDS patients in the clinic?

Lastly, I will provide some suggestions for further research building from our studies.

5.2 Defining and diagnosing FGDS

Throughout the papers presented in this thesis, the diagnosis of FGDS was based on carriage of the c.2519G→T mutation in *GUCY2C.* While mutation carrier status is a logical diagnostic test that correlated closely with the presence of diarrhoea, some criteria could help the clinician to decide when to suspect and how to proceed in diagnosing FGDS in a patient. Before defining these criteria I will outline the main clinical observations that they rest on.

Diarrhoea

Reported stool frequency ranged from 0.5 to 20/day, showing that diarrhoea is not a constant trait in mutation carriers. The reported age of onset of diarrhoea varied, and holds intrinsic recall bias as to the memory of the participant or a parent. The overall ratio of hospitalisation for dehydration within one month of birth was approximately 30%, but more than 50% among the younger participants, supporting that diarrhoea presents in very early age. In those hospitalized in neonatal age for dehydration, improvement occurred in most within a couple of days to weeks, an observation that differs from other congenital diarrheal disorders where neonatal dehydration may herald severe continuous diarrhoea.[30]

Meteorism

Almost all FGDS patients reported abnormal amount of intestinal gas, a rather subjective symptom, but this symptom was often felt socially disturbing due to noise or associated soiling of underwear and the symptom seems required for a diagnostic algorithm of FGDS. The meteorism could be explained by intestinal dysmotility.[174, 175]

Inflammatory Bowel Disease

Depending on how strict criteria we used, 20-25 % of adult FGDS patients have developed IBD. While the absolute numbers are small, our data indicate that young people with FGDS are at increased risk for developing IBD, categorized as or resembling CD, and diagnostic criteria for FGDS must account for this complication. Delineating potential pathophysiological aspects of the IBD manifestation was a major emphasis of this thesis.

Small bowel obstruction

Among adult FGDS patients, one in three have experienced small bowel obstruction needing surgery at least once in life and diagnostic criteria for FGDS must include the possibility of this complication. Apart from a strong tendency to involve the distal small bowel, no universal pattern was seen in the surgical reports. Observations made

during surgery support that a tendency towards intestinal torsion results from impaired intestinal motility. Potential mechanisms for dysmotility and bowel obstruction are discussed in chapter 5.3.

Heritability

In our study chronic diarrhoea was inherited in an autosomal dominant pattern with penetrance at >30% by one month of age and 100% by adolescence. The c.2519G→T variant in *GUCY2C* may arise *de novo*, and absence of relatives with chronic diarrhoea does not rule out FGDS. Other variants in *GUCY2C* might also give rise to diarrhoea with a similar mechanism but a different pattern of inheritance. When three brothers in the current family were evaluated 20 years ago, the geneticist noted that their history resembled CSD,[176] but differed from this disorder that was thought to be autosomal recessive, by showing autosomal dominant heritability and a milder phenotype.[177]

A couple of years after our first paper was published, Janecke and coworkers showed that *de novo* heterozygote gain-of-function variants in *GUCY2C* or homozygote / compound heterozygote loss-of-function variants in the *SLC9A3* gene (encoding NHE3 immediately downstream of *GUCY2C*) cause CSD.[178, 179] These patients generally showed more severe diarrhoea than FGDS, and earlier development of IBD and intestinal obstructions.[179] The severity of presentations correlated with the functional effects of different activating mutations in *GUCY2C.* These findings suggest that CSD and FGDS could be regarded as variants of the same disorder.

Terminology

In this work we have agreed to maintain FGDS as a distinct disorder, but acknowledge that the terminology might be reconsidered. The name Familial *GUCY2C* diarrhoea syndrome would not include disorders caused by variants in *SLC9A3.* Dysmotility was reported in FGDS as well as CSD caused by *GUCY2C* variants but not in CSD patients with *SLC9A3* variants. However, intestinal dysmotility may not be easy to diagnose, and whether it is a useful clinical criterion

for distinguishing one entity from another could be questioned. If we in the clinic opt to classify FGDS as variant(s) of CSD, it could be useful to subdivide the entity according to age of onset: neonatal, infant, juvenile. The CSD entity should still be maintained distinct from other diarrheal disorders with sodium loss, such as the type associated with congenital malformations and variants in *SPINT2,* as well a Microvillus Inclusion Disease, caused by variants in *MYO5B*.[179, 180]

Genetic testing

It seems relevant that genetic testing for suspected FGDS should include scanning of the *SLC9A3* gene. In fact, I would propose that for all conditions with a presentation form similar to FGDS and CSD, genetic testing for *GUCY2C, SLC9A3, CFTR*, *SLC26A3, SPINT2* and *MYO5B* should be performed as a gene panel assessment.

Proposed diagnostic criteria for FGDS:

Major criteria:

1. Chronic loose stools with frequency ≥ 3 times a day lasting more than 3 months, onset in childhood
2. Neonatal onset diarrhoea with dehydration (loss of >10% of birth weight)
3. Increased loss of sodium in the stools.

Minor criteria:

1. At least one first-degree relative affected by chronic diarrhoea
2. Intestinal gas at levels perceived to be socially disturbing
3. Neonatal dehydration with hyponatraemia
4. Spontaneous improvement of neonatal diarrhoea
5. Dehydration needing hospitalization during intercurrent gastrointestinal infections
6. IBD following a long period (>1 year) of chronic diarrhoea
7. Small bowel obstruction with signs of chronic intestinal dysmotility
8. Ileocoecal volvulus
9. Any one factor of minor criteria 3.-8. in a first-degree relative

These criteria represent characteristics of use to suspect FGDS but with a large number of different inheritable diarrheal disorders, a final diagnosis should only be made in support of genetic variation compatible with gain-of-function in *GUCY2C*.[30] For patients that I meet in the clinic I have used the rule of thumb that at least one of the major criteria, and three of the minor criteria should be fulfilled prior to initiating genetic testing for FGDS. This rule has however not been formally evaluated as to its sensitivity and specificity. As in any clinical decision, sound judgement should be exercised. Irritable bowel syndrome (IBS) is a common disorder that run in families and may resemble FGDS, but unlike the latter, IBS tend to present later.[181] In patients with relatively late-onset diarrhoea, I would recommend that at least four minor criteria are needed to support suspicion of FGDS and the initiation of genetic testing.

5.3 What do our studies tell about *GUCY2C* function?

The involvement of GC-C (as encoded by *GUCY2C*) in diarrhoea was discovered almost 40 years ago, following the efforts to identify the receptor for ETEC toxins.[182] Apart from the toxins, GC-C is also activated by small peptides, guanylin and uroguanylin secreted from the intestinal wall.[183, 184] As the GC-C receptor itself, these peptides are mainly located along the small and large bowel lining.[185, 186] Regulation and effects of GC-C activation have been outlined in subsequent studies using cell-models and genetically engineered mice.[148, 187, 188, 189, 190] Secretion of the guanylins is regulated by nutrients, and GC-C seems to be part of a sensing-effector circuit involved in feeding.[191]

The genetic models provide extensive opportunities to carry out experimental interventions and strictly control for effects of confounding environmental factors such as microbes and nutrients. The model-organism principle was exploited in paper I where experiments in a cell-line revealed that the c.2519G→T mutation in *GUCY2C* results in increased activation of GC-C and that the increased activation depends on stimulation from the guanylins.

Unlike the animal and cell-based model systems, monogenic disorders such as FGDS offer the opportunity to record symptoms and complications that occur over the life span of a complex human organism, and may provide new and unexpected insights in the biology of single genes (a human model system).[3, 192] As such, the present study of FGDS allowed us to gain insights in the long-term effects of GC-C activation in humans, and also to perform proof-of-concept interventions. Although patient size is limited, such proof-of-concept interventions may further inform the pathophysiology and open for treatment of even more common disorders with similar symptoms. In our case, an observational intervention with metformin had been tried on the basis of its potential to activate AMPK and thereby impair chloride efflux through CFTR (**Figure 7**) with some possible beneficial effects.

Our observations suggesting dysmotility, i.e. perturbations of gut peristaltic movements in FGDS, motivated the extensive work on this topic performed by dr. Von Volkmann.[174, 175] Patients with FGDS display increased non-occlusive small bowel contractions and delayed colonic transit.[174, 175] Intestinal motility is regulated through a complex and still insufficiently understood crosstalk involving mechano- and chemo-sensors in the gut, the entero-endocrine as well as the autonomic and the enteric nervous system. [193] Supporting that the GC-C pathway interacts with the enteric nervous system, FGDS patients display different circulating levels of the neurotransmitter serotonin than healthy following a meal.[175] However the GC-C activating ligands stored in various cell types throughout the gut show little degree of co-localization with serotonin-secreting cells.[194]

Intestinal dysmotility has to our knowledge not previously been linked with *GUCY2C,* and this finding emphasizes the multiple functional impacts of this gene. Considering GC-C as a member of the guanylate cyclase family, the finding may not be surprising. Guanylate cyclase-A, is one membrane-bound member of this family, expressed in the heart and blood vessels and a key regulator of vascular smooth muscle tone.[195]

Cyclic GMP generated by soluble guanylate cyclase (another member of the family) decreases smooth muscle tone in blood vessels as well as the gut by facilitating calcium transport.[196 197] There is evidence of a cross talk between membrane bound GC and soluble GC, which could provide an explanation for the effect of GC-C activation on intestinal motility.[198] Also cGMP generated in the epithelium by GC-C may be transported across cell-boundaries and thereby regulate intestinal smooth muscle contractility more directly.[197]

Dysmotility probably explains delayed intestinal transit, gas and susceptibility to twisting of the gut in FGDS. This is the extreme of a continuum, and effects from *GUCY2C* under normal physiological conditions likely serve beneficial purposes. GC-C signalling appears to be coordinated according to feeding patterns, and could help nutrient uptake by tuning peristaltic movements according to the need for propulsion yet assuring optimal absorbtion.[191, 199, 200] This is in line with the notion that GC-C activation promotes food absorbance also by moderating the mucus barrier.[191]

In the biopsies from the distal ileum of FGDS patients in paper II, only a modest number of genes were differentially expressed compared to healthy controls. Down-regulation of metallothionein-genes was a striking observation in these data. The metallothioneins are low molecular proteins with strong antioxidant and metal-binding properties due to a high content of cysteine residues. While protecting the organism from metal toxicity and oxidative stress, metallothioneins are also crucial for storage and shuttling of physiologic metals, most importantly zinc.[201] The availability of zinc is a key regulator of tissue metallothionein expression, and we asked whether the gene-expression findings might mirror zinc depletion e.g. due to diarrhoea.[201] Zinc depletion could amplify GC-C activation and thus aggravate diarrhoea in FGDS, but while shortage of tissue samples forbade the optimal assessment of this matter, such an explanation was neither supported by plasma zinc measurements nor any reported clinical response to oral zinc supplements given to a few patients.[202]

Tissue architecture may impact on global gene expression recordings in whole-tissue samples due to cell-type specific transcriptional activity.[19] For example GC-C activation could promote cell-differentiation and decrease epithelial proliferation, whereas intestinal inflammation may increase epithelial proliferation.[4, 203] In our study, deviations from normal architecture (fibrosis/granulation tissue/inflammation) were only seen in three samples (FGDS patients with CD). When testing for bias by performing an analysis without these samples, findings remained stable.

We believe that regulation of metallothionein expression by physiological levels of GC-C activity may serve a signal transducer function influencing epithelial zinc and redox status,[204] and thereby host-microbiota interactions.[205, 206]

The differences in overall microbiota composition in FGDS compared to controls we found in paper III did suggest a role of *GUCY2C* in shaping the intestinal microbiota. While *GUCY2C* has not figured among the genetic factors associated with the intestinal microbiota in population-based studies, clearly these studies hold major limitations as shown by little overlap in gene-microbiota associations between the different studies.[107, 207, 208] On the other hand, population studies have shown stool consistency to be one of the strongest determinants of the intestinal microbiota, supporting that even small transepithelial fluid shifts induced by GC-C may influence intestinal ecology.[209, 210] GC-C induced ion transport across the gut epithelium may also impact on the intestinal microbiota by changing the pH.[211] We have indeed verified that the intestinal lumen is more alkaline in FGDS compared to healthy controls, suggesting another factor linking GC-C activation and the intestinal microbiota.[175, 211]

Taken together our findings align with mounting evidence in human as well as animal studies indicating that GC-C signalling may serve important host benefits, including the defence against pathogens by regulation of the intestinal microbiota.[212, 213]

The involvement of GC-C in intestinal immune function has been suggested by studies in *Gucy2c* knock-out mice, which show increased susceptibility to invasive enteric infections, malignancy and inflammation.[4, 212, 214] A finely balanced GC-C

signalling may be especially important in the neonatal period.[186] In contrast to FGDS, children with GC-C inactivating mutations experienced neonatal severe constipation (meconium ileus) mimicking CF and retained liability to severe infections throughout infancy.[215, 216] Some of this liability can been attributed to perturbations of epithelial differentiation and barrier function (including the mucus layer) in the absence of appropriate GC-C signalling.[212, 213, 217] Our studies of FGDS has added a new piece to this picture, by showing that also increased GC-C signalling may adversely affect immune function and contribute to the development of intestinal inflammation in humans.

5.4 Can FGDS provide insight of general relevance and relevant to other diseases?

Although afflicted by a rare disorder, FGDS patients display clinical similarities with the common condition IBS with diarrhoea (IBS-D), and are liable to develop other common clinical conditions such as IBD and small bowel obstruction. Monogenic disorders may overlap with common disorders regarding single genes and pathways involved,[3, 218] and the demarcation between monogenic and complex diseases is most appropriately considered a continuum **(Figure 2).** Along this concept, studies of monogenic disorders may help to elucidate mechanisms of relevance also to complex disorders involving similar pathophysiological components or similar symptoms. Such studies thus hold general relevant, and may ultimately pave the way for therapies of relevance far beyond the monogenic condition in which they are performed.[3] My main emphasis has been the identification of factors relevant to understand why some FGDS patients develop IBD, and whether these factors are shared with IBD patients without FGDS.

Inflammatory Bowel Disease

In the gene set enrichment analysis (reported in paper II) we found that genes belonging to the GC-C signalling pathway were overrepresented among the GWAS findings in IBD. Rather than considering association statistics of single genes, but as members of biological pathways, we found this method appropriate to examine

whether the GC-C gene set may overlap with pathways involved in regular IBD.[99, 218] As noted above, gene set selection may be biased and the putative involvement of genes in more than one pathway (which commonly is the case) may confound GSEA outcomes.[102] Regardless of the GSEA, even individual members of the GC-C pathway were associated with both types of IBD at genome-wide significance levels (i.e. *SLC9A3, SCL9A2, SLC26A3* and *PRKAR2A, PRKAA1*), supporting our observations from the GSEA that IBD in FGDS shares pathophysiological mechanisms with conventional IBD.

A notable segregation of risk variants in *NOD2* was found in FGDS patients with concomitant IBD (**Figure 9)**. *NOD2* encodes a cytosolic receptor for bacterial patterns and unfolded proteins. Coordination of the responses to these exposures e.g. engulfment and neutralization by autophagy are key functions of *NOD2* relevant to inflammatory pathways.[63] Variants in *NOD2* are the strongest genetic risk factors for CD and are mainly associated with the ileal subtype, i.e. a phenotype similar to that observed in FGDS patients with IBD. [68, 219, 220, 221] Since the FGDS patients are closely related and therefore share a common genetic background, segregation of genetic variants beyond *GUCY2C* may occur by chance and unrelated to FGDS pathophysiology. However, in the genotyping of IBD risk variants, *NOD2* was a solitary and substantial outlier. This observation is supported by the fact that the *NOD2* variants appear to originate from five independent founders. Further supporting a disease modifying role of *NOD2* in FGDS, and even suggesting an allele dosage effect, the only two FGDS patients homozygous for *NOD2* risk variants displayed a particularly severe IBD phenotype requiring extensive surgical interventions. All but one FGDS patient carrying *NOD2* risk variants had been diagnosed and treated for IBD before genotyping, minimizing the risk of ascertainment bias.

Metallothionein down-regulation has previously been reported in studies of the intestinal mucosa of patients with IBD, and its occurrence in FGDS raises the question whether it may represent a common factor of the inflammatory manifestations in IBD and FGDS.[172, 173, 222, 223] The number of unrelated disease

controls with CD was limited, and we were only able to determine a non-significant trend towards metallothionein down-regulation in our conventional CD samples. Epithelial metallothionein expression may be affected by the intestinal microbiota, but may also serve as a signal transducer important for the epithelial clearance of bacteria.[224, 225, 226, 227] The underlying biology of metallothionein involvement in gut inflammation may indeed relate to its role as a mediator between NOD2 bacterial sensing and subsequent bacterial removal by autophagy.[227] Impaired autophagy favours an inflammatory response to bacteria, and is an established risk factor for development of CD.[62, 64] Metallothionein down regulation and *NOD2* risk variants may thus in sum constitute additive liability factors for CD, linked to GC-C activation in FGDS, potentially providing an explanation for the clinical affections.

In paper III we reported that FGDS patients display changes in the gut microbiota that could pose a relevant contribution to development of IBD. Most strikingly, FGDS patients displayed remarkable loss of the bacterial species *F. prausnitzii* and increased abundance of Enterobacteriaceae. These changes are often seen in CD, but while both can be speculated to contribute to inflammation, they may also both be secondary markers of inflammation and gut barrier alterations.[64, 92, 93, 228, 229] This "chicken and the egg problem" is inherent to the interpretation of all cross-sectional data on the intestinal microbiota. In our study, stratification of FGDS patients according to IBD status combined with stool inflammatory markers helped to control for the potential confounding effect of inflammation, however similarities between FGDS and IBD could partially be explained by the fact that both groups are affected by diarrhoea, which is associated with total loss of bacteria and increased oxygen tension, again favouring an ecology enriched for Enterobacteriaceae and disfavouring *F. prausnitzii.*[206, 230]

Similar to FGDS, diarrhoea in IBD may partially depend on impaired sodium/hydrogen exchange through NHE3.[20, 231] Whether this is a primary or a secondary event in IBD may be hard to settle, since this ion channel display complex multidirectional relationship with the intestinal microbiota, microbial metabolites and inflammation.[20, 151, 232] In line with findings of ileal and colonic inflammation in mice

deficient of NHE3-activity, our studies support that transepithelial fluid shifts may initiate and sustain inflammation. Such pro-inflammatory effects may be mediated by changes in the intestinal microbiota.[233, 234] Finally, other human studies support that alterations of the intestinal microbiota may precede development of IBD in individuals with genetic risk.[235, 236, 237] In our data, the similarities between FGDS patients with and without IBD, may indeed indicate that intestinal microbiota alterations precede the intestinal inflammation, but longitudinal studies are needed to confirm this possibility.

Bowel obstruction

The high lifetime incidence of acute bowel obstruction observed in FGDS and CSD has to our knowledge not been reported in other diarrheal disorders, suggesting that it is a complication specifically linked to GC-C signalling. We were particularly intrigued by the constant location (ileocoecal junction), but apparently variable aetiology of small bowel obstruction (adhesional bands, inflammation, volvulus).

Torsion of the intestine at the junction between small and large bowel has been seen during surgery in several FGDS patients and characterised as coecal volvulus in three patients. This kind of intestinal obstruction may resolve if the bowel untwists in a timely manner.[238, 239, 240] Recurrent intestinal torsion could explain spontaneously resolving episodes of strong abdominal pain needing hospitalisation in some FGDS patients. Apart from a freely movable ileocoecum observed during surgery, the tendency to develop volvulus in FGDS is likely explained by GC-C related gut dysmotility, which may predispose for twisting and impair untwisting of the gut.[178, 241]

Intestinal torsion may eventually lead to compromised blood flow and intestinal ischemia and potentially to a so-called ischemia/reperfusion (I/R) injury.[242] Reperfusion of hypoxic tissues following surgical treatment releases reactive oxygen species and inflammatory mediators creating inflammation and scars locally, and sometimes a systemic inflammatory response with multi-organ failure.[243, 244, 245]

Twisting of the coecum would not necessarily require surgery if spontaneous untwisting takes place, but an I/R injury could still be sustained, resulting in epithelial apoptosis, bacterial translocation across the bowel wall, inflammation and/or fibrotic adhesions. It could be speculated that such events induce different types of pathology observed in FGDS, including intestinal inflammation, as illustrated in **Figure 10**.

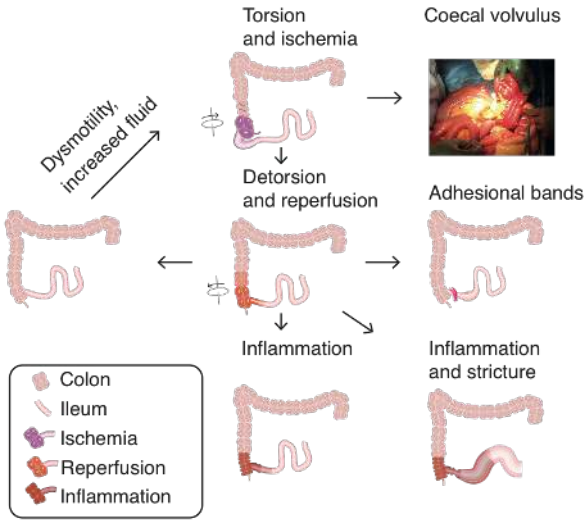


Figure 10. Concept of ischemia / reperfusion injury due to intestinal torsion as an event leading to variable pathology observed in the ileocecal region in FGDS. Photo used with permission from the patient.

That hypoxia could have general relevance in the pathogenesis of IBD has been suggested e.g. by an association between high-altitude travelling and disease flares, as well as over-expression of hypoxia inducible pathways in animal IBD models.[246, 247] Interestingly the liability to develop injury after I/R may be modulated by the intestinal microbiota through NOD2 signalling and autophagy.[248] We noted that among FGDS patients with intestinal obstruction, *NOD2* variants demarcated those that also developed inflammation. We are unable to determine whether metallothionein down regulation impacts on autophagy in FGDS, but we note that metallothionein up regulation has been associated with better outcomes after I/R injury.[249] Moreover animal models have suggested a role of guanylate cyclase signalling in modulation of the I/R injury.[250] GC-C impacts on the mucus layer, the

fluid status as well as cellular energy utilization, which may all be critical factors in the I/R injury.[7, 245] The fluid losses that follow ion-channel dysfunction in ischemia could be further aggravated by GC-C deregulation in patients with FGDS.

Irritable bowel syndrome

IBS affects almost 10% of the general population and is usually categorized according to the predominant stool quality, diarrhoea (IBS-D) or constipation (IBS-C). FGDS shows striking similarities with IBS-D regarding abdominal pain, stool consistency, meteorism, and influence of diet upon symptoms.[150, 251] On the other hand, pharmacological GC-C activation is employed to treat IBS-C, showing that IBS like symptoms can arise at both extremes of GC-C activation.[6] Supporting that GC-C signalling contribute mechanistically to IBS, its ligands were found up regulated in a study of mucosal gene expression in IBS-D.[252]

Intestinal dysmotility is thought to be one important factor in IBS, but its exact role and underlying mechanisms are not known.[251] That loose stools correlate with fast intestinal transit has been a prevailing belief.[210] Extensive studies of intestinal motility in FGDS show that this clearly cannot be regarded as a universal principle.[174, 175] While affected by diarrhoea, FGDS patients have prominent delay of colonic transit. Also as reported in IBS-C, and unlike healthy controls, FGDS patients displayed no rise in postprandial blood serotonin level.[175, 253] Moreover the FGDS patients have some stool characteristics that are usually associated with constipation rather than diarrhoea, including a normal alpha-diversity and enrichment of Methanobacteria.[210, 254]

Methanobacteria are the sole producers of methane in the gut, and the presence of this gas has been used as diagnostic marker for IBS-C.[253, 255] There are data indicating that methanogenic bacteria are favoured by slow intestinal transit, but also that methane can slow intestinal propulsions by changing smooth muscle contractility.[254, 255] We have not assessed whether intestinal methane production induces dysmotility in FGDS. However our findings suggest that changes in intestinal motility and the microbiota induced by GC-C signalling may contribute to IBS like symptoms.

5.5 How should we manage FGDS patients in the clinic?

Conceding that the observations made in the family studied in this project may only partially be representative for patients with other *GUCY2C* variants and still different genetic and environmental background, I will present some suggestions for the clinical management of FGDS based on current pathophysiological insight.

5.5.1 Causal treatment

Any treatment given to FGDS patients should be supported by a perceived benefit weighed against potential side effects. A novel class of GC-C inhibiting drug is under development, but given the generally long process of drug development pipelines, these will probably not be commercially available for some time.[256] An available candidate, N-acetylcystein, has been shown to supress GC-C signalling in vivo and in vitro, and alleviate intestinal inflammation in an animal model. It is well tolerated during long-term treatment in patients with chronic airways disease.[257, 258, 259, 260] As such, off-label prescription of N-acetylcystein in patients with FGDS may be considered, awaiting controlled clinical studies.

5.5.2 Diarrhoea

FGDS patients are particularly vulnerable to severe diarrhoea during the neonatal period, during intestinal infections and following acute intestinal obstructions. Since weight loss may be the only sign of dehydration in the neonate,[261] weight and symptoms of new born offspring of patients with FGDS should be monitored the first weeks of life in order to avoid fluid and electrolyte derangements.[262]

Rotavirus infections have been associated with severe complications in FGDS, and given the safety profile of the current rotavirus vaccine, we recommend its use in infants with FGDS.[263]

We have not confirmed any cases of enteritis caused by ETEC infections in FGDS, which theoretically could cause severe symptoms in the context of the gain-of-function *GUCY2C* variant. An effective ETEC vaccine is not yet available, [10] and in

order to prevent any infectious diarrhoea, FGDS patients are recommended to pay attention to hygiene and avoid unsafe foods, especially during international travel.

With approval of the regional Ethical Committee, we performed an open treatment trial in FGDS with the drug metformin. This drug activates AMPK and inhibits PKA and could thus potentially diminish fluid efflux through CFTR (**Figure 7**). A significant, but modest decrease in diarrhoea was reported in the patients. The anti-diarrheal effect of this drug was later corroborated by an independent group.[264] However, one patient developed small bowel obstruction during treatment. As this is a complication experienced by several relatives with FGDS not taking the drug, and not a known side effect of the drug, any causal link is unclear. Still, the event discouraged any further use of metformin in our patients.

The intestinal dysmotility in FGDS is a known predisposing factor for diarrhoea secondary to small bowel bacterial overgrowth.[174, 175] This phenomenon, often diagnosed with the support of a lactulose-hydrogen breath tests, might be considered when intestinal gas and fluid losses are extensive. While not a constant finding, several FGDS patients reported relief of diarrhoea after courses of metronidazole or ciprofloxacin, antibiotics typically used to suppress small bowel bacterial overgrowth. One particular case warrants mention: An elderly FGDS patient underwent ileocoecal resection due to coecal volvus, and following surgery she continued to pass large amounts of diarrhoea, even after the ileostomy was converted to an ileocolonic anastomosis. She showed renal insufficiency related to dehydration and was treated regularly with intravenous rehydration. Shortly after a course of metronidazole fluid losses subsided and renal function normalised.

The observations related to antibiotics are of interest in connection with our own findings of a perturbed intestinal microbiota in FGDS. One common approach in altering the intestinal microbiota has been probiotics. The probiotic genus *Bifidobacterium,* which was almost absent in FGDS may relieve diarrhoea by a stimulating effect on ion channels.[265] Whether *Bifidobacterium* provided in probiotic formula can have such effect in FGDS could be evaluated. Longer term, the

possibility to specifically correct disturbances in the intestinal microbiota, may also hold relevance for FGDS patients.[266]

5.5.3 Bowel obstruction

When a patient with FGDS ceases to pass stools and experience abdominal pain, bowel obstruction should be suspected. While true intestinal obstructions needing surgery have occurred in ten FGDS patients, several have also been hospitalised due to pending bowel obstruction that resolved spontaneously. Aberrations of intestinal motility likely predispose for intestinal twisting as well as intussusception, which both may lead to the frank or intermittent bowel obstruction observed. Small bowel intussusceptions have in fact been observed transiently during routine ultrasonography in children and one adult with FGDS (unpublished observations), but these were not accompanied by pain or other symptoms of obstruction.

Surgery pre-disposes for adhesive bands, which again may lead to bowel obstructions. The decision to perform explorative abdominal surgery should thus be carefully contemplated, probably using standard guidelines. To our knowledge all but one case of bowel obstruction in need of intestinal resection have been treated with a primary anastomosis. In the one case where ileostomy was preferred, a subsequent early conversion to ileocolonic anastomosis was found necessary to decrease fluid losses. This observation is of relevance to choice of anastomosis in FGDS patients undergoing intestinal surgery. Our data suggest that FGDS patients may be particularly susceptible to I/R injury, with local as well as systemic inflammation (with multi-organ failure) following bowel obstruction.

We recommend that FGDS patients with suspected bowel obstruction should be admitted without delay to a hospital that can provide state-of-the art surgical and perioperative management.[267] Particular attention to fluid replacement regimens is crucial.

5.5.4 Inflammatory Bowel Disease

Patients with FGDS should avoid tobacco, a known risk factor for development of CD. We have insufficient evidence to claim that IBD-like microbiota changes (dysbiosis) in FGDS play a role in their risk for development of CD. While some patients have reported a benefit of antibiotics against diarrhoea (see above), it would still be prudent to use antibiotics judicially, since these may exacerbate dysbiosis and enhance the risk of IBD.[268, 269] Our data do not justify deviations from standard guidelines regarding treatment of IBD in patient with FGDS. Given the high proportion of FGDS that experience long-term remission from CD, the need for continuing drugs (e.g. TNFα blockers) should regularly be assessed.

6. Conclusions and future perspectives

6.1 Conclusions

The current project has outlined the clinical characteristics of a new monogenic disorder and its causal factor in the form of an activating genetic variant in *GUCY2C,* and thereby provided a new source of knowledge about this gene and a target for genetic testing in patients with similar symptoms. Our investigations of FGDS patients have highlighted changes in small bowel gene expression and the gut luminal microbiota supporting that *GUCY2C* contributes to the interplay between the human host and its resident microbiota.[212, 213] The relevance of such interaction for development of IBD in a subset of FGDS patients is supported by segregation in this group of genetic risk variants in *NOD2*, a receptor for microbial patterns. Lastly, pathway based analysis supports that changes in the *GUCY2C* pathway may be risk factors for development of IBD in the general population.

6.2 Future perspectives

6.2.1 Treatment

Development of treatments for GC-C induced diarrhoea might not only benefit individuals with the rare disorders FGDS and CSD. A drug that partially blocks GC-C activity could also help to reduce morbidity due to infections with ETEC, and thus potentially save thousands of lives.[1, 256] Based on its GC-C suppressing effect in mouse intestine and few known side effects in humans, N-acetylcystein should be tested in controlled clinical trials in FGDS.

Our studies support the potential of GC-C targeting drugs in IBD as well as in IBS. Advancement and preclinical evaluations of such drugs may be facilitated by novel methods of gene editing and model organisms, such as intestinal organoids or knock-in mouse.[270, 271]

6.2.2 Microbiota

Further studies that explore the relationship between diet, microbiota and GC-C signalling are awaited and more studies are needed to clarify whether the influence of *GUCY2C* on the intestinal microbiota is relevant for clinical outcomes. Future studies should evaluate the potential to identify subsets of GC-C related disorders (e.g. IBD and IBS) based on microbiota profile, and to tailor treatment accordingly.

The lessons from FGDS and the animal models, ask for prospective studies that account for multiple levels of GC-C signalling. The availability of GC-C activating drugs provides an excellent opportunity and a strong rationale to do so.[197]

6.2.3 GC-C and IBD

The existing paradigm concerning IBD genetic risk has mainly considered effects elicited in the intestinal wall, e.g. on epithelial barrier, innate and adaptive immune responses towards the microbiota.[21] *GUCY2C* and other genes involved in ion transport do impact on epithelial barrier, but our studies supports the concept that these inheritable factors can contribute to development and sustainment of IBD also by changing the luminal compartment and thereby the microbiota itself.[235, 236, 237] Future studies should evaluate mechanistically this hypothesis and whether applicable to *GUCY2C* or any of the other IBD risk variants. Such studies should not ignore effects on the mucus layer, which may be a crucial factor in the interaction between GC-C signalling and the microbiota.[212]

Based on the disease trajectories in FGDS patients we have proposed that episodes of hypoxia due to twisting of the gut could be one initiating factor for development of IBD. This hypothesis can be further evaluated by providing careful follow-up of our patients and obtaining good symptom descriptions.

6.2.4 GC-C and IBS

IBS is a heterogeneous disorder in terms of stool pattern, severity, response to available treatments and probably also in terms of the underlying mechanisms. Based

on our findings in FGDS we believe that assessment of the microbiota and diet should be combined with genetics as well as new modalities for investigation of gut motility in the search to uncover the pathogeneses of IBS. Such studies may clarify whether IBS comprises aetiologically distinct entities beyond the main subtypes IBS-C and IBS-D, e.g. regarding the constitution of the intestinal microbiota, regarding small bowel motility and intestinal transit, and finally define the role of GC-C.

6.2.5 Genetic testing and stratified medicine

Treatment of the monogenic disorder CF is now stratified according to genetic variants within the *CFTR* gene, providing carriers of the so-called gating mutations a specific treatment that restores CFTR function, improves life quality, lung function, nutrition and will probably give the patients a longer life.[24] This is just one example of how genetics have helped to identify subtypes of known disorders and to develop specific treatments. We have in the present work described a rare variant in the GC-C pathway linked to the common disorders IBS and IBD, suggesting that genetic testing could help to identify relevant subtypes of these disorders that could be amenable for stratified treatment.

Fortunately stratified treatment needs not be expensive. Colleagues at our own institution have done seminal work in monogenic diabetes showing that based on genetics some patients are better of with one tablet a day with an inexpensive drug, rather than the standard treatment, many injections per day of insulin.[272]

6.2.6 Safeguarding the FGDS patients

FGDS patients may continue to provide precious insights in effects of *GUCY2C* but potential gains should be weighed against the inconvenience of repeated investigations. The FGDS cohort should now be followed prospectively, and all patients should be able to easily get in contact with physicians with expertise on their disorder.[273] This could help to optimize patient management and identify factors associated with phenotypic changes (e.g. development of IBD).

7. References

1 Kotloff KL, Nataro JP, Blackwelder WC*, et al.* Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. Lancet 2013;**382**:209-22.

2 Basu N, Arshad N, Visweswariah SS. Receptor guanylyl cyclase C (GC-C): regulation and signal transduction. Molecular and cellular biochemistry 2010;**334**:67-80.

3 Antonarakis SE, Beckmann JS. Mendelian disorders deserve more attention. Nature reviews Genetics 2006;**7**:277-82.

4 Lin JE, Snook AE, Li P*, et al.* GUCY2C opposes systemic genotoxic tumorigenesis by regulating AKT-dependent intestinal barrier integrity. PLoS One 2012;**7**:e31686.

5 Steinbrecher KA, Harmel-Laws E, Garin-Laflam MP*, et al.* Murine guanylate cyclase C regulates colonic injury and inflammation. J Immunol 2011;**186**:7205-14.

6 Andresen V, Camilleri M, Busciglio IA*, et al.* Effect of 5 days linaclotide on transit and bowel function in females with constipation-predominant irritable bowel syndrome. Gastroenterology 2007;**133**:761-8.

7 Lin JE, Li P, Snook AE*, et al.* The hormone receptor GUCY2C suppresses intestinal tumor formation by inhibiting AKT signaling. Gastroenterology 2010;**138**:241-54.

8 Pitari GM. Pharmacology and clinical potential of guanylyl cyclase C agonists in the treatment of ulcerative colitis. Drug design, development and therapy 2013;**7**:351-60.

9 Liu L, Johnson HL, Cousens S*, et al.* Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. Lancet 2012;**379**:2151-61.

10 Taxt AM, Diaz Y, Aasland R*, et al.* Towards Rational Design of a Toxoid Vaccine against the Heat-Stable Toxin of Escherichia coli. Infection and immunity 2016;**84**:1239-49.

11 Viswanathan VK, Hodges K, Hecht G. Enteric infection meets intestinal function: how bacterial pathogens cause diarrhoea. Nat Rev Microbiol 2009;**7**:110-9.

12 Walker CLF, Rudan I, Liu L*, et al.* Global burden of childhood pneumonia and diarrhoea. Lancet 2013;**381**:1405-16.

13 White LE, Merrick VM, Bannerman E*, et al.* The rising incidence of celiac disease in Scotland. Pediatrics 2013;**132**:e924-31.

14 Molodecky NA, Soon IS, Rabi DM*, et al.* Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. Gastroenterology 2012;**142**:46-54 e42; quiz e30.

15 Verdu EF, Galipeau HJ, Jabri B. Novel players in coeliac disease pathogenesis: role of the gut microbiota. Nat Rev Gastroenterol Hepatol 2015;**12**:497-506.

16 Sartor RB. Microbial influences in inflammatory bowel diseases. Gastroenterology 2008;**134**:577-94.

17 Liu JZ, van Sommeren S, Huang H*, et al.* Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nat Genet 2015;**47**:979-86.

18 Haberman Y, Tickle TL, Dexheimer PJ*, et al.* Pediatric Crohn disease patients exhibit specific ileal transcriptome and microbiome signature. J Clin Invest 2014;**124**:3617-33.

19 Granlund A, Flatberg A, Ostvik AE*, et al.* Whole genome gene expression meta-analysis of inflammatory bowel disease colon mucosa demonstrates lack of major differences between Crohn's disease and ulcerative colitis. PLoS One 2013;**8**:e56818.

20 Priyamvada S, Gomes R, Gill RK*, et al.* Mechanisms Underlying Dysregulation of Electrolyte Absorption in Inflammatory Bowel Disease-Associated Diarrhea. Inflamm Bowel Dis 2015;**21**:2926-35.
21 Khor B, Gardet A, Xavier RJ. Genetics and pathogenesis of inflammatory bowel disease. Nature 2011;**474**:307-17.
22 MacKenzie T, Gifford AH, Sabadosa KA*, et al.* Longevity of patients with cystic fibrosis in 2000 to 2010 and beyond: survival analysis of the Cystic Fibrosis Foundation patient registry. Annals of internal medicine 2014;**161**:233-41.
23 Riordan JR, Rommens JM, Kerem B*, et al.* Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 1989;**245**:1066-73.
24 Harrison MJ, Murphy DM, Plant BJ. Ivacaftor in a G551D homozygote with cystic fibrosis. N Engl J Med 2013;**369**:1280-2.
25 Wainwright CE, Elborn JS, Ramsey BW*, et al.* Lumacaftor-Ivacaftor in Patients with Cystic Fibrosis Homozygous for Phe508del CFTR. N Engl J Med 2015;**373**:220-31.
26 Grasemann H. CFTR Modulator Therapy for Cystic Fibrosis. N Engl J Med 2017;**377**:2085-8.
27 Vernooij-van Langen AM, Loeber JG, Elvers B*, et al.* Novel strategies in newborn screening for cystic fibrosis: a prospective controlled study. Thorax 2012;**67**:289-95.
28 Berni Canani R, Terrin G, Cardillo G*, et al.* Congenital diarrheal disorders: improved understanding of gene defects is leading to advances in intestinal physiology and clinical management. Journal of pediatric gastroenterology and nutrition 2010;**50**:360-6.
29 Terrin G, Tomaiuolo R, Passariello A*, et al.* Congenital diarrheal disorders: an updated diagnostic approach. Int J Mol Sci 2012;**13**:4168-85.
30 Canani RB, Castaldo G, Bacchetta R*, et al.* Congenital diarrhoeal disorders: advances in this evolving web of inherited enteropathies. Nat Rev Gastroenterol Hepatol 2015;**12**:293-302.
31 Haas JT, Winter HS, Lim E*, et al.* DGAT1 mutation is linked to a congenital diarrheal disorder. J Clin Invest 2012;**122**:4680-4.
32 Hihnala S, Hoglund P, Lammi L*, et al.* Long-term clinical outcome in patients with congenital chloride diarrhea. Journal of pediatric gastroenterology and nutrition 2006;**42**:369-75.
33 Yourshaw M, Solorzano-Vargas RS, Pickett LA*, et al.* Exome sequencing finds a novel PCSK1 mutation in a child with generalized malabsorptive diarrhea and diabetes insipidus. Journal of pediatric gastroenterology and nutrition 2013;**57**:759-67.
34 Boycott KM, Vanstone MR, Bulman DE*, et al.* Rare-disease genetics in the era of next-generation sequencing: discovery to translation. Nature reviews Genetics 2013;**14**:681-91.
35 International Human Genome Sequencing C. Finishing the euchromatic sequence of the human genome. Nature 2004;**431**:931-45.
36 Genomes Project C, Abecasis GR, Auton A*, et al.* An integrated map of genetic variation from 1,092 human genomes. Nature 2012;**491**:56-65.
37 Li W, Soave D, Miller MR*, et al.* Unraveling the complex genetic model for cystic fibrosis: pleiotropic effects of modifier genes on early cystic fibrosis-related morbidities. Human genetics 2014;**133**:151-61.
38 Bartlett JR, Friedman KJ, Ling SC*, et al.* Genetic modifiers of liver disease in cystic fibrosis. JAMA 2009;**302**:1076-83.
39 Genin E, Feingold J, Clerget-Darpoux F. Identifying modifier genes of monogenic disease: strategies and difficulties. Human genetics 2008;**124**:357-68.
40 Glocker EO, Kotlarz D, Boztug K*, et al.* Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. N Engl J Med 2009;**361**:2033-45.

41 Franke A, Balschun T, Karlsen TH*, et al.* Sequence variants in IL10, ARPC2 and multiple other loci contribute to ulcerative colitis susceptibility. Nat Genet 2008;**40**:1319-23.
42 Loftus EV, Jr. Clinical epidemiology of inflammatory bowel disease: Incidence, prevalence, and environmental influences. Gastroenterology 2004;**126**:1504-17.
43 Panes J, O'Connor M, Peyrin-Biroulet L*, et al.* Improving quality of care in inflammatory bowel disease: what changes can be made today? Journal of Crohn's & colitis 2014;**8**:919-26.
44 Hovde O, Hoivik ML, Henriksen M*, et al.* Malignancies in Patients with Inflammatory Bowel Disease: Results from 20 Years of Follow-up in the IBSEN Study. Journal of Crohn's & colitis 2017;**11**:571-7.
45 Lunder AK, Hov JR, Borthne A*, et al.* Prevalence of Sclerosing Cholangitis Detected by Magnetic Resonance Cholangiography in Patients With Long-term Inflammatory Bowel Disease. Gastroenterology 2016;**151**:660-9 e4.
46 Ossum AM, Palm O, Lunder AK*, et al.* Ankylosing Spondylitis and Axial Spondyloarthritis in Patients With Long-term Inflammatory Bowel Disease: Results From 20 Years of Follow-up in the IBSEN Study. Journal of Crohn's & colitis 2018;**12**:96-104.
47 Bernstein CN, Loftus EV, Jr., Ng SC*, et al.* Hospitalisations and surgery in Crohn's disease. Gut 2012;**61**:622-9.
48 Chan HC, Ng SC. Emerging biologics in inflammatory bowel disease. J Gastroenterol 2017;**52**:141-50.
49 Borrelli O, Cordischi L, Cirulli M*, et al.* Polymeric diet alone versus corticosteroids in the treatment of active pediatric Crohn's disease: a randomized controlled open-label trial. Clin Gastroenterol Hepatol 2006;**4**:744-53.
50 Andersen V, Chan S, Luben R*, et al.* Fibre intake and the development of inflammatory bowel disease: A European prospective multi-centre cohort study (EPIC-IBD). Journal of Crohn's & colitis 2018;**12**:129-36.
51 Lee D, Albenberg L, Compher C*, et al.* Diet in the pathogenesis and treatment of inflammatory bowel diseases. Gastroenterology 2015;**148**:1087-106.
52 Tysk C, Lindberg E, Jarnerot G*, et al.* 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 1988;**29**:990-6.
53 Hugot JP, Laurent-Puig P, Gower-Rousseau C*, et al.* Mapping of a susceptibility locus for Crohn's disease on chromosome 16. Nature 1996;**379**:821-3.
54 Hugot JP, Chamaillard M, Zouali H*, et al.* Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature 2001;**411**:599-603.
55 Ogura Y, Bonen DK, Inohara N*, et al.* A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature 2001;**411**:603-6.
56 Anderson CA, Boucher G, Lees CW*, et al.* Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. Nat Genet 2011;**43**:246-52.
57 Franke A, McGovern DP, Barrett JC*, et al.* Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet 2010;**42**:1118-25.
58 Bonen DK, Ogura Y, Nicolae DL*, et al.* Crohn's disease-associated NOD2 variants share a signaling defect in response to lipopolysaccharide and peptidoglycan. Gastroenterology 2003;**124**:140-6.
59 Satsangi J, Welsh KI, Bunce M*, et al.* Contribution of genes of the major histocompatibility complex to susceptibility and disease phenotype in inflammatory bowel disease. Lancet 1996;**347**:1212-7.
60 Adams DH, Eksteen B. Aberrant homing of mucosal T cells and extra-intestinal manifestations of inflammatory bowel disease. Nature reviews Immunology 2006;**6**:244-51.

61 Hampe J, Franke A, Rosenstiel P*, et al.* A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. Nat Genet 2007;**39**:207-11.
62 Rioux JD, Xavier RJ, Taylor KD*, et al.* Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. Nat Genet 2007;**39**:596-604.
63 Fritz T, Niederreiter L, Adolph T*, et al.* Crohn's disease: NOD2, autophagy and ER stress converge. Gut 2011;**60**:1580-8.
64 Lapaquette P, Glasser AL, Huett A*, et al.* Crohn's disease-associated adherent-invasive E. coli are selectively favoured by impaired autophagy to replicate intracellularly. Cell Microbiol 2010;**12**:99-113.
65 Thorn M, Rorsman F, Ronnblom A*, et al.* Active cytomegalovirus infection diagnosed by real-time PCR in patients with inflammatory bowel disease: a prospective, controlled observational study (.). Scand J Gastroenterol 2016;**51**:1075-80.
66 Cleynen I, Boucher G, Jostins L*, et al.* Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. Lancet 2016;**387**:156-67.
67 Ellinghaus D, Jostins L, Spain SL*, et al.* Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. Nat Genet 2016;**48**:510-8.
68 Jostins L, Ripke S, Weersma RK*, et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature 2012;**491**:119-24.
69 Wedenoja S, Pekansaari E, Hoglund P*, et al.* Update on SLC26A3 mutations in congenital chloride diarrhea. Hum Mutat 2011;**32**:715-22.
70 Asano K, Matsushita T, Umeno J*, et al.* A genome-wide association study identifies three new susceptibility loci for ulcerative colitis in the Japanese population. Nat Genet 2009;**41**:1325-9.
71 Keely S, Kelly CJ, Weissmueller T*, et al.* Activated fluid transport regulates bacterial-epithelial interactions and significantly shifts the murine colonic microbiome. Gut Microbes 2012;**3**:250-60.
72 Juul FE, Garborg K, Bretthauer M*, et al.* Fecal Microbiota Transplantation for Primary Clostridium difficile Infection. N Engl J Med 2018;**378**:2535-6.
73 Marchesi JR, Adams DH, Fava F*, et al.* The gut microbiota and host health: a new clinical frontier. Gut 2016;**65**:330-9.
74 Blattner FR, Plunkett G, 3rd, Bloch CA*, et al.* The complete genome sequence of Escherichia coli K-12. Science 1997;**277**:1453-62.
75 Human Microbiome Project C. Structure, function and diversity of the healthy human microbiome. Nature 2012;**486**:207-14.
76 Xu J, Bjursell MK, Himrod J*, et al.* A genomic view of the human-Bacteroides thetaiotaomicron symbiosis. Science 2003;**299**:2074-6.
77 McDonald D, Price MN, Goodrich J*, et al.* An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. The ISME journal 2012;**6**:610-8.
78 Matamoros S, Gras-Leguen C, Le Vacon F*, et al.* Development of intestinal microbiota in infants and its impact on health. Trends Microbiol 2013;**21**:167-73.
79 David LA, Maurice CF, Carmody RN*, et al.* Diet rapidly and reproducibly alters the human gut microbiome. Nature 2014;**505**:559-63.
80 Vatanen T, Kostic AD, d'Hennezel E*, et al.* Variation in Microbiome LPS Immunogenicity Contributes to Autoimmunity in Humans. Cell 2016;**165**:842-53.
81 Penders J, Stobberingh EE, Brandt PAvd*, et al.* The role of the intestinal microbiota in the development of atopic disorders. Allergy 2007;**62**:1223-36.

82 Fanning S, Hall LJ, Cronin M*, et al.* Bifidobacterial surface-exopolysaccharide facilitates commensal-host interaction through immune modulation and pathogen protection. Proc Natl Acad Sci U S A 2012;**109**:2108-13.
83 Jorgensen SF, Troseid M, Kummen M*, et al.* Altered gut microbiota profile in common variable immunodeficiency associates with levels of lipopolysaccharide and markers of systemic immune activation. Mucosal immunology 2016;**9**:1455-65.
84 Kummen M, Holm K, Anmarkrud JA*, et al.* The gut microbial profile in patients with primary sclerosing cholangitis is distinct from patients with ulcerative colitis without biliary disease and healthy controls. Gut 2017;**66**:611-9.
85 Le Chatelier E, Nielsen T, Qin J*, et al.* Richness of human gut microbiome correlates with metabolic markers. Nature 2013;**500**:541-6.
86 De Filippo C, Cavalieri D, Di Paola M*, et al.* Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proc Natl Acad Sci U S A 2010;**107**:14691-6.
87 Yatsunenko T, Rey FE, Manary MJ*, et al.* Human gut microbiome viewed across age and geography. Nature 2012;**486**:222-7.
88 Vrieze A, Van Nood E, Holleman F*, et al.* Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. Gastroenterology 2012;**143**:913-6 e7.
89 Shimazu T, Hirschey MD, Newman J*, et al.* Suppression of oxidative stress by beta-hydroxybutyrate, an endogenous histone deacetylase inhibitor. Science 2013;**339**:211-4.
90 Kiela PR, Kuscuoglu N, Midura AJ*, et al.* Molecular mechanism of rat NHE3 gene promoter regulation by sodium butyrate. American journal of physiology Cell physiology 2007;**293**:C64-74.
91 Rajca S, Grondin V, Louis E*, et al.* Alterations in the intestinal microbiome (dysbiosis) as a predictor of relapse after infliximab withdrawal in Crohn's disease. Inflamm Bowel Dis 2014;**20**:978-86.
92 Quevrain E, Maubert MA, Michon C*, et al.* Identification of an anti-inflammatory protein from Faecalibacterium prausnitzii, a commensal bacterium deficient in Crohn's disease. Gut 2016;**65**:415-25.
93 Rossi O, van Berkel LA, Chain F*, et al.* Faecalibacterium prausnitzii A2-165 has a high capacity to induce IL-10 in human and murine dendritic cells and modulates T cell responses. Scientific reports 2016;**6**:18507.
94 Gerasimidis K, Bertz M, Hanske L*, et al.* Decline in presumptively protective gut bacterial species and metabolites are paradoxically associated with disease improvement in pediatric Crohn's disease during enteral nutrition. Inflamm Bowel Dis 2014;**20**:861-71.
95 Sartor RB. Key questions to guide a better understanding of host-commensal microbiota interactions in intestinal inflammation. Mucosal immunology 2011;**4**:127-32.
96 Frank DN, Robertson CE, Hamm CM*, et al.* Disease phenotype and genotype are associated with shifts in intestinal-associated microbiota in inflammatory bowel diseases. Inflamm Bowel Dis 2011;**17**:179-84.
97 Ott J, Wang J, Leal SM. Genetic linkage analysis in the age of whole-genome sequencing. Nature reviews Genetics 2015;**16**:275-84.
98 Duncan E, Brown M, Shore EM. The revolution in human monogenic disease mapping. Genes (Basel) 2014;**5**:792-803.
99 Cantor RM, Lange K, Sinsheimer JS. Prioritizing GWAS results: A review of statistical methods and recommendations for their application. Am J Hum Genet 2010;**86**:6-22.
100 Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide association studies. Nature reviews Genetics 2010;**11**:843-54.

101 Network, Pathway Analysis Subgroup of Psychiatric Genomics C. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nature neuroscience 2015;**18**:199-209.
102 de Leeuw CA, Neale BM, Heskes T*, et al.* The statistical properties of gene-set analysis. Nature reviews Genetics 2016;**17**:353-64.
103 Tyler AD, Smith MI, Silverberg MS. Analyzing the human microbiome: a "how to" guide for physicians. Am J Gastroenterol 2014;**109**:983-93.
104 Kevans D, Tyler AD, Holm K*, et al.* Characterization of Intestinal Microbiota in Ulcerative Colitis Patients with and without Primary Sclerosing Cholangitis. Journal of Crohn's & colitis 2016;**10**:330-7.
105 Silverberg MS, Satsangi J, Ahmad T*, et al.* Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. Can J Gastroenterol 2005;**19 Suppl A**:5A-36A.
106 Slattery ML, Pellatt DF, Mullany LE*, et al.* Differential Gene Expression in Colon Tissue Associated With Diet, Lifestyle, and Related Oxidative Stress. PLoS One 2015;**10**:e0134406.
107 Wang J, Thingholm LB, Skieceviciene J*, et al.* Genome-wide association analysis identifies variation in vitamin D receptor and other host factors influencing the gut microbiota. Nat Genet 2016;**48**:1396-406.
108 Andersen LF, Solvoll K, Johansson LR*, et al.* Evaluation of a food frequency questionnaire with weighed records, fatty acids, and alpha-tocopherol in adipose tissue and serum. American journal of epidemiology 1999;**150**:75-87.
109 Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr 1997;**65**:1220S-8S; discussion 9S-31S.
110 Graf D, Di Cagno R, Fak F*, et al.* Contribution of diet to the composition of the human gut microbiota. Microbial ecology in health and disease 2015;**26**:26164.
111 Wu GD, Chen J, Hoffmann C*, et al.* Linking long-term dietary patterns with gut microbial enterotypes. Science 2011;**334**:105-8.
112 Schofield WN. Predicting basal metabolic rate, new standards and review of previous work. Hum Nutr Clin Nutr 1985;**39 Suppl 1**:5-41.
113 Svedlund J, Sjodin I, Dotevall G. GSRS--a clinical rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and peptic ulcer disease. Dig Dis Sci 1988;**33**:129-34.
114 Harvey RF, Bradshaw JM. A simple index of Crohn's-disease activity. Lancet 1980;**1**:514.
115 Walmsley RS, Ayres RC, Pounder RE*, et al.* A simple clinical colitis activity index. Gut 1998;**43**:29-32.
116 Langhorst J, Elsenbruch S, Koelzer J*, et al.* Noninvasive markers in the assessment of intestinal inflammation in inflammatory bowel diseases: performance of fecal lactoferrin, calprotectin, and PMN-elastase, CRP, and clinical indices. Am J Gastroenterol 2008;**103**:162-9.
117 Thatcher SA. DNA/RNA preparation for molecular detection. Clin Chem 2015;**61**:89-99.
118 Lockhart DJ, Winzeler EA. Genomics, gene expression and DNA arrays. Nature 2000;**405**:827-36.
119 Slonim DK, Yanai I. Getting started in gene expression microarray analysis. PLoS Comput Biol 2009;**5**:e1000543.

120 Jorgensen SF, Reims HM, Frydenlund D*, et al.* A Cross-Sectional Study of the Prevalence of Gastrointestinal Symptoms and Pathology in Patients With Common Variable Immunodeficiency. Am J Gastroenterol 2016;**111**:1467-75.
121 Rismo R, Olsen T, Ciu G*, et al.* The effect of adalimumab for induction of endoscopic healing and normalization of mucosal cytokine gene expression in Crohn's disease. Scand J Gastroenterol 2012;**47**:1200-10.
122 Salonen A, Nikkila J, Jalanka-Tuovinen J*, et al.* Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis. J Microbiol Methods 2010;**81**:127-34.
123 Ahlquist DA, Schwartz S, Isaacson J*, et al.* A stool collection device: the first step in occult blood testing. Annals of internal medicine 1988;**108**:609-12.
124 Wu GD, Lewis JD, Hoffmann C*, et al.* Sampling and pyrosequencing methods for characterizing bacterial communities in the human gut using 16S sequence tags. BMC Microbiol 2010;**10**:206.
125 Genomes Project C, Auton A, Brooks LD*, et al.* A global reference for human genetic variation. Nature 2015;**526**:68-74.
126 Trynka G, Hunt KA, Bockett NA*, et al.* Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. Nat Genet 2011;**43**:1193-201.
127 Steemers FJ, Chang W, Lee G*, et al.* Whole-genome genotyping with the single-base extension assay. Nature methods 2006;**3**:31-3.
128 Bolstad BM, Irizarry RA, Astrand M*, et al.* A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 2003;**19**:185-93.
129 Fellenberg K, Hauser NC, Brors B*, et al.* Correspondence analysis applied to microarray data. Proc Natl Acad Sci U S A 2001;**98**:10781-6.
130 Dysvik B, Jonassen I. J-Express: exploring gene expression data using Java. Bioinformatics 2001;**17**:369-70.
131 Bustin SA, Benes V, Garson JA*, et al.* The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 2009;**55**:611-22.
132 Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001;**25**:402-8.
133 Matos LL, Trufelli DC, de Matos MG*, et al.* Immunohistochemistry as an important tool in biomarkers detection and clinical practice. Biomark Insights 2010;**5**:9-20.
134 Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 1977;**74**:5463-7.
135 McKenna A, Hanna M, Banks E*, et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research 2010;**20**:1297-303.
136 Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;**25**:1754-60.
137 Li H, Handsaker B, Wysoker A*, et al.* The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;**25**:2078-9.
138 Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research 2010;**38**:e164.
139 Muzzey D, Evans EA, Lieber C. Understanding the Basics of NGS: From Mechanism to Variant Calling. Curr Genet Med Rep 2015;**3**:158-65.
140 Koboldt DC, Steinberg KM, Larson DE*, et al.* The next-generation sequencing revolution and its impact on genomics. Cell 2013;**155**:27-38.

141 Bentley DR, Balasubramanian S, Swerdlow HP*, et al.* Accurate whole human genome sequencing using reversible terminator chemistry. Nature 2008;**456**:53-9.
142 Kozich JJ, Westcott SL, Baxter NT*, et al.* Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Applied and environmental microbiology 2013;**79**:5112-20.
143 DeSantis TZ, Hugenholtz P, Larsen N*, et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Applied and environmental microbiology 2006;**72**:5069-72.
144 Caporaso JG, Kuczynski J, Stombaugh J*, et al.* QIIME allows analysis of high-throughput community sequencing data. Nature methods 2010;**7**:335-6.
145 Magoc T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 2011;**27**:2957-63.
146 Fadrosh DW, Ma B, Gajer P*, et al.* An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. Microbiome 2014;**2**:6.
147 Shenoy AR, Visweswariah SS. Site-directed mutagenesis using a single mutagenic oligonucleotide and DpnI digestion of template DNA. Anal Biochem 2003;**319**:335-6.
148 Saha S, Biswas KH, Kondapalli C*, et al.* The linker region in receptor guanylyl cyclases is a key regulatory module: mutational analysis of guanylyl cyclase C. The Journal of biological chemistry 2009;**284**:27135-45.
149 Fraser AG, Marcotte EM. A probabilistic view of gene function. Nat Genet 2004;**36**:559-64.
150 Fiskerstrand T, Arshad N, Haukanes BI*, et al.* Familial diarrhea syndrome caused by an activating GUCY2C mutation. N Engl J Med 2012;**366**:1586-95.
151 Kato A, Romero MF. Regulation of electroneutral NaCl absorption by the small intestine. Annu Rev Physiol 2011;**73**:261-81.
152 Gudbjartsson DF, Thorvaldsson T, Kong A*, et al.* Allegro version 2. Nat Genet 2005;**37**:1015-6.
153 Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 2001;**98**:5116-21.
154 Breitling R, Armengaud P, Amtmann A*, et al.* Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Lett 2004;**573**:83-92.
155 Stavrum AK, Petersen K, Jonassen I*, et al.* Analysis of gene-expression data using J-Express. Current protocols in bioinformatics / editoral board, Andreas D Baxevanis [et al] 2008;**Chapter 7**:Unit 7 3.
156 Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research 2009;**37**:1-13.
157 Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;**4**:44-57.
158 Subramanian A, Tamayo P, Mootha VK*, et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;**102**:15545-50.
159 Golub TR, Slonim DK, Tamayo P*, et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999;**286**:531-7.
160 Christoforou A, Dondrup M, Mattingsdal M*, et al.* Linkage-disequilibrium-based binning affects the interpretation of GWASs. Am J Hum Genet 2012;**90**:727-33.

161 Moskvina V, O'Dushlaine C, Purcell S*, et al.* Evaluation of an approximation method for assessment of overall significance of multiple-dependent tests in a genomewide association study. Genetic epidemiology 2011;**35**:861-6.
162 Ersland KM, Christoforou A, Stansberg C*, et al.* Gene-based analysis of regionally enriched cortical genes in GWAS data sets of cognitive traits and psychiatric disorders. PLoS One 2012;**7**:e31687.
163 Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. Applied and environmental microbiology 2005;**71**:8228-35.
164 Navas-Molina JA, Peralta-Sanchez JM, Gonzalez A*, et al.* Advancing our understanding of the human microbiome using QIIME. Methods in enzymology 2013;**531**:371-444.
165 Goodrich JK, Di Rienzi SC, Poole AC*, et al.* Conducting a microbiome study. Cell 2014;**158**:250-62.
166 Segata N, Izard J, Waldron L*, et al.* Metagenomic biomarker discovery and explanation. Genome Biol 2011;**12**:R60.
167 Morgan XC, Tickle TL, Sokol H*, et al.* Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. Genome Biol 2012;**13**:R79.
168 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society 1995;**57**:11.
169 Wei C, Li J, Bumgarner RE. Sample size for detecting differentially expressed genes in microarray experiments. BMC Genomics 2004;**5**:87.
170 Yazdanyar S, Kamstrup PR, Tybjaerg-Hansen A*, et al.* Penetrance of NOD2/CARD15 genetic variants in the general population. CMAJ 2010;**182**:661-5.
171 Ananthakrishnan AN, Huang H, Nguyen DD*, et al.* Differential effect of genetic burden on disease phenotypes in Crohn's disease and ulcerative colitis: analysis of a North American cohort. Am J Gastroenterol 2014;**109**:395-400.
172 Lawrance IC, Fiocchi C, Chakravarti S. Ulcerative colitis and Crohn's disease: distinctive gene expression profiles and novel susceptibility candidate genes. Hum Mol Genet 2001;**10**:445-56.
173 Waeytens A, De Vos M, Laukens D. Evidence for a potential role of metallothioneins in inflammatory bowel diseases. Mediators Inflamm 2009;**2009**:729172.
174 von Volkmann HL, Nylund K, Tronstad RR*, et al.* An activating gucy2c mutation causes impaired contractility and fluid stagnation in the small bowel. Scand J Gastroenterol 2016;**51**:1308-15.
175 von Volkmann HL, Bronstad I, Gilja OH*, et al.* Prolonged intestinal transit and diarrhea in patients with an activating GUCY2C mutation. PLoS One 2017;**12**:e0185496.
176 Holmberg C, Perheentupa J. Congenital Na+ diarrhea: a new type of secretory diarrhea. J Pediatr 1985;**106**:56-61.
177 Muller T, Wijmenga C, Phillips AD*, et al.* Congenital sodium diarrhea is an autosomal recessive disorder of sodium/proton exchange but unrelated to known candidate genes. Gastroenterology 2000;**119**:1506-13.
178 Muller T, Rasool I, Heinz-Erian P*, et al.* Congenital secretory diarrhoea caused by activating germline mutations in GUCY2C. Gut 2016;**65**:1306-13.
179 Janecke AR, Heinz-Erian P, Yin J*, et al.* Reduced sodium/proton exchanger NHE3 activity causes congenital sodium diarrhea. Hum Mol Genet 2015;**24**:6614-23.
180 Engevik AC, Kaji I, Engevik MA*, et al.* Loss of MYO5B Leads to Reductions in Na(+) Absorption With Maintenance of CFTR-Dependent Cl(-) Secretion in Enterocytes. Gastroenterology 2018;**155**:1883-97 e10.
181 Miele E, Simeone D, Marino A*, et al.* Functional gastrointestinal disorders in children: an Italian prospective survey. Pediatrics 2004;**114**:73-8.

182 Schulz S, Green CK, Yuen PS*, et al.* Guanylyl cyclase is a heat-stable enterotoxin receptor. Cell 1990;**63**:941-8.
183 Currie MG, Fok KF, Kato J*, et al.* Guanylin: an endogenous activator of intestinal guanylate cyclase. Proc Natl Acad Sci U S A 1992;**89**:947-51.
184 Hamra FK, Forte LR, Eber SL*, et al.* Uroguanylin: structure and activity of a second endogenous peptide that stimulates intestinal guanylate cyclase. Proc Natl Acad Sci U S A 1993;**90**:10464-8.
185 Krause WJ, Cullingford GL, Freeman RH*, et al.* Distribution of heat-stable enterotoxin/guanylin receptors in the intestinal tract of man and other mammals. J Anat 1994;**184 (Pt 2)**:407-17.
186 Cohen MB, Guarino A, Shukla R*, et al.* Age-related differences in receptors for Escherichia coli heat-stable enterotoxin in the small and large intestine of children. Gastroenterology 1988;**94**:367-73.
187 Hughes JM, Murad F, Chang B*, et al.* Role of cyclic GMP in the action of heat-stable enterotoxin of Escherichia coli. Nature 1978;**271**:755-6.
188 Roy N, Guruprasad MR, Kondaiah P*, et al.* Protein kinase C regulates transcription of the human guanylate cyclase C gene. Eur J Biochem 2001;**268**:2160-71.
189 Steinbrecher KA, Mann EA, Giannella RA*, et al.* Increases in guanylin and uroguanylin in a mouse model of osmotic diarrhea are guanylate cyclase C-independent. Gastroenterology 2001;**121**:1191-202.
190 Lucas KA, Pitari GM, Kazerounian S*, et al.* Guanylyl cyclases and signaling by cyclic GMP. Pharmacol Rev 2000;**52**:375-414.
191 Valentino MA, Lin JE, Snook AE*, et al.* A uroguanylin-GUCY2C endocrine axis regulates feeding in mice. J Clin Invest 2011;**121**:3578-88.
192 Brierley SM. Guanylate cyclase-C receptor activation: unexpected biology. Curr Opin Pharmacol 2012;**12**:632-40.
193 Grundy D, Al-Chaer ED, Aziz Q*, et al.* Fundamentals of neurogastroenterology: basic science. Gastroenterology 2006;**130**:1391-411.
194 Brenna O, Furnes MW, Munkvold B*, et al.* Cellular localization of guanylin and uroguanylin mRNAs in human and rat duodenal and colonic mucosa. Cell Tissue Res 2016;**365**:331-41.
195 Kuhn M. Molecular Physiology of Membrane Guanylyl Cyclase Receptors. Physiol Rev 2016;**96**:751-804.
196 Friebe A, Mergia E, Dangel O*, et al.* Fatal gastrointestinal obstruction and hypertension in mice lacking nitric oxide-sensitive guanylyl cyclase. Proc Natl Acad Sci U S A 2007;**104**:7699-704.
197 Castro J, Harrington AM, Hughes PA*, et al.* Linaclotide inhibits colonic nociceptors and relieves abdominal pain via guanylate cyclase-C and extracellular cyclic guanosine 3',5'-monophosphate. Gastroenterology 2013;**145**:1334-46 e1-11.
198 Kotlo KU, Rasenick MM, Danziger RS. Evidence for cross-talk between atrial natriuretic peptide and nitric oxide receptors. Molecular and cellular biochemistry 2010;**338**:183-9.
199 Folgueira C, Sanchez-Rebordelo E, Barja-Fernandez S*, et al.* Uroguanylin levels in intestine and plasma are regulated by nutritional status in a leptin-dependent manner. European journal of nutrition 2016;**55**:529-36.
200 Lin JE, Colon-Gonzalez F, Blomain E*, et al.* Obesity-Induced Colorectal Cancer Is Driven by Caloric Silencing of the Guanylin-GUCY2C Paracrine Signaling Axis. Cancer Res 2016;**76**:339-46.
201 King JC. Zinc: an essential but elusive nutrient. Am J Clin Nutr 2011;**94**:679S-84S.

202 Blanchard RK, Cousins RJ. Upregulation of rat intestinal uroguanylin mRNA by dietary zinc restriction. Am J Physiol 1997;**272**:G972-8.
203 Basu N, Saha S, Khan I*, et al.* Intestinal cell proliferation and senescence are regulated by receptor guanylyl cyclase C and p21. The Journal of biological chemistry 2014;**289**:581-93.
204 Krezel A, Hao Q, Maret W. The zinc/thiolate redox biochemistry of metallothionein and the control of zinc ion fluctuations in cell signaling. Arch Biochem Biophys 2007;**463**:188-200.
205 Neish AS, Jones RM. Redox signaling mediates symbiosis between the gut microbiota and the intestine. Gut microbes 2014;**5**:250-3.
206 Khan MT, van Dijl JM, Harmsen HJ. Antioxidants keep the potentially probiotic but highly oxygen-sensitive human gut bacterium Faecalibacterium prausnitzii alive at ambient air. PLoS One 2014;**9**:e96097.
207 Bonder MJ, Kurilshikov A, Tigchelaar EF*, et al.* The effect of host genetics on the gut microbiome. Nat Genet 2016;**48**:1407-12.
208 Turpin W, Espin-Garcia O, Xu W*, et al.* Association of host genome with intestinal microbial composition in a large healthy cohort. Nat Genet 2016;**48**:1413-7.
209 Falony G, Joossens M, Vieira-Silva S*, et al.* Population-level analysis of gut microbiome variation. Science 2016;**352**:560-4.
210 Vandeputte D, Falony G, Vieira-Silva S*, et al.* Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. Gut 2016;**65**:57-62.
211 Duncan SH, Louis P, Thomson JM*, et al.* The role of pH in determining the species composition of the human colonic microbiota. Environ Microbiol 2009;**11**:2112-22.
212 Amarachintha S, Harmel-Laws E, Steinbrecher KA. Guanylate cyclase C reduces invasion of intestinal epithelial cells by bacterial pathogens. Scientific reports 2018;**8**:1521.
213 Majumdar S, Mishra V, Nandi S*, et al.* Absence of Receptor Guanylyl Cyclase C Enhances Ileal Damage and Reduces Cytokine and Antimicrobial Peptide Production during Oral Salmonella enterica Serovar Typhimurium Infection. Infection and immunity 2018;**86**.
214 Harmel-Laws E, Mann EA, Cohen MB*, et al.* Guanylate cyclase C deficiency causes severe inflammation in a murine model of spontaneous colitis. PLoS One 2013;**8**:e79180.
215 Tal A, Carmi R, Chai-Am E*, et al.* Familial meconium ileus with normal sweat electrolytes. Clin Pediatr (Phila) 1985;**24**:460-2.
216 Romi H, Cohen I, Landau D*, et al.* Meconium ileus caused by mutations in GUCY2C, encoding the CFTR-activating guanylate cyclase 2C. Am J Hum Genet 2012;**90**:893-9.
217 Mann EA, Harmel-Laws E, Cohen MB*, et al.* Guanylate cyclase C limits systemic dissemination of a murine enteric pathogen. BMC Gastroenterol 2013;**13**:135.
218 Uhlig HH. Monogenic diseases associated with intestinal inflammation: implications for the understanding of inflammatory bowel disease. Gut 2013;**62**:1795-805.
219 Heresbach D, Gicquel-Douabin V, Birebent B*, et al.* NOD2/CARD15 gene polymorphisms in Crohn's disease: a genotype- phenotype analysis. Eur J Gastroenterol Hepatol 2004;**16**:55-62.
220 Vermeire S, Wild G, Kocher K*, et al.* CARD15 genetic variation in a Quebec population: prevalence, genotype-phenotype relationship, and haplotype structure. Am J Hum Genet 2002;**71**:74-83.
221 Hampe J, Grebe J, Nikolaus S*, et al.* Association of NOD2 (CARD 15) genotype with clinical course of Crohn's disease: a cohort study. Lancet 2002;**359**:1661-5.
222 Ioachim E, Michael M, Katsanos C*, et al.* The immunohistochemical expression of metallothionein in inflammatory bowel disease. Correlation with HLA-DR antigen

expression, lymphocyte subpopulations and proliferation-associated indices. Histology and histopathology 2003;**18**:75-82.
223 Perez-Torras S, Iglesias I, Llopis M*, et al.* Transportome Profiling Identifies Profound Alterations in Crohn's Disease Partially Restored by Commensal Bacteria. Journal of Crohn's & colitis 2016;**10**:850-9.
224 Fukushima K, Ogawa H, Takahashi K*, et al.* Non-pathogenic bacteria modulate colonic epithelial gene expression in germ-free mice. Scand J Gastroenterol 2003;**38**:626-34.
225 Sommer F, Nookaew I, Sommer N*, et al.* Site-specific programming of the host epithelial transcriptome by the gut microbiota. Genome Biol 2015;**16**:62.
226 Reikvam DH, Erofeev A, Sandvik A*, et al.* Depletion of murine intestinal microbiota: effects on gut mucosa and epithelial gene expression. PLoS One 2011;**6**:e17996.
227 Lahiri A, Abraham C. Activation of pattern recognition receptors up-regulates metallothioneins, thereby increasing intracellular accumulation of zinc, autophagy, and bacterial clearance by macrophages. Gastroenterology 2014;**147**:835-46.
228 Naftali T, Reshef L, Kovacs A*, et al.* Distinct Microbiotas are Associated with Ileum-Restricted and Colon-Involving Crohn's Disease. Inflamm Bowel Dis 2016;**22**:293-302.
229 Sokol H, Pigneur B, Watterlot L*, et al.* Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. Proc Natl Acad Sci U S A 2008;**105**:16731-6.
230 David LA, Weil A, Ryan ET*, et al.* Gut microbial succession follows acute secretory diarrhea in humans. mBio 2015;**6**:e00381-15.
231 Brenna O, Bruland T, Furnes MW*, et al.* The guanylate cyclase-C signaling pathway is down-regulated in inflammatory bowel disease. Scand J Gastroenterol 2015;**50**:1241-52.
232 Amin MR, Dudeja PK, Ramaswamy K*, et al.* Involvement of Sp1 and Sp3 in differential regulation of human NHE3 promoter activity by sodium butyrate and IFN-gamma/TNF-alpha. Am J Physiol Gastrointest Liver Physiol 2007;**293**:G374-82.
233 Larmonier CB, Laubitz D, Hill FM*, et al.* Reduced colonic microbial diversity is associated with colitis in NHE3-deficient mice. Am J Physiol Gastrointest Liver Physiol 2013;**305**:G667-77.
234 Laubitz D, Harrison CA, Midura-Kiela MT*, et al.* Reduced Epithelial Na+/H+ Exchange Drives Gut Microbial Dysbiosis and Promotes Inflammatory Response in T Cell-Mediated Murine Colitis. PLoS One 2016;**11**:e0152044.
235 Imhann F, Vich Vila A, Bonder MJ*, et al.* Interplay of host genetics and gut microbiota underlying the onset and clinical presentation of inflammatory bowel disease. Gut 2018;**67**:108-19.
236 Hedin C, van der Gast CJ, Rogers GB*, et al.* Siblings of patients with Crohn's disease exhibit a biologically relevant dysbiosis in mucosal microbial metacommunities. Gut 2016;**65**:944-53.
237 Li D, Achkar JP, Haritunians T*, et al.* A Pleiotropic Missense Variant in SLC39A8 Is Associated With Crohn's Disease and Human Gut Microbiome Composition. Gastroenterology 2016;**151**:724-32.
238 Kapadia MR. Volvulus of the Small Bowel and Colon. Clin Colon Rectal Surg 2017;**30**:40-5.
239 Large AM. Partial intermittent volvulus of the cecum. Ann Surg 1968;**167**:609-11.
240 Tirol FT. Recurrent cecocolic torsion: radiological diagnosis and treatment. JSLS 2003;**7**:23-31.
241 de Betue CT, Boersma D, Oomen MW*, et al.* Volvulus as a complication of chronic intestinal pseudo-obstruction syndrome. Eur J Pediatr 2011;**170**:1591-5.

242 Gonzalez LM, Moeser AJ, Blikslager AT. Animal models of ischemia-reperfusion-induced intestinal injury: progress and promise for translational research. Am J Physiol Gastrointest Liver Physiol 2015;**308**:G63-75.
243 Tsimoyiannis EC, Tsimoyiannis JC, Sarros CJ*, et al.* The role of oxygen-derived free radicals in peritoneal adhesion formation induced by ileal ischaemia/reperfusion. Acta Chir Scand 1989;**155**:171-4.
244 Grootjans J, Lenaerts K, Derikx JP*, et al.* Human intestinal ischemia-reperfusion-induced inflammation characterized: experiences from a new translational model. The American journal of pathology 2010;**176**:2283-91.
245 Hassoun HT, Kone BC, Mercer DW*, et al.* Post-injury multiple organ failure: the role of the gut. Shock 2001;**15**:1-10.
246 Glover LE, Colgan SP. Hypoxia and metabolic factors that influence inflammatory bowel disease pathogenesis. Gastroenterology 2011;**140**:1748-55.
247 Vavricka SR, Rogler G, Maetzler S*, et al.* High altitude journeys and flights are associated with an increased risk of flares in inflammatory bowel disease patients. Journal of Crohn's & colitis 2014;**8**:191-9.
248 Perez-Chanona E, Muhlbauer M, Jobin C. The microbiota protects against ischemia/reperfusion-induced intestinal injury through nucleotide-binding oligomerization domain-containing protein 2 (NOD2) signaling. The American journal of pathology 2014;**184**:2965-75.
249 van Lookeren Campagne M, Thibodeaux H, van Bruggen N*, et al.* Evidence for a protective role of metallothionein-1 in focal cerebral ischemia. Proc Natl Acad Sci U S A 1999;**96**:12870-5.
250 Lo HC, Yang RB, Lee CH. Guanylyl cyclase-G modulates jejunal apoptosis and inflammation in mice with intestinal ischemia and reperfusion. PLoS One 2014;**9**:e101314.
251 Rajilic-Stojanovic M, Jonkers DM, Salonen A*, et al.* Intestinal microbiota and diet in IBS: causes, consequences, or epiphenomena? Am J Gastroenterol 2015;**110**:278-87.
252 Camilleri M, Carlson P, Acosta A*, et al.* RNA sequencing shows transcriptomic changes in rectosigmoid mucosa in patients with irritable bowel syndrome-diarrhea: a pilot case-control study. Am J Physiol Gastrointest Liver Physiol 2014;**306**:G1089-98.
253 Pimentel M, Kong Y, Park S. IBS subjects with methane on lactulose breath test have lower postprandial serotonin levels than subjects with hydrogen. Dig Dis Sci 2004;**49**:84-7.
254 El Oufir L, Flourie B, Bruley des Varannes S*, et al.* Relations between transit time, fermentation products, and hydrogen consuming flora in healthy humans. Gut 1996;**38**:870-7.
255 Pimentel M, Lin HC, Enayati P*, et al.* Methane, a gas produced by enteric bacteria, slows intestinal transit and augments small intestinal contractile activity. Am J Physiol Gastrointest Liver Physiol 2006;**290**:G1089-95.
256 Bijvelds MJ, Loos M, Bronsveld I*, et al.* Inhibition of Heat-Stable Toxin-Induced Intestinal Salt and Water Secretion by a Novel Class of Guanylyl Cyclase C Inhibitors. J Infect Dis 2015;**212**:1806-15.
257 Greenberg RN, Dunn JA, Guerrant RL. Reduction of the secretory response to Escherichia coli heat-stable enterotoxin by thiol and disulfide compounds. Infection and immunity 1983;**41**:174-80.
258 Siddiqui A, Ancha H, Tedesco D*, et al.* Antioxidant therapy with N-acetylcysteine plus mesalamine accelerates mucosal healing in a rodent model of colitis. Dig Dis Sci 2006;**51**:698-705.
259 Cetinkaya A, Bulbuloglu E, Kurutas EB*, et al.* Beneficial effects of N-acetylcysteine on acetic acid-induced colitis in rats. The Tohoku journal of experimental medicine 2005;**206**:131-9.

260 Tam J, Nash EF, Ratjen F*, et al.* Nebulized and oral thiol derivatives for pulmonary disease in cystic fibrosis. Cochrane Database Syst Rev 2013;**7**:CD007168.
261 Tjora E, Karlsen LC, Moster D*, et al.* Early severe weight loss in newborns after discharge from regular nurseries. Acta Paediatr 2010;**99**:654-7.
262 Guarino A, Cohen MB, Giannella RA. Small and large intestinal guanylate cyclase activity in children: effect of age and stimulation by Escherichia coli heat-stable enterotoxin. Pediatr Res 1987;**21**:551-5.
263 Buyse H, Vinals C, Karkada N*, et al.* The human rotavirus vaccine Rotarix in infants: an integrated analysis of safety and reactogenicity. Hum Vaccin Immunother 2014;**10**:19-24.
264 Rogers AC, Huetter L, Hoekstra N*, et al.* Activation of AMPK inhibits cholera toxin stimulated chloride secretion in human and murine intestine. PLoS One 2013;**8**:e69050.
265 Kumar A, Hecht C, Priyamvada S*, et al.* Probiotic Bifidobacterium species stimulate human SLC26A3 gene function and expression in intestinal epithelial cells. American journal of physiology Cell physiology 2014;**307**:C1084-92.
266 Atarashi K, Tanoue T, Oshima K*, et al.* Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. Nature 2013;**500**:232-6.
267 Pujahari AK. Decision Making in Bowel Obstruction: A Review. Journal of Clinical and Diagnostic Research 2016;**10**:PE07-PE12.
268 Lewis JD, Chen EZ, Baldassano RN*, et al.* Inflammation, Antibiotics, and Diet as Environmental Stressors of the Gut Microbiome in Pediatric Crohn's Disease. Cell Host Microbe 2017;**22**:247.
269 Card T, Logan RF, Rodrigues LC*, et al.* Antibiotic use and the development of Crohn's disease. Gut 2004;**53**:246-50.
270 Sato T, Vries RG, Snippert HJ*, et al.* Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature 2009;**459**:262-5.
271 Neggers JE, Vercruysse T, Jacquemyn M*, et al.* Identifying drug-target selectivity of small-molecule CRM1/XPO1 inhibitors by CRISPR/Cas9 genome editing. Chem Biol 2015;**22**:107-16.
272 Pearson ER, Flechtner I, Njolstad PR*, et al.* Switching from insulin to oral sulfonylureas in patients with diabetes due to Kir6.2 mutations. N Engl J Med 2006;**355**:467-77.
273 von der Lippe C, Diesen PS, Feragen KB. Living with a rare disorder: a systematic review of the qualitative literature. Mol Genet Genomic Med 2017;**5**:758-73.

Appendices

Rome II Integrative Questionnaire
Norsk Versjon

Del A
F¿rst vil vi stille deg noen sp¿rsmŒl om smerter eller ubehag i magen i l¿pet av det siste Œret.

Med magesmerter eller ubehag, mener vi smerte eller ubehag i omrŒdene merket A, B, C, D slik det er vist i diagram 1. (Vennligst ikke ta med krampe eller smerte i forbindelse med menstruasjonen og heller ikke smerter i brystet).

A1. Har du i l¿pet av det siste Œret hatt smerter eller ubehag i magen?

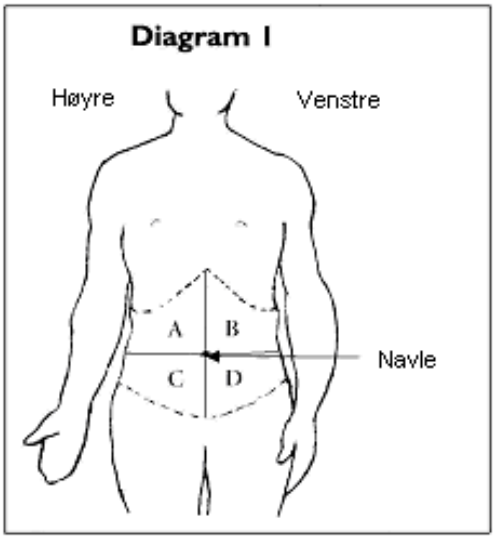
- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

A 2. Har du i l¿pet av de tre siste mŒnedene hatt ubehag eller smerte i magen?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

Dersom du svarer Sjelden eller aldri pŒ sp¿rsmŒl A1 og A2,
vennligst gŒ til sp¿rsmŒl B 1.

Diagram I



Noen ganger: mer enn en av ti uker
Ofte: mer enn en av fire uker
Svært ofte: mer enn halvparten av ukene

A 3. Hvordan er ubehaget eller smerten du har hatt de siste tre mŒnedene sammenliknet med hva du har hatt det siste Œret?

- ❑ **Bedre de siste tre mŒnedene**
- ❑ **Verre de siste tre mŒnedene**
- ❑ **Omtrent det samme**
- ❑ **Har ikke hatt ubehag eller smerte de siste tre mŒnedene**

A 4. NŒr du i l¿pet av det siste Œret hadde ubehag eller smerte i magen, var de da kontinuerlig eller nesten kontinuerlig (uten opphold)?

- ❑ **Ja**
- ❑ **Nei**

A 5. Hvor gammel var du da ubehaget eller smerten begynte? (sŒ n¾rt som du kan huske)

Alder_____

A 6. Hvor er magen din lokalisert i forhold til diagrammet pŒ forutgŒende side (omrŒdene A, B, C, D)? Du kan markere mer enn ett svar.

- ❑ **A**
- ❑ **B**
- ❑ **C**
- ❑ **D**

A 7. Hvis du har smerter flere typer smerter, hvor sitter den mest plagsomme smerten? Velg bare ett av omrŒdene (A, B, C, D) fra diagram 1.

- ❑ **A**
- ❑ **B**
- ❑ **C**
- ❑ **D**

A 8. I l¿pet av det siste Œret, har dine daglige aktiviteter blitt pŒvirket av ubehag eller smerte i magen?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

De f¿lgende sp¿rsmŒlene angŒr smerte i magen som du mŒtte hatt det siste Œret .

A 9. I l¿pet av det siste Œret, hvor alvorlig var ubehaget eller smerten i magen?

- ❑ **Veldig mild**
- ❑ **Mild**
- ❑ **Moderat**
- ❑ **Sterk**

Noen ganger: mer enn en av ti uker
Ofte: mer enn en av fire uker
Svært ofte: mer enn halvparten av ukene

A 10. I l¿pet av det siste Œret, ble ubehaget eller smerten i magen bedre eller forsvant den etter at du hadde hatt avf¿ring?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

A 11. I l¿pet av et siste Œret, hadde du hyppigere avf¿ring enn vanlig nŒr ubehaget eller smerten begynte?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

A 12. I l¿pet av det siste Œret, hadde du sjeldnere avf¿ring enn vanlig nŒr ubehaget eller smerten begynte?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

A 13. I l¿pet av det siste Œret, hadde du l¿sere avf¿ring enn vanlig nŒr ubehaget eller smerten begynte?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

A 14. I l¿pet av det siste Œret, hadde du hardere avf¿ring enn vanlig nŒr ubehaget eller smerten begynte?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

A 15. I l¿pet av det siste Œret, ble ubehaget eller smerten bedre etter raping?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

Noen ganger: mer enn en av ti uker
Ofte: mer enn en av fire uker
Svært ofte: mer enn halvparten av ukene

A 16. I l¿pet av det siste Œret, forekom ubehaget eller smerten etter mŒltider?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

A 17. I l¿pet av det siste Œret, har ubehaget eller smerten vekket deg nŒr du sov?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

A 18. I l¿pet av det siste Œret, har ubehaget eller smerten noen gang strŒlt mot ryggen eller skuldrene?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

A 19. I l¿pet av det siste Œret, ble ubehaget eller smerten mindre nŒr du b¿yde deg fremover?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

A 20. I l¿pet av det siste Œret, nŒr du hadde ubehaget eller smerten, varte det lenger enn 20 minutter?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

A 21. Dersom du har ubehag eller smerte i ¿vre del av magen (omrŒdene merket A og B i diagram 1), blir det bedre nŒr du har hatt avf¿ring?

- ❑ **Ja**
- ❑ **Ikke i det hele tatt eller sjelden**
- ❑ **Jeg har ikke slik smerte.**

A 21. 1. NŒr du hadde ubehaget eller smerten i ¿vre del av magen, hadde du da oftere avf¿ring?

- ❑ **Ja**
- ❑ **Ikke i det hele tatt eller sjelden**

Noen ganger: mer enn en av ti uker
Ofte: mer enn en av fire uker
Svært ofte: mer enn halvparten av ukene

A 21. 2. NŒr du hadde ubehaget eller smerten i ¿vre del av magen, hadde du da sjeldnere avf¿ringer?

- ❑ **Ja**
- ❑ **Ikke i det hele tatt eller sjelden**

A 21. 3. NŒr du hadde ubehaget eller smerten i ¿vre del av magen, hadde du da hardere avf¿ringer?

- ❑ **Ja**
- ❑ **Ikke i det hele tatt eller sjelden**

A 21. 4. NŒr du hadde ubehaget eller smerten i ¿vre del av magen, hadde du da l¿sere avf¿ringer?

- ❑ **Ja**
- ❑ **Ikke i det hele tatt eller sjelden**

A 22. I l¿pet av det siste Œret nŒr du hadde ubehaget eller smerten, var det vanligvis i et lite avgrenset omrŒde som du kunne peket pŒ med en eller to fingre?

- ❑ **Ja**
- ❑ **Nei**

A 23. I l¿pet av det siste Œret nŒr du hadde ubehaget eller smerten, var det jevnt og kontinuerlig (det gikk ikke i b¿lger)?

- ❑ **Ja**
- ❑ **Nei**

A 24. I l¿pet av det siste Œret, har du hatt episoder med kontinuerlig sterk smerte (i midten av ¿vre av magen eller i h¿yre ¿vre del av magen) som varte i 30 minutter eller mer?

- ❑ **Ikke i det hele tatt**
- ❑ **1 eller 2 episoder med slik smerte**
- ❑ **Mer enn 2 episoder med slik smerte**

A 25. Har du noen gang hatt en operasjon der gallebl¾ren ble fjernet?

- ❑ **Ja**
- ❑ **Nei**

Vennligst sjekk at du har svart pŒ alle aktuelle sp¿rsmŒl i del A.
Vennligst begynn pŒ del B.

Noen ganger: mer enn en av ti uker
Ofte: mer enn en av fire uker
Svært ofte: mer enn halvparten av ukene

Del B
Vi vil gjerne stille deg sp¿rsmŒl om andre mageproblemer du mŒtte ha hatt det siste Œret .

B 1. I l¿pet av det siste Œret, har du f¿lt deg ubehagelig mett eller oppfylt like etter at du begynte Œ spise, slik at du ikke klarte Œ spise opp et normalt mŒltid?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

B 2. I l¿pet av det siste Œret, etter at du har spist normale mŒltider, har du hatt en ubehagelig f¿lelse av at maten har stoppet opp i magen?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

B 3. I l¿pet av det siste Œret, har du v¾rt kvalm (ville kaste opp men gjorde det ikke)?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

B 4. I l¿pet av det siste Œret, har du kastet opp uten at du selv eller medisiner du tok var Œrsaken til dette?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

B 5. I l¿pet av det siste Œret, har du kastet opp minst tre forskjellige dager hver uken i en tre mŒneds periode?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

Noen ganger: mer enn en av ti uker
Ofte: mer enn en av fire uker
Svært ofte: mer enn halvparten av ukene

B 6. I l¿pet av det siste Œret, har du hatt brekninger (som om du kaster opp, uten at noe kommer opp)?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

B 7. I l¿pet av et siste Œret, har du hatt f¿lelsen av Œ v¾re ÓstinnÓ, oppblŒst eller at hele magen eser ut?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

B 8. I l¿pet av det siste Œret, har du sett at magen din har eset ut?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

B 9. I l¿pet av det siste Œret, hva har v¾rt ditt mest plagsomme symptom i ¿vre del av magen? (omrŒdet A eller B i figur 1) (velg bare ett svar).

- ❑ **Magesmerter**
- ❑ **Svie, Óbrenningó i magen**
- ❑ **Ubehag i magen (ikke smerte)**
- ❑ **OppblŒsthet i hele magen**
- ❑ **Kvalme**
- ❑ **Maten Óstopper oppó i magen**
- ❑ **ÓStinnÓ, full i magen**
- ❑ **Ubehagelig mett like etter at du begynte Œ spise**
- ❑ **Jeg har ikke noen symptomer fra ¿vre del av magen**

Vennligst sjekk at du har svart pŒ alle sp¿rsmŒlene i del B.
Vennligst begynn pŒ del C.

Noen ganger: mer enn en av ti uker
Ofte: mer enn en av fire uker
Svært ofte: mer enn halvparten av ukene

Del C
Vi vil gjerne stille deg noen sp¿rsmŒl om andre helseproblemer du mŒtte ha hatt i l¿pet av det siste Œret.

C 1. I l¿pet av det siste Œret, har du v¾rt plaget med gjentatt raping?
- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

C 2. I l¿pet av det siste Œret, har du har svelget mye luft?
- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

C 3. I l¿pet av det siste Œret, har du hatt f¿lelse av klump i halsen nŒr du ikke svelger?
- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

C 4. I l¿pet av det siste Œret, har du hatt vanskeligheter med Œ svelge fast eller flytende f¿de (mat eller drikke setter seg fast eller passerer unormalt nedover)?
- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

C 5. I l¿pet av det siste Œret, har du mŒttet gulpe opp mat, tygge den pŒ nytt og enten spyttet det ut eller svelget den igjen?
- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

Hvis du har svart Sjelden eller aldri pŒ sp¿rsmŒl C 5, v¾r sŒ snill Œ gŒ til sp¿rsmŒl C 6.

Noen ganger: mer enn en av ti uker
Ofte: mer enn en av fire uker
Svært ofte: mer enn halvparten av ukene

C 5.1. Sjedde dette ved anledninger hvor du hadde v¾rt kvalm eller kastet opp?

- ❑ **Ja**
- ❑ **Nei**

C 5. 2. Sluttet du med dette etter at du hadde gulpet opp mat som smakte syrlig eller surt?

- ❑ **Ja**
- ❑ **Nei**

C 6. I l¿pet av det siste Æret, har du hatt brystbrann? Med dette mener vi brennende smerte eller ubehag bak brystbeinet stigende opp mot svelget. (Ikke ta med smerte pÆ grunn av angina eller hjerteproblemer).

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

C 7. I l¿pet av et siste Æret, har du hatt smerte eller ubehag midt i brystet?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

Vennligst sjekk at du har svart pÆ alle aktuelle sp¿rsmÆl i del C.
Vennligst begynn pÆ del D.

Noen ganger: mer enn en av ti uker
Ofte: mer enn en av fire uker
Svært ofte: mer enn halvparten av ukene

Del D
Vi vil gjerne stille deg noen sp¿rsmŒl om tarmproblemer du mŒtte ha hatt i l¿pet av det siste Œret.

D 1. I l¿pet av et siste Œret, har du hatt noen tarmproblemer?
- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

D 2. I l¿pet av de siste tre mŒnedene, har du hatt noen tarmproblemer?
- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

D 3. I l¿pet av et siste Œret, har du hatt flere enn tre avf¿ringer hver dag?
- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

D 4. I l¿pet av et siste Œret, har du hatt f¾rre enn tre avf¿ringer hver uke?
- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

D 5. I l¿pet av et siste Œret, har du hatt klumpete eller hard avf¿ring?
- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

D 6. I l¿pet av et siste Œret, har du hatt l¿se, gr¿tete eller vandig avf¿ring?
- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

Noen ganger: mer enn en av ti uker
Ofte: mer enn en av fire uker
Svært ofte: mer enn halvparten av ukene

D 7. I l¿pet av et siste Œret, har du f¿lt at du fortsatt ikke var helt t¿mt etter at du har hatt avf¿ring?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

D 8. I l¿pet av et siste Œret, har du mŒtte trykke hardt for Œ fŒ avf¿ring?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

D9. I l¿pet av det siste Œret, har du opplevd sŒ ¿yeblikkelig trang til Œ ha avf¿ring, at du mŒtte l¿pe til toalettet?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

D10 I l¿pet av det siste Œret, har du lagt merke til slim (hvitt slim) i avf¿ring?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

D11. I l¿pet av det siste Œret, har du hatt en f¿lelse av at endetarmsŒpning din har v¾rt blokkert mens du hadde avf¿ring

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

D12. I l¿pet av det siste Œret, mŒttet presse fingeren inn i eller rundt endetarmsŒpningen eller skjeden for Œ hjelpe avf¿ringen ut?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

Noen ganger: mer enn en av ti uker
Ofte: mer enn en av fire uker
Svært ofte: mer enn halvparten av ukene

D13. I l¿pet av det siste Œret, har du hatt vedvarende eller tilbakevendende verkende smerte eller trykk i endetarmsŒpningen eller endetarmsomrŒdet?

- ❑ **Sjelden eller aldri**
- ❑ **Noen ganger**
- ❑ **Ofte**
- ❑ **Sv¾rt ofte**
- ❑ **Nesten alltid**

Hvis du svarte Sjelden aller aldri pŒ sp¿rsmŒl D13, kan du gŒ videre til sp¿rsmŒl D16

D14. Da du hadde vedvarende eller tilbakevendende verkende smerte eller trykk i endetarmsŒpningen eller endetarmsomrŒdet, varte det da vanligvis

- ❑ **Fra sekunder til minutter og forsvant helt?**
- ❑ **20 minutter eller lenger?**

D15. Var det perioder av minst to ukers varighet mellom episodene med smerte eller trykk i endetarmsŒpningen eller endetarmsomrŒdet?

- ❑ **Ja**
- ❑ **Nei**

D16. I l¿pet av det siste Œret, nŒr du hadde treg eller l¿s avf¿ring, har du lekket, hatt avf¿ring eller blitt tils¿lt uten ville det?

- ❑ **Aldri**
- ❑ **En gang hver mŒned**
- ❑ **To ganger hver mŒned**
- ❑ **Mer enn to ganger hver mŒned**

D17. Skjedde dette (at du hadde avf¿ring uten at du hadde kontroll over det) oftest nŒr du f¿lte deg obstipert (treg avf¿ring)?

- ❑ **Ja**
- ❑ **Nei**

D18. Skjedde dette (at du hadde avf¿ring uten kontroll) oftest nŒr du hadde l¿s avf¿ring?

- ❑ **Ja**
- ❑ **Nei**

Vennligst sjekk at du svart pŒ alle aktuelle sp¿rsmŒl i del D
Takk.

Noen ganger: mer enn en av ti uker
Ofte: mer enn en av fire uker
Svært ofte: mer enn halvparten av ukene

(fylles ut av pasienten selv)

THE GASTROINTESTINAL SYMPTOM RATING SCALE (GSRS)
(Irritable bowel syndrome (IBS)-versjon)

Les dette først:
Undersøkelsen inneholder spørsmål om hvordan du har følt deg og hvordan du har hatt det DE SISTE TRE DAGENE. Sett kryss (X) ved det alternativet som passer best på deg og din situasjon.

Navn: _____

Fødselsdato: _____

Hvilke datoer dette gjelder: _____

1. Har du i løpet av de siste tre dagene vært plaget av MAGESMERTER?

- ☐ Ingen plager i det hele tatt
- ☐ Ubetydelige plager
- ☐ Milde plager
- ☐ Moderate plager
- ☐ Ganske alvorlige plager
- ☐ Alvorlige plager
- ☐ Meget alvorlige plager

2. Har du i løpet av den siste tre dagene vært plaget av SMERTER ELLER UBEHAG I MAGEN SOM GIR SEG NÅR DU HAR HATT AVFØRING?

- ☐ Ingen plager i det hele tatt
- ☐ Ubetydelige plager
- ☐ Milde plager
- ☐ Moderate plager
- ☐ Ganske alvorlige plager
- ☐ Alvorlige plager
- ☐ Meget alvorlige plager

(fylles ut av pasienten selv)

3. Har du i løpet av den siste tre dagene vært plaget av OPPBLÅSTHET?

- ☐ Ingen plager i det hele tatt
- ☐ Ubetydelige plager
- ☐ Milde plager
- ☐ Moderate plager
- ☐ Ganske alvorlige plager
- ☐ Alvorlige plager
- ☐ Meget alvorlige plager

4. Har du i løpet av den siste tre dagene vært plaget av LUFTAVGANG?

- ☐ Ingen plager i det hele tatt
- ☐ Ubetydelige plager
- ☐ Milde plager
- ☐ Moderate plager
- ☐ Ganske alvorlige plager
- ☐ Alvorlige plager
- ☐ Meget alvorlige plager

5. Har du i løpet av den siste tre dagene vært plaget av FORSTOPPELSE (problemer med å tømme tarmen)?

- ☐ Ingen plager i det hele tatt
- ☐ Ubetydelige plager
- ☐ Milde plager
- ☐ Moderate plager
- ☐ Ganske alvorlige plager
- ☐ Alvorlige plager
- ☐ Meget alvorlige plager

(fylles ut av pasienten selv)

6. Har du i løpet av den siste tre dagene vært plaget av DIARÉ (hyppig avføring)?

- ☐ Ingen plager i det hele tatt
- ☐ Ubetydelige plager
- ☐ Milde plager
- ☐ Moderate plager
- ☐ Ganske alvorlige plager
- ☐ Alvorlige plager
- ☐ Meget alvorlige plager

7. Har du i løpet av den siste tre dagene vært plaget av LØS AVFØRING?

- ☐ Ingen plager i det hele tatt
- ☐ Ubetydelige plager
- ☐ Milde plager
- ☐ Moderate plager
- ☐ Ganske alvorlige plager
- ☐ Alvorlige plager
- ☐ Meget alvorlige plager

8. Har du i løpet av de siste tre dagene vært plaget av HARD AVFØRING?

- ☐ Ingen plager i det hele tatt
- ☐ Ubetydelige plager
- ☐ Milde plager
- ☐ Moderate plager
- ☐ Ganske alvorlige plager
- ☐ Alvorlige plager
- ☐ Meget alvorlige plager

(fylles ut av pasienten selv)

9. Har du i løpet av den siste tre dagene vært plaget av TVINGENDE AVFØRINGSBEHOV (plutselig behov for å gå på toalettet for å tømme tarmen)?

- ☐ Ingen plager i det hele tatt
- ☐ Ubetydelige plager
- ☐ Milde plager
- ☐ Moderate plager
- ☐ Ganske alvorlige plager
- ☐ Alvorlige plager
- ☐ Meget alvorlige plager

10. Har du i løpet av de siste tre dagene vært plaget av en FØLELSE AV UFULLSTENDIG TØMMING AV TARMEN ETTER AVFØRING?

- ☐ Ingen plager i det hele tatt
- ☐ Ubetydelige plager
- ☐ Milde plager
- ☐ Moderate plager
- ☐ Ganske alvorlige plager
- ☐ Alvorlige plager
- ☐ Meget alvorlige plager

11. Har du i løpet av den siste tre dagene vært plaget av at du FØLER DEG METT LIKE ETTER AT DU HAR BEGYNT PÅ ET MÅLTID?

- ☐ Ingen plager i det hele tatt
- ☐ Ubetydelige plager
- ☐ Milde plager
- ☐ Moderate plager
- ☐ Ganske alvorlige plager
- ☐ Alvorlige plager
- ☐ Meget alvorlige plager

(fylles ut av pasienten selv)

12. Har du i løpet av den siste tre dagene vært plaget av at du FØLER DEG METT SELV LENGE ETTER AT DU ER FERDIG MED Å SPISE?

- ☐ Ingen plager i det hele tatt
- ☐ Ubetydelige plager
- ☐ Milde plager
- ☐ Moderate plager
- ☐ Ganske alvorlige plager
- ☐ Alvorlige plager
- ☐ Meget alvorlige plager

13. Har du i løpet av den siste tre dagene vært plaget av at MAGEN ER SYNLIG OPPBLÅST?

- ☐ Ingen plager i det hele tatt
- ☐ Ubetydelige plager
- ☐ Milde plager
- ☐ Moderate plager
- ☐ Ganske alvorlige plager
- ☐ Alvorlige plager
- ☐ Meget alvorlige plager

VENNLIGST KONTROLLER AT ALLE SPØRSMÅLENE ER BESVART!

(fylles ut av pasienten selv)

SHC Ð subjective health complains

Helseproblemer siste 30 d¿gn

PŒ den neste siden nevnes noen vanlige helseplager. Vi vil be deg om Œ vurdere hvert enkelt problem/symptom, og oppgi i hvilken grad du har v¾rt plaget **av dette i l¿pet av de siste tretti d¿gn, og** antall dager **du har v¾rt plaget.**

Eksempel
Hvis du f¿ler at du har v¾rt en del **plaget med forkj¿lelse/influensa siste mŒned og varigheten av plagene var** ca. en uke**, fylles dette ut pŒ f¿lgende mŒte:**
Sett ring rundt tallet som passer best.

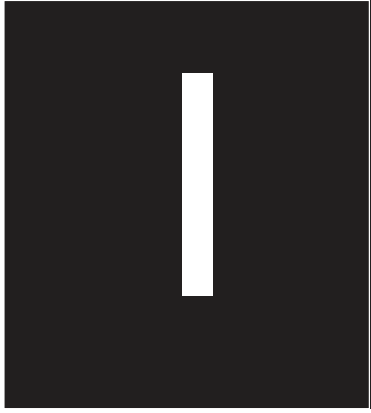
Nedenfor nevnes noen alminnelige helseproblemer	**Ikke plaget**	**Litt plaget**	**En del plaget**	**Alvorlig plaget**	**Antall dager plagene varte (omtrent)**
1. Forkj¿lelse, influensa	**0**	**1**	**(2)**	**3**	**7**

NB! Det er viktig at du fyller ut bŒde **hvor plaget** du har v¾rt, og **omtrent antall dager** du har v¾rt plaget siste tretti d¿gn.

(fylles ut av pasienten selv)

Helseproblemer siste 30 døgn

Nedenfor nevnes noen alminnelige helseproblemer **(sett ring rundt tallet som passer)**		Ikke plaget	Litt plaget	Endel plaget	Alvorlig plaget	Antall dager plagene varte (omtrent)
1.	**Forkj¿lelse, influensa**	**0**	**1**	**2**	**3**	…….
2.	**Hoste, bronkitt**	**0**	**1**	**2**	**3**	…….
3.	**Astma**	**0**	**1**	**2**	**3**	…….
4.	**Hodepine**	**0**	**1**	**2**	**3**	…….
5.	**Nakkesmerter**	**0**	**1**	**2**	**3**	…….
6.	**Smerter ¿verst i ryggen**	**0**	**1**	**2**	**3**	…….
7.	**Smerter i korsrygg**	**0**	**1**	**2**	**3**	…….
8.	**Smerter i armer**	**0**	**1**	**2**	**3**	…….
9.	**Smerter i skuldre**	**0**	**1**	**2**	**3**	…….
10.	**Migrene**	**0**	**1**	**2**	**3**	…….
11.	**Hjertebank, ekstraslag**	**0**	**1**	**2**	**3**	…….
12.	**Brystsmerter**	**0**	**1**	**2**	**3**	…….
13.	**Pustevansker**	**0**	**1**	**2**	**3**	…….
14.	**Smerter i f¿ttene ved anstrengelser**	**0**	**1**	**2**	**3**	…….
15.	**Sure oppst¿t, ÇhalsbrannÈ**	**0**	**1**	**2**	**3**	…….
16.	**Sug eller svie i magen**	**0**	**1**	**2**	**3**	…….
17.	**Magekatarr, magesŒr**	**0**	**1**	**2**	**3**	…….
18.	**Mageknip**	**0**	**1**	**2**	**3**	…….
19.	**ÇLuftplagerÈ**	**0**	**1**	**2**	**3**	…….
20.	**L¿s avf¿ring, diarŽ**	**0**	**1**	**2**	**3**	…….
21.	**Forstoppelse**	**0**	**1**	**2**	**3**	…….
22.	**Eksem**	**0**	**1**	**2**	**3**	…….
23.	**Allergi**	**0**	**1**	**2**	**3**	…….
24.	**Hetetokter**	**0**	**1**	**2**	**3**	…….
25.	**S¿vnproblemer**	**0**	**1**	**2**	**3**	…….
26.	**Tretthet**	**0**	**1**	**2**	**3**	…….
27.	**Svimmelhet**	**0**	**1**	**2**	**3**	…….
28.	**Angst**	**0**	**1**	**2**	**3**	…….
29.	**Nedtrykt, depresjon**	**0**	**1**	**2**	**3**	…….



ORIGINAL ARTICLE

Familial Diarrhea Syndrome Caused by an Activating *GUCY2C* Mutation

Torunn Fiskerstrand, M.D., Ph.D., Najla Arshad, Ph.D., Bjørn Ivar Haukanes, Ph.D., Rune Rose Tronstad, M.D., Khanh Do-Cong Pham, M.D., Stefan Johansson, Ph.D., Bjarte Håvik, Ph.D., Siv L. Tønder, M.Sc., Shawn E. Levy, Ph.D., Damien Brackman, M.D., Ph.D., Helge Boman, M.D., Ph.D., Kabir Hassan Biswas, Ph.D., Jaran Apold, M.D., Ph.D., Nils Hovdenak, M.D., Ph.D., Sandhya S. Visweswariah, Ph.D., and Per M. Knappskog, Ph.D.

ABSTRACT

BACKGROUND

Familial diarrhea disorders are, in most cases, severe and caused by recessive mutations. We describe the cause of a novel dominant disease in 32 members of a Norwegian family. The affected members have chronic diarrhea that is of early onset, is relatively mild, and is associated with increased susceptibility to inflammatory bowel disease, small-bowel obstruction, and esophagitis.

METHODS

We used linkage analysis, based on arrays with single-nucleotide polymorphisms, to identify a candidate region on chromosome 12 and then sequenced *GUCY2C*, encoding guanylate cyclase C (GC-C), an intestinal receptor for bacterial heat-stable enterotoxins. We performed exome sequencing of the entire candidate region from three affected family members, to exclude the possibility that mutations in genes other than *GUCY2C* could cause or contribute to susceptibility to the disease. We carried out functional studies of mutant GC-C using HEK293T cells.

RESULTS

We identified a heterozygous missense mutation (c.2519G→T) in *GUCY2C* in all affected family members and observed no other rare variants in the exons of genes in the candidate region. Exposure of the mutant receptor to its ligands resulted in markedly increased production of cyclic guanosine monophosphate (cGMP). This may cause hyperactivation of the cystic fibrosis transmembrane regulator (CFTR), leading to increased chloride and water secretion from the enterocytes, and may thus explain the chronic diarrhea in the affected family members.

CONCLUSIONS

Increased GC-C signaling disturbs normal bowel function and appears to have a proinflammatory effect, either through increased chloride secretion or additional effects of elevated cellular cGMP. Further investigation of the relevance of genetic variants affecting the GC-C–CFTR pathway to conditions such as Crohn's disease is warranted. (Funded by Helse Vest [Western Norway Regional Health Authority] and the Department of Science and Technology, Government of India.)

From the Center for Medical Genetics and Molecular Medicine (T.F., B.I.H., S.J., B.H., S.L.T., H.B., J.A., P.M.K.) and the Departments of Internal Medicine (K.D.C.-P., N.H.) and Pediatrics (R.R.T., D.B.), Haukeland University Hospital, and the Departments of Clinical Medicine (T.F., B.H., P.M.K.) and Biomedicine (S.J.), University of Bergen — both in Bergen, Norway; the Department of Molecular Reproduction, Development and Genetics, Indian Institute of Science, Bangalore, India (N.A., K.H.B., S.S.V.); and HudsonAlpha Institute for Biotechnology, Huntsville, AL (S.E.L.). Address reprint requests to Dr. Fiskerstrand at the Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, 5021 Bergen, Norway, or at torunn.fiskerstrand@helse-bergen.no; or to Dr. Visweswariah at the Department of Molecular Reproduction, Development and Genetics, Indian Institute of Science, Bangalore 560012, India, or at sandhya@mrdg.iisc.ernet.in.

This article (10.1056/NEJMoa1110132) was published on March 21, 2012, at NEJM.org.

Drs. Fiskerstrand and Arshad, and Drs. Haukanes and Tronstad, contributed equally to this article.

N Engl J Med 2012;366:1586-95.

Chronic diarrhea is a health problem that poses challenges with respect to both diagnosis and treatment. The irritable bowel syndrome affects 15 to 20% of adults and is a common cause of diarrhea.[1] Other causes include inflammatory bowel disease, infections, paraneoplastic hormones, celiac disease, malabsorption syndromes, and bacterial overgrowth in the small intestine.[2] In addition to organic causes, psychological factors have an important effect on bowel function.[1]

Recent studies have focused on the importance of genetic factors in the development of chronic diarrhea[3-5] because genetic factors can provide insight into the pathophysiology of intestinal diseases and point to new treatments. The irritable bowel syndrome aggregates strongly in families,[6] and the genetic predisposition to Crohn's disease, particularly ileitis,[7] is well documented,[3] but no major genes causing these disorders have been found. Both diseases are considered to be multifactorial, and the causes include genetic susceptibility variants and environmental factors. Rare inherited forms of chronic diarrhea have been reported[4]; nearly all of these are severe, autosomal recessive, single-gene diseases that are manifested in the newborn period.

There is a fine balance between intestinal absorption and secretion of water and electrolytes.[2,8,9] The secretory capacity of the small intestine is substantial, as evidenced by the potentially life-threatening secretory diarrhea that results from enterotoxigenic *Escherichia coli* and *Vibrio cholerae* infections.[9] Heat-stable enterotoxins from *E. coli* bind to intestinal guanylate cyclase C (GC-C) receptors, resulting in elevated levels of cellular cyclic guanosine monophosphate (cGMP). cGMP elicits a signaling cascade involving protein kinases and the cystic fibrosis transmembrane conductance regulator (CFTR) (Fig. 1), ultimately causing increased chloride secretion through CFTR. This creates an osmotic drive that results in the secretion of sodium ions, and hence water, into the intestinal lumen (Fig. 1).[2,9,10]

We evaluated a large family (Fig. 2A) with a dominantly inherited, fully penetrant syndrome of chronic diarrhea and dysmotility. Other conditions present in some members of this family included Crohn's disease, small-bowel obstruction, and esophagitis with or without esophageal hernia. We performed whole-genome single-nucleotide-polymorphism (SNP)–based linkage analysis, as well as exome sequencing, and identified a heterozygous activating mutation in *GUCY2C*, encoding GC-C. We thus established a genetic cause for this novel inherited disease.

METHODS

PARTICIPANTS

We studied 32 affected persons from three branches of the same family (Fig. 2A), as well as 14 unaffected family members. We obtained written informed consent from all the participants in the study. The affected family members were examined by a gastroenterologist and completed questionnaires (Rome II and ad hoc forms) regarding bowel symptoms. The study was approved by the regional ethics committee of Western Norway.

DETECTION OF MUTATIONS

Genomic DNA was purified from blood with the use of the QIAsymphony system (Qiagen). Whole-genome genotyping of SNPs was performed with the use of GeneChip Human Mapping 250K NspI array (Affymetrix). All exons and flanking intron sequences of *GUCY2C* (National Center for Biotechnology Information [NCBI] reference sequence NM_004963.3), were amplified and sequenced by means of standard methods. Exome sequencing was performed with the use of HiSeq (Illumina) (for details, see the Supplementary Appendix, available with the full text of this article at NEJM.org).

SITE-DIRECTED MUTAGENESIS

We generated mutant complementary DNA (cDNA) containing the c.2519G→T [p.Ser840Ile] as described previously.[11] Nonmutant and mutant GC-C cDNAs were cloned into the mammalian expression vector pcDNA3 (Invitrogen), and the respective proteins (nonmutant GC-C and mutant GC-CS840I) were transiently expressed in HEK293T cells.[11]

CHARACTERIZATION OF MUTANT PROTEIN GC-CS840I

We measured ligand-stimulated GC-C activity in intact cells 72 hours after transfection, after the addition of varying concentrations of heat-stable enterotoxin (STa) or uroguanylin (for 15 minutes) or guanylin (for 60 minutes).[11] For in vitro guanylate cyclase assays, membranes were prepared from transfected cells.[11] Details are provided in the Supplementary Appendix.

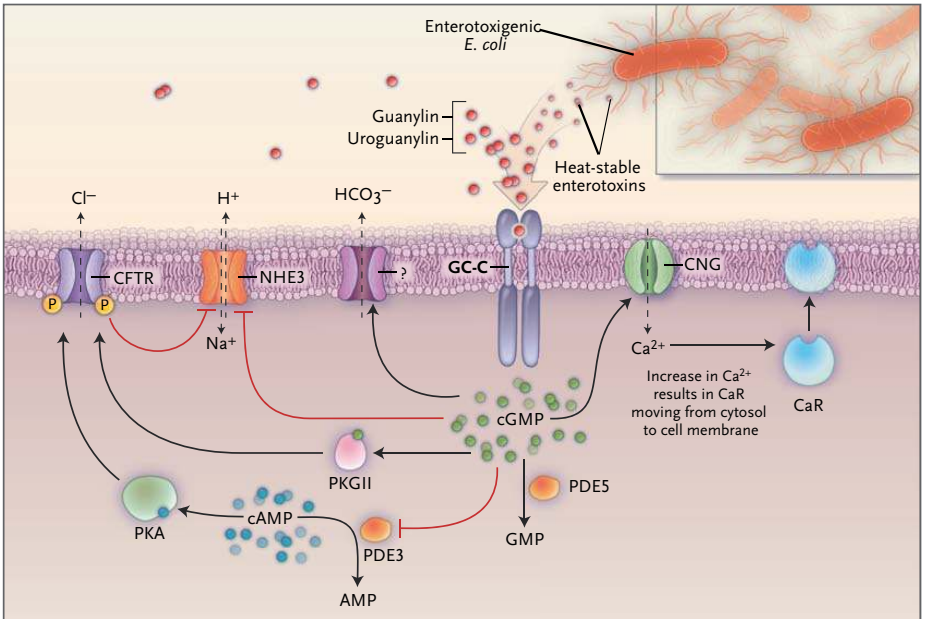


Figure 1. Signaling Mechanisms Mediated by Guanylate Cyclase C (GC-C).

GC-C expressed on the surface of enterocytes serves as the receptor for the endogenous ligands uroguanylin and guanylin or for heat-stable enterotoxins produced by enterotoxigenic *Escherichia coli*. Guanylin-family hormones are synthesized in the intestine and released both luminally and into the circulation. Uroguanylin exerts a natriuretic effect in the kidney. Ligand binding to GC-C increases intracellular levels of cyclic guanosine monophosphate (cGMP). The cGMP activates cGMP-dependent protein kinase II (PKGII) and inhibits the activity of a cyclic AMP (cAMP) phosphodiesterase, PDE3, thereby cross-activating cAMP–dependent protein kinase (PKA). PKGII and PKA phosphorylate the cystic fibrosis transmembrane conductance regulator (CFTR), increasing its chloride-secreting activity. In addition, cGMP enhances duodenal bicarbonate secretion through an unknown channel in a CFTR-dependent manner. These processes are involved in the maintenance of fluid and ion homeostasis. The cGMP also directly activates cyclic nucleotide gated channels (CNGs), leading to Ca^{2+} influx. Elevated intracellular Ca^{2+} levels activate calcium-sensing receptors (CaRs), promoting cell differentiation and migration. GC-C signaling is terminated by hydrolysis of cGMP to GMP by a cGMP-dependent phosphodiesterase, PDE5.[10] NHE3 denotes sodium–hydrogen exchanger 3.

HETERODIMERIZATION AND INTERACTION OF GC-C

To monitor heterodimerization, plasmids harboring either nonmutant GC-C fused at the C-terminal to glutathione S-transferase[11] or mutant GC-CS840I were mixed in varying ratios and cotransfected. Seventy-two hours after transfection, solubilized membrane protein was incubated with glutathione beads for 1 hour and then washed, and protein bound to the beads was detected by Western blot analysis (for details, see the Supplementary Appendix).

STATISTICAL ANALYSIS

Linkage analysis was used to determine the candidate region for the mutation in the family. Multipoint parametric linkage analysis and haplotyping were performed with the use of the Allegro program, version 2, on a set of 45,000 SNPs pruned for strong local linkage disequilibrium. Details of the analysis are provided in the Supplementary Appendix.

RESULTS

PARTICIPANTS

We studied 32 affected family members (14 females and 18 males), with a mean age of 44 years (range, 2 to 89), and examined their medical histories. Family branch A came to our attention first, when the 88-year-old index member (A-IV1) was admit-

ted to the hospital owing to dehydration (Fig. 2A, and Table 1 in the Supplementary Appendix). The affected family members in branch A typically had a history of chronic diarrhea that started in infancy and was fairly constant over the years but that tended to subside by middle age in some persons. The results of colonoscopic examinations of affected family members in branch A were essentially normal (Table 1 in the Supplementary Appendix). Across all branches, the affected family members had an average of 3.6 stools per day (range, 0.3 to 20.0); the stools typically had a watery or loose consistency and were accompanied by meteorism and in some cases abdominal pain. The pattern of inheritance was autosomal dominant, with some variation in expression. Four of the 32 cases (in family members C-VI2, B-VI1, B-VI3, and C-VII3) had previously been diagnosed as the irritable bowel syndrome, although they did not strictly meet the Rome II criteria.[12] However, 5 other cases (in family members A-IV1, A-V3, A-VI6, B-VII4, and B-VII5) did meet these criteria.

Some members of family branches B and C had more severe phenotypes (Table 1 in the Supplementary Appendix). Ten family members underwent laparotomy for suspected bowel obstruction; in eight of them (including family member B-IV1, who is not included in Table 1 in the Supplementary Appendix), obstruction resulting from volvulus, adhesional bands, or ileal inflammation was confirmed. Two family members with bowel obstruction had anatomical variants in the ileocecal region, such as a partly nonfixated ascending colon and slits in the ileal mesenterium. Three family members with bowel obstruction (B-V2, B-V3, and C-VI3) underwent a second operation to resolve adhesions, and family members C-VI3 and B-IV1 were described as having adhesional-band obstruction, even on the basis of the first laparotomy. Five family members with bowel obstruction (B-V2, B-V3, B-VI5, B-VI7, and B-VI9) eventually underwent resection of the terminal ileum and in some cases also the cecum; four of the five had verified or suspected Crohn's disease (Table 1 in the Supplementary Appendix). In addition, Crohn's disease was diagnosed in family members B-VI8 and C-VI2, and this diagnosis was suspected in family member C-V1. Family member A-VI6 received a diagnosis of possible eosinophilic enteritis.

Eight family members had been hospitalized for dehydration, metabolic acidosis, and electrolyte disturbances when they were newborns (Table 1 in the Supplementary Appendix). They were found to have hyponatremia, hypokalemia, and in some cases also hypomagnesemia and hypocalcemia, accompanied by abdominal distention and dilatation of the small intestine. None had a confirmed infectious disease. Hirschsprung's disease was ruled out in three of these family members, including B-VII1, who at 2 years of age had periods of constipation alternating with diarrhea. Seven other family members were hospitalized for dehydration at various times later in life, usually with infectious gastroenteritis, which, according to some of the family members, was followed by a prolonged period of recovery (Table 2 in the Supplementary Appendix). Several other conditions that members of this family had may be associated with this inherited condition (Table 2 in the Supplementary Appendix), including urolithiasis (in four family members), vitamin B12 deficiency (in six) and esophagitis with or without esophageal hernia (in five). We observed no evidence of behavioral disturbances[13] or a tendency toward obesity or excessive leanness.[14] Most of the family members reported food sensitivity, and several limit their intake of fruits, vegetables, and sweets.

POTENTIALLY CAUSATIVE MUTATION

Linkage analysis of samples from 11 affected members and 14 healthy members of family branch A revealed only one significant shared region in the affected members, on the short arm of chromosome 12 (12p), with a maximum LOD score of 5.1 (Fig. 1 in the Supplementary Appendix). A haplotype spanning approximately 2.9 megabases (Mb) (base pairs 14,466,726 to 17,410,570 from the start of 12p) showed complete cosegregation with the disease in the family. The region contained 28 putative protein-coding genes (Fig. 2B). Among these, *GUCY2C* (Fig. 2B) was the best candidate, because it encodes GC-C, an intestinal transmembrane receptor with known function in heat-stable enterotoxin–mediated diarrhea.[10] Sequence analysis identified a heterozygous base substitution, c.2519G→T, in exon 22 (Fig. 2C), predicting the replacement of the amino acid serine in codon 840 with isoleucine (p.Ser840Ile). Whole-exome sequencing in 3 persons (1 from each family branch) did not identify any other rare coding variant (Table 3 in the Supplementary Appendix). Using Sanger sequencing, we found the *GUCY2C* c.2519G→T mis-

Figure 2 (facing page). Identification of a Pathogenic Mutation in a Large Family with Chronic Diarrhea.
Panel A shows the pedigree of the family. All affected males (black squares) except B-IV1 and C-IV1 and all affected females (black circles) except A-IV5 and B-IV3 were investigated. The inheritance pattern is autosomal dominant, and linkage analysis was performed on samples from 14 healthy members and 11 affected members of branch A. The numbers of the generations are shown in Roman numerals to the left. Family members in each generation are designated by Arabic numbers from left to right, starting on 1 within each branch (with numbers shown only for the first and last person within each branch). Information about disease status in the affected members in generations I, II, and III was limited; therefore these persons are shown as unaffected (white squares or circles). Slashes denote deceased family members. Panel B shows the map of chromosome 12 and the 2.9-Mb region identified by linkage analysis (www.ensembl.org). The region on the short arm of chromosome 12 (p13.1 to p12.3) extends from base pair 14,466,726 to base pair 17,410,570 from the start of the short arm (according to NCBI Build 36.3). It contains 28 putative protein-coding genes, including *GUCY2C* (red circle), which is known to be the receptor for the *E. coli* heat-stable enterotoxin. Panel C shows the heterozygous missense mutation (c.2519G→T, arrow) in exon 22 in *GUCY2C*, which was found in all affected family members. This base change predicts the replacement of the amino acid serine in codon 840 with isoleucine [p.Ser840Ile] in the GC-C protein. Panel D shows the domain organization of GC-C. The GC-C protein contains an extracellular domain that binds its ligands. A single transmembrane domain is followed by a juxtamembrane domain and a kinase homology domain that bears sequence similarity to protein kinases. This is followed by a linker region and finally the C-terminal guanylate cyclase domain. Ligand binding to the extracellular domain results in activation of the catalytic domain, leading to increased cGMP production. GC-C forms oligomers in cells. A homodimer is shown; the catalytic domains of guanylate cyclases are functional only when in the context of a dimer. Highlighted in the red box is the position of S840 (red sphere). Numbers indicate amino acid positions of human GC-C, which includes a signal sequence predicted to be 23 amino acids long. The modeling of the various domains was performed as described previously.[10]

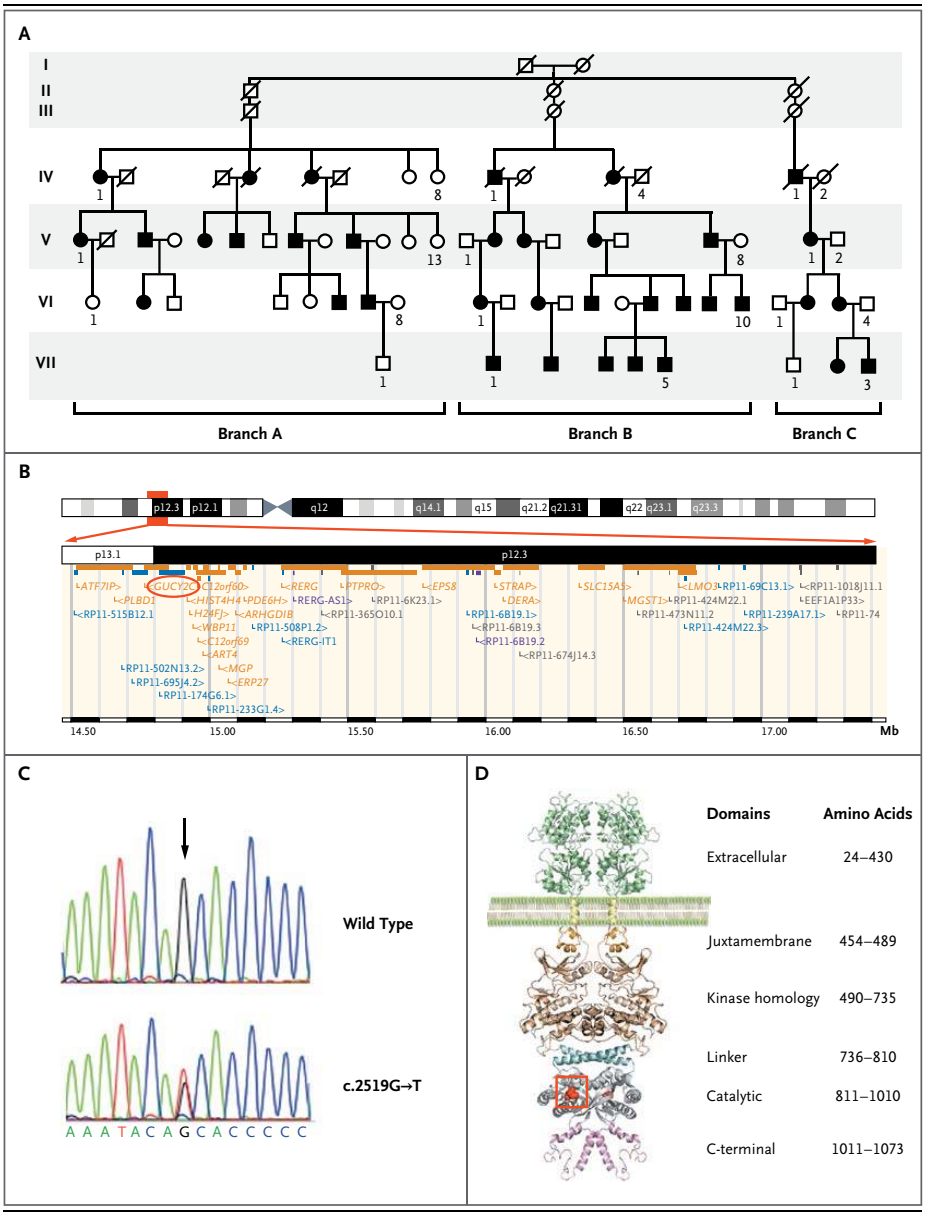
sense mutation to be present in all affected family members. We did not find this mutation in the NCBI human dbSNP database (build 132) or in 190 local healthy blood donors. None of the 14 unaffected family members we tested carried the mutation.

The amino acid Ser840 is highly conserved among mammalian GC-C proteins[15] and also in chicken and zebrafish (Fig. 2 in the Supplementary Appendix). The substitution is located in the catalytic domain (Fig. 2D), and we hypothesized that it may alter the guanylate cyclase activity of the mutant receptor.

FUNCTIONAL CHARACTERIZATION OF THE MUTATION

We performed biochemical experiments using intact transfected cells (Fig. 3, and Fig. 3 in the Supplementary Appendix) and membranes isolated from these transfected cells (Fig. 4 through 9 in the Supplementary Appendix), expressing either normal (nonmutant) or mutant (S840I) protein in equal amounts; the results of Western blot analysis are shown in Figure 3A, and in Figures 3 and 4 in the Supplementary Appendix. The basal GC-C enzyme activity and cellular cGMP levels (Fig. 3A, and Fig. 4 and 5 in the Supplementary Appendix) and affinities of ligands (heat-stable enterotoxin, uroguanylin, and guanylin) (Fig. 7 and 8 in the Supplementary Appendix) were similar in cells expressing the normal receptor and those expressing the mutant receptor. However, heat-stable enterotoxin (Fig. 3A and 3B, and Fig. 4 and 6 in the Supplementary Appendix), uroguanylin (Fig. 3C), and guanylin (Fig. 3 in the Supplementary Appendix) activated the mutant receptor to a greater extent than the nonmutant receptor. When cells were treated with 10^{-7} M heat-stable enterotoxin (Fig. 3B), 10^{-6} M uroguanylin (Fig. 3C), or 10^{-6} M guanylin (Fig. 3 in the Supplementary Appendix), cGMP production was increased by a factor of 7 to 8 (in the case of heat-stable enterotoxin), 5 (in the case of uroguanylin), or 3 to 4 (in the case of guanylin) with the mutant receptor, as compared with the nonmutant receptor. With respect to heat-stable enterotoxin, there was no significant difference between mutant and nonmutant receptors in the concentration required for half-maximal activation (EC_{50}) of the receptor (Fig. 3B). In contrast, with respect to uroguanylin, the EC_{50} of the mutant receptor was lower than that of the nonmutant receptor by a factor of about 5 to 6 (Fig. 3C), indicating that uroguanylin acts more potently on the mutant receptor. This suggests that the concentrations of uroguanylin present in the intestine could result in abnormally elevated levels of cellular cGMP in intestinal cells harboring the mutant receptor.

In patients who are heterozygous for the mutation, nonmutant and mutant receptors may be co-



expressed in a single cell. We therefore engineered the coexpression of a nonmutant receptor (fused with glutathione *S*-transferase) and the mutant receptor in HEK293T cells. The mutant receptor along with the glutathione *S*-transferase–tagged nonmutant receptor was captured by glutathione beads (Fig. 9 in the Supplementary Appendix), indicating the formation of heterodimers. There was a significant increase in the production of cGMP when the two receptors were coexpressed (in a 1:1 ratio), as compared with the production of cGMP when the nonmutant receptor was ex-

Figure 3 (facing page). Functional Analyses of Mutant GC-C Protein in HEK293T Cells.

Panel A shows heat-stable enterotoxin (ST)–mediated cyclic GMP (cGMP) production by GC-C. HEK293T cells were transfected with plasmids that allowed expression of either nonmutant GC-C or mutant GC-C (S840I). Equivalent expression of nonmutant or mutant GC-C was observed on Western blot analysis with the use of a monoclonal antibody to GC-C (inset). GC-C migrates as two differentially glycosylated proteins of molecular weight 130 and 145 kD.[11] Cells were treated with medium alone or with ST (10^{-7} M) for 15 minutes, and intracellular cGMP levels were measured by radioimmunoassay. Data shown are from one experiment (two measurements), and the experiment was repeated twice. T bars indicate standard errors. Panel B shows the dose response with ST. Cells expressing either nonmutant GC-C or mutant GC-C (S840I) were treated for 15 minutes with varying concentrations of ST as indicated, and intracellular cGMP levels were measured. The values in the graph are the mean values of a representative experiment; I bars indicate standard errors. The experiment (two measurements) was repeated three times with similar findings, and the concentration required for half-maximal activation (EC_{50} nanomolar values) was calculated from six measurements. EC_{50} values did not differ significantly between nonmutant and mutant GC-C. Panel C shows the dose response with uroguanylin. Cells expressing either nonmutant GC-C or mutant GC-C (S840I) were treated for 15 minutes with varying concentrations of uroguanylin as indicated, and intracellular cGMP levels were measured. The values in the graph are the mean values in a representative experiment; I bars indicate standard errors. The experiment (two measurements) was repeated three times with similar findings, and the EC_{50} values were calculated from six measurements. EC_{50} (micromolar values) differed significantly between nonmutant and mutant GC-C ($P<0.001$). Panel D shows heterodimerization of nonmutant GC-C and mutant GC-C (S840I). Indicated ratios of plasmids expressing either nonmutant GC-C tagged with glutathione *S*-transferase (GST) or mutant GC-C were cotransfected in HEK293T cells. Nonmutant GC-C–GST migrates at a size of approximately 160 kD owing to the C-terminal fusion of GST, and mutant GC-C migrates at 130 and 145 kD. The relative expression of the two proteins was monitored by Western blot analysis with the use of a monoclonal antibody to GC-C. At 72 hours after transfection, cells were treated with ST (10^{-7} M) for 15 minutes, and intracellular cGMP was measured by radioimmunoassay. Data shown represent the means of duplicate measurements, and the experiments were repeated twice. T bars indicate standard errors.

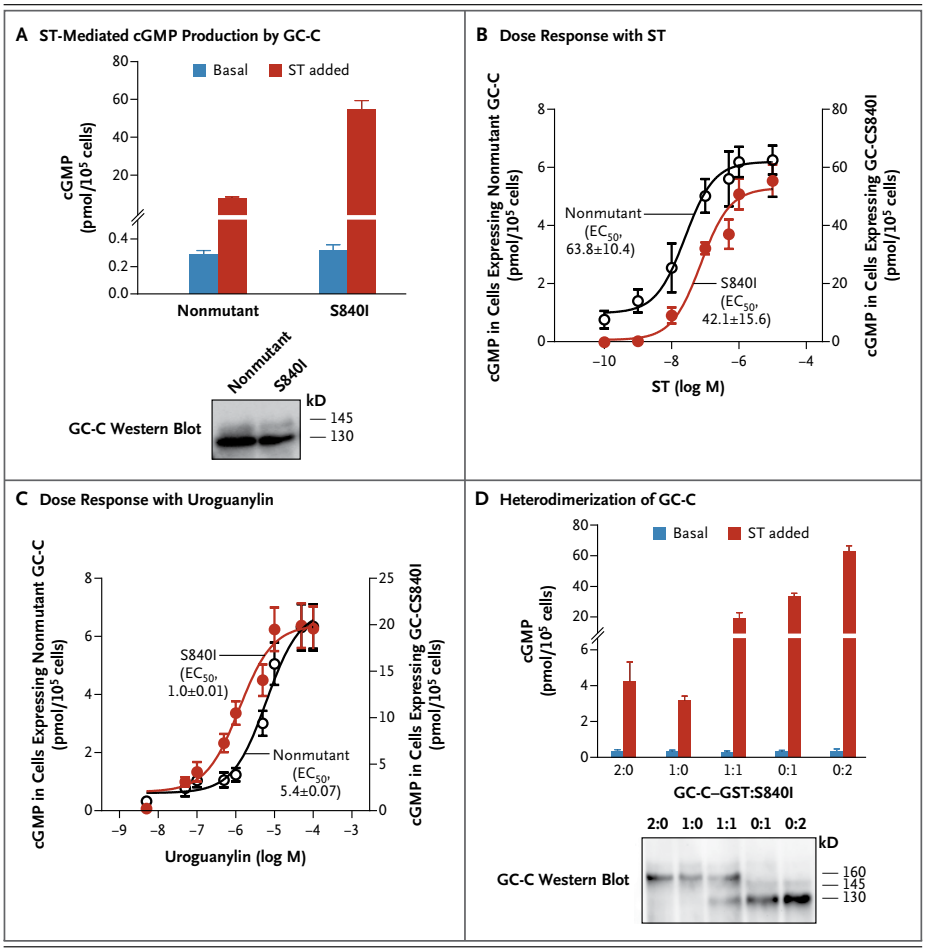
pressed alone (2:0 or 1:0 ratio) (Fig. 3D). This elevated activity was not as high as the activity observed with the mutant receptor alone (0:1 or 0:2 ratio), indicating that heterodimerization of the nonmutant and mutant receptors may blunt the ligand-mediated hyperactivation of the mutant receptor.

DISCUSSION

We describe a previously unrecognized disease characterized by familial diarrhea and define its genetic cause. The chronic diarrhea showed an inheritance pattern that was consistent with a dominant mutation in an autosomal gene, and we observed a heterozygous missense mutation, c.2519G→T [p.Ser840Ile], in *GUCY2C* in all affected family members (Fig. 2C and 2D). Subsequent functional analyses of the mutation in cell culture suggested that it effects a gain of function, in that it increases ligand-mediated activation of GC-C with subsequent intracellular accumulation of cGMP (Fig. 3). Persons with this mutation would therefore be prone to the production of comparatively high levels of cGMP in response to normal levels of endogenous ligands.

The importance of GC-C signaling and CFTR function in infectious bowel disease is well established.[8-10,16] The binding of heat-stable enterotoxin to GC-C greatly increases the formation of intracellular cGMP, which activates protein kinase GII, leading to phosphorylation of the CFTR channel.[10] This results in efflux of Cl^- (or HCO_3^- in the duodenum) and water into the intestinal lumen (Fig. 1), as well as reduced sodium absorption through inhibition of the sodium–hydrogen exchanger 3 (NHE3).[16] The disease we describe here may thus represent a "knock-in" genetic model of enterotoxigenic *E. coli* infection, with increased signaling through cGMP explaining the diarrhea in the affected family members, who were particularly prone to diarrhea, dehydration, and electrolyte disturbances soon after birth (Table 1 in the Supplementary Appendix). This is consistent with the observation that GC-C in the intestine is most highly expressed in newborns and declines during the first months of life to the level found in adults.[17]

The symptoms in the family members we examined overlapped with those of common disorders such as the irritable bowel syndrome (particularly in nine family members) and Crohn's disease (which was confirmed or suspected in seven members). This familial diarrhea syndrome also appears to confer a predisposition in the affected family members to small-bowel obstruc-



tion (which has occurred in eight family members) and esophagitis or esophageal hernia (in five members). Thus, the spectrum of symptoms seen in our patients suggests that the increased level of cGMP present in cells harboring the c.2519G→T [p.Ser840Ile] mutation affects inflammation and motility in the gut.

cGMP signaling has been linked to the regulation of inflammation and cell proliferation.[18,19] It has been shown that mice deficient in *Gucy2c* have a reduced inflammatory response to a colitis-inducing agent, owing to reduced expression of proinflammatory molecules.[20] The possibility that increased cGMP signaling enhances the expression of proinflammatory cytokines should be tested, since that may explain the susceptibility to ileitis in our patients. An increased risk of inflammatory bowel disease is also seen in connection with other conditions that cause intestinal electrolyte disturbances. This increased risk is observed in mice that are deficient in *Nhe3*[21] and in patients with congenital chloride diarrhea.[22] Changes in electrolyte homeostasis and increased permeability[23,24] may cause a disturbance in intestinal barrier function, which was recently proposed as a primary contributor to the development of inflammatory bowel disease.[25,26] Disturbed intestinal barrier function may also be caused by chronic changes in tight-junction function or assembly (or both),[26] which are affected by GC-C signaling.[27]

The terminal ileum in the affected family members we examined appeared to be susceptible to

small-bowel obstruction. Eight of 33 family members underwent laparotomy owing to small-bowel obstruction (Table 1 in the Supplementary Appendix); some family members required a second or even third laparotomy owing to reobstruction, which is frequently related to the formation of adhesions. All the affected family members had meteorism, and dilatation of the small intestine was seen during quiescent phases in many of them (Table 1 in the Supplementary Appendix). The latter, which is also observed in children with congenital chloride diarrhea,[22] may result from increased fluid in the intestine, reduced tone of smooth muscle, or both. The ileocecal junction was not dilated on radiographs or at surgery, and it may represent a bottleneck for the increased fluid formation. Mechanical strain or abnormal mobility in the ileocecal region could increase the risk of recurrent colic and bowel obstruction.[28] The molecular mechanisms underlying the disturbed motility in the gut and the meteorism in these patients are not known.

Both guanylin and cGMP have a relaxing effect on smooth-muscle cells in the gastrointestinal tract.[29,30] Levels of cGMP are altered by several effectors in the intestine, including other natriuretic peptides and the strong muscle relaxant and secretory agent, vasoactive intestinal peptide,[31] suggesting the possibility of cross-talk between these signaling molecules and GC-C.[32,33] Such cross-talk is of interest in a wider perspective, because (pro)uroguanylin is also secreted into the bloodstream and affects renal electrolyte transport in humans,[34] as well as behavior[13] and satiety (as part of the gut–brain axis)[14] in animals. *GUCY2C* is expressed in the brain,[13,14] but we observed no symptoms consistent with an effect of the mutant protein on the central nervous system.

Our data are consistent with those obtained through study of the role of GC-C in intestinal hydration, meteorism, and bowel function.[18,20,35] Both low activity[35] and high activity of this receptor-cyclase may interfere with homeostasis in the intestine. To our knowledge, no common low-risk susceptibility variants in *GUCY2C* for either the irritable bowel syndrome or inflammatory bowel disease have been reported.[3,5] Forthcoming genomewide association studies that use next-generation sequencing may reveal rare susceptibility variants. Also, variants in any of the genes related to the GC-C–CFTR pathway probably contribute to overall signaling through GC-C, making the whole pathway an interesting target for the investigation of diarrhea or constipation of unknown cause.

In conclusion, we have identified an activating mutation in *GUCY2C* as the cause of a novel familial diarrhea syndrome, characterized by meteorism, abdominal pain, dysmotility, and inflammatory bowel disease. This finding highlights *GUCY2C* as a susceptibility gene for Crohn's disease, small-bowel obstruction, and functional gastrointestinal diseases such as the irritable bowel syndrome and points to the GC-C–CFTR pathway as a target for a further search of mechanisms underlying these conditions.

Supported by grants from Helse Vest (Western Norway Regional Health Authority) (911308, to Drs. Fiskerstrand, Haukanes, Johansson, Boman, and Knappskog) and the Department of Science and Technology, Government of India (to Dr. Visweswariah). Dr. Arshad was supported by a fellowship from the Indian Institute of Science, and Dr. Biswas was supported by a Senior Research Fellowship from the Council of Scientific and Industrial Research, Government of India.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank all the patients and their families for participating in this study; Guri Matre, Jorunn Skeie Bringsli, Paal Borge, Hilde Rusaas, Suhas Ballal, and Vani R. Iyer for technical assistance; Dr. Francisco Roque for assistance with exome sequencing data analysis; Dr. Birgitte Emken for assistance with clinical characterization of the patients; and Drs. Ian M. Graham, Gunnar Houge, and Vidar M. Steen for critical reading of an earlier draft of the manuscript.

REFERENCES

1. Ohman L, Simrén M. Pathogenesis of IBS: role of inflammation, immunity and neuroimmune interactions. Nat Rev Gastroenterol Hepatol 2010;7:163-73.
2. Field M. Intestinal ion transport and the pathophysiology of diarrhea. J Clin Invest 2003;111:931-43.
3. Franke A, McGovern DP, Barrett JC, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet 2010;42:1118-25.
4. Berni Canani R, Terrin G, Cardillo G, Tomaiuolo R, Castaldo G. Congenital diarrheal disorders: improved understanding of gene defects is leading to advances in intestinal physiology and clinical management. J Pediatr Gastroenterol Nutr 2010;50:360-6.
5. Saito YA. The role of genetics in IBS. Gastroenterol Clin North Am 2011;40:45-67.
6. Saito YA, Petersen GM, Larson JJ, et al. Familial aggregation of irritable bowel syndrome: a family case-control study. Am J Gastroenterol 2010;105:833-41.
7. Caprilli R. Why does Crohn's disease usually occur in terminal ileum? J Crohns Colitis 2008;2:352-6.
8. Murek M, Kopic S, Geibel J. Evidence for intestinal chloride secretion. Exp Physiol 2010;95:471-8.
9. Navaneethan U, Giannella RA. Mechanisms of infectious diarrhea. Nat Clin Pract Gastroenterol Hepatol 2008;5:637-47.
10. Basu N, Arshad N, Visweswariah SS.

Receptor guanylyl cyclase C (GC-C): regulation and signal transduction. Mol Cell Biochem 2010;334:67-80.
11. Saha S, Biswas KH, Kondapalli C, Isloor N, Visweswariah SS. The linker region in receptor guanylyl cyclases is a key regulatory module: mutational analysis of guanylyl cyclase C. J Biol Chem 2009;284:27135-45.
12. Thompson WG, Longstreth GF, Drossman DA, Heaton KW, Irvine EJ, Müller-Lissner SA. Functional bowel disorders and functional abdominal pain. Gut 1999;45:Suppl 2:II-43–II-47.
13. Gong R, Ding C, Hu J, et al. Role for the membrane receptor guanylyl cyclase-C in attention deficiency and hyperactive behavior. Science 2011;333:1642-6.
14. Valentino MA, Lin JE, Snook AE, et al. A uroguanylin-GUCY2C endocrine axis regulates feeding in mice. J Clin Invest 2011;121:3578-88.
15. Biswas KH, Shenoy AR, Dutta A, Visweswariah SS. The evolution of guanylyl cyclases as multidomain proteins: conserved features of kinase-cyclase domain fusions. J Mol Evol 2009;68:587-602.
16. Li C, Naren AP. CFTR chloride channel in the apical compartments: spatiotemporal coupling to its interacting partners. Integr Biol (Camb) 2010;2:161-77.
17. Cohen MB, Guarino A, Shukla R, Giannella RA. Age-related differences in receptors for Escherichia coli heat-stable enterotoxin in the small and large intestine of children. Gastroenterology 1988;94:367-73.
18. Steinbrecher KA, Cohen MB. Transmembrane guanylate cyclase in intestinal pathophysiology. Curr Opin Gastroenterol 2011;27:139-45.
19. Kuhn M. Function and dysfunction of mammalian membrane guanylyl cyclase receptors: lessons from genetic mouse models and implications for human diseases. Handb Exp Pharmacol 2009;191:47-69.
20. Steinbrecher KA, Harmel-Laws E, Garin-Laflam MP, et al. Murine guanylate cyclase C regulates colonic injury and inflammation. J Immunol 2011;186:7205-14.
21. Laubitz D, Larmonier CB, Bai A, et al. Colonic gene expression profile in NHE3-deficient mice: evidence for spontaneous distal colitis. Am J Physiol Gastrointest Liver Physiol 2008;295:G63-G77.
22. Wedenoja S, Höglund P, Holmberg C. The clinical management of congenital chloride diarrhoea. Aliment Pharmacol Ther 2010;31:477-85.
23. Fries W, Renda MC, Lo Presti MA, et al. Intestinal permeability and genetic determinants in patients, first-degree relatives, and controls in a high-incidence area of Crohn's disease in Southern Italy. Am J Gastroenterol 2005;100:2730-6.
24. Shulman RJ, Eakin MN, Czyzewski DI, Jarrett M, Ou CN. Increased gastrointestinal permeability and gut inflammation in children with functional abdominal pain and irritable bowel syndrome. J Pediatr 2008;153:646-50.
25. McGuckin MA, Eri R, Simms LA, Florin TH, Radford-Smith G. Intestinal barrier dysfunction in inflammatory bowel diseases. Inflamm Bowel Dis 2009;15:100-13.
26. Shen L, Su L, Turner JR. Mechanisms and functional implications of intestinal barrier defects. Dig Dis 2009;27:443-9.
27. Han X, Mann E, Gilbert S, et al. Loss of guanylyl cyclase C (GCC) signaling leads to dysfunctional intestinal barrier. PLoS One 2011;6(1):e16139.
28. Habre J, Sautot-Vial N, Marcotte C, Benchimol D. Caecal volvulus. Am J Surg 2008;196(5):e48-e49.
29. Ochiai T, Chijiiwa Y, Motomura Y, Iwakiri Y, Nawata H. Direct inhibitory effect of adrenomedullin and guanylin on isolated caecal circular smooth muscle cells of guinea pig. Life Sci 1997;61:1479-85.
30. Francis SH, Busch JL, Corbin JD, Sibley D. cGMP-dependent protein kinases and cGMP phosphodiesterases in nitric oxide and cGMP action. Pharmacol Rev 2010;62:525-63.
31. Lelievre V, Favrais G, Abad C, et al. Gastrointestinal dysfunction in mice with a targeted mutation in the gene encoding vasoactive intestinal polypeptide: a model for the study of intestinal ileus and Hirschsprung's disease. Peptides 2007;28:1688-99.
32. Mourad FH, Nassar CF. Effect of vasoactive intestinal polypeptide (VIP) antagonism on rat jejunal fluid and electrolyte secretion induced by cholera and Escherichia coli enterotoxins. Gut 2000;47:382-6.
33. Kurjak M, Fritsch R, Saur D, Schusdziarra V, Allescher HD. Functional coupling between nitric oxide synthesis and VIP release within enteric nerve terminals of the rat: involvement of protein kinase G and phosphodiesterase 5. J Physiol 2001;534:827-36.
34. Qian X, Moss NG, Fellner RC, Goy MF. Circulating prouroguanylin is processed to its active natriuretic form exclusively within the renal tubules. Endocrinology 2008;149:4499-509.
35. Lembo AJ, Schneier HA, Shiff SJ, et al. Two randomized trials of linaclotide for chronic constipation. N Engl J Med 2011;365:527-36.

Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Fiskerstrand T, Arshad N, Haukanes BI, et al. Familial diarrhea syndrome caused by an activating *GUCY2C* mutation. N Engl J Med 2012;366:1586-95. DOI: 10.1056/NEJMoa1110132.

Supplementary appendix.

Table of contents

Supplementary Material and Methods

Authors' contributions

Clinical and molecular genetic aspects of this study were designed by TF, PMK, JA and NH. Studies on the molecular analysis of the mutant receptor were designed by SSV, NA and KH. Data were gathered by NA, BIH, RRT, KDP, SLT, SEL, DB, HB, JA and TF. Data were analyzed by NA, BIH, RRT, SJ, BH, HB, JA, NH, SSV, PMK and TF. PMK, SSV, NH and SEL vouch for the data and the analyses. TF was responsible for writing the paper, including the first draft, in close collaboration with SSV and with contributions mainly from NA, PMK, RRT, SJ, BH, BIH, HB and NH. The decision to publish the paper was a joint one amongst all authors. The generated data belong to the authors. There are no other agreements concerning confidentiality of the data.

DNA analyses and mutation detection

Whole genome SNP genotyping data are deposited in the Gene expression omnibus database (http://www.ncbi.nlm.nih.gov/geo/), with accession number GSE31260. PCR primers for amplification of exons and flanking intronic sequences of *GUCY2C* were designed using the Oligo 6.3 program (National Bioscience, Plymouth, USA). Reference sequence for *GUCY2C* was NM_004963.3. PCR amplification was performed using standard procedures. DNA sequencing was performed using the ABI BigDye kit v1.1 and the ABI 3730 sequencer. Data analysis was assisted by use of the Seqscape software (ABI, Foster City, USA).

Exome sequencing

Samples were prepared for whole exome capture using Roche-Nimblegen's SeqCap EZ Exome v3 (64Mb capture region) and sequenced using the Illumina HiSeq to a median

coverage of 74x according to manufacturer's protocol. [1] Following sequencing using a paired-end 100 nt read length, sequences were analyzed with CASAVA v1.8 (Illumina Inc) followed by alignment with Burrows-Wheeler transformation. [2] PCR duplicates were removed with PICARD (http://picard.sourceforge.net) followed by base quality recalibration using Genome analysis tool kit. [3] SNPs and indels were called by SAMtools. [4] Variants with less than 8X sequencing depth and quality score ≤ 30 were excluded from analysis. Annovar [5] was used for variant annotation. Variant analysis was restricted to the chromosome 12 linkage interval at 14,466,726-17,410,570. Median coverage of the CCDS-bases in the linkage interval was 95X, with 99% of bases covered at ≥10X.

Linkage analysis

Multi-point parametric linkage analysis and haplotyping was performed using the program Allegro v2[6] on a set of 45000 SNPs pruned for strong local LD. Files were handled by the Uni Computing Linux cluster FIMM at the Bergen Center for Computational Science. We used a fully penetrant dominant inheritance mode with population disease allele frequency of 0.001. Marker allele frequencies were estimated using data from all individuals in the analysis. Physical marker positions are given according to NCBI Build 36.3 (http://www.ncbi.nlm.nih.gov/mapview/).

Characterization of mutant protein GC-CS840I

Ligand-stimulated GC-C activity was measured in cells 72 h post-transfection, following addition of varying concentrations of uroguanylin or heat-stable enterotoxin (ST) (for 15 min), or guanylin (for 60 min). ST was purified in the laboratory by procedures described earlier .[7] Cells were then lysed in 0.1N HCl and cGMP measured by radioimmunoassay as described earlier [8], or by ELISA (Enzo Life Sciences). Cells were also directly lysed in SDS

sample buffer and subjected to Western blot analysis using a monoclonal antibody to GC-C. [9] For *in vitro* guanylate cyclase assays, membranes were prepared from transfected cells. [9] Protein concentration was estimated by using a modification of the Bradford protein assay. [10]

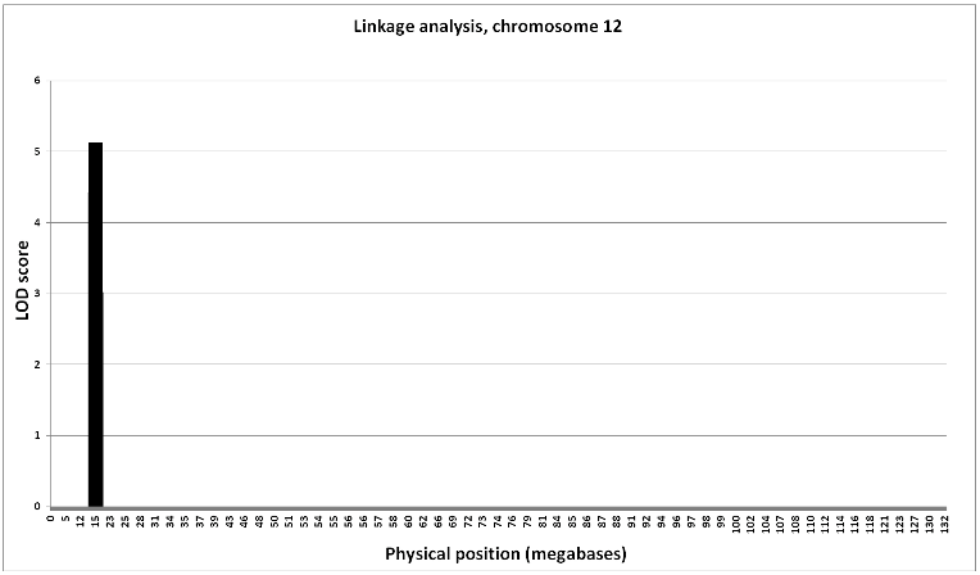
Heterodimerization and interaction of wild-type (WT) and mutant GC-C

To monitor heterodimerization, plasmids harboring WT GC-C fused at the C-terminus to GST (GC-C-GST)[8] and mutant GC-CS840I were mixed in varying ratios for a total of 400 ng of DNA/well and co-transfected. In instances where only WT or mutant DNA (1:0 or 0:1) were to be transfected, vector control DNA was co-transfected to maintain the ratio of total DNA to lipid. 72 h post transfection, membrane protein (400 ∝g) was solubilized in interaction buffer (10 mM Tris-HCl pH 7.5, 5 μg/ml each of aprotinin, leupeptin and soybean trypsin inhibitor, 1% NP-40 and 5 mM β-mercaptoethanol), and solubilized protein incubated with 10 μL GSH beads for 1 h at 4 °C. Following washing of the beads, protein bound to the beads was subjected to Western blot analysis using the GC-C monoclonal antibody.

Supplementary Figure 1

LODscore on chromosome 12

Whole genome linkage analysis using SNP-arrays identified only one significant region (4.13 centimorgan; 14,466,726-17,410,570 basepairs from end of 12p) on chromosome 12 with a max LODscore of 5.1. This means that the odds for linkage versus no linkage for this region with the disease is 125892/1. Physical position in Mb (chromosome 12) is given in the figure below, with start of the short (p) arm to the left.



Supplementary Figure 2

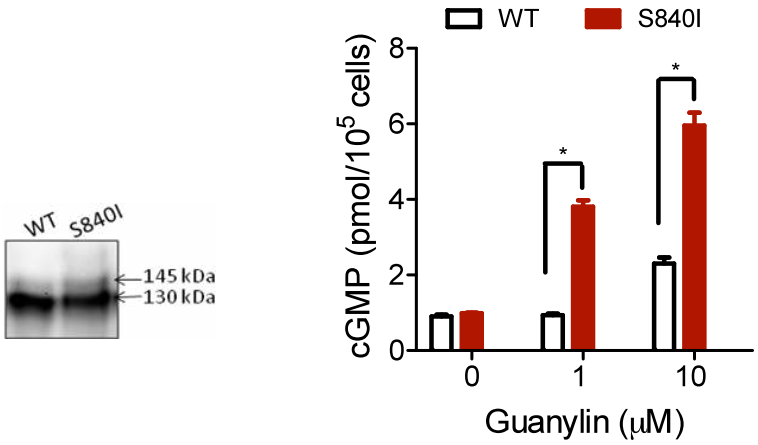
Multiple protein sequence alignment of GC-C

Homo sapiens	794	**DRLNFMLLPRLVVKSLKEKGFVEPELYEEVTIYFSDIVGFTTICKYSTPM**	843
Pan troglodytes	794	DRLNFMLLPRLVVKSLKEKGFVEPELYEEVTIYFSDIVGFTTICKYSTPM	843
Canis lupus	792	DRLNFMLLPRLVVKSLKEKGIVEPELYEEVTIYFSDIVGFTTICKYSTPM	841
Bos taurus	793	DQLNFMLLPRLVVKSLKEKGFVEPELYEEVTVYFSDIVGFTTICKYSTPM	842
Mus musculus	769	DHLNFMLLPRLVVKSLKEKGIVEPELYEEVTIYFSDIVGFTTICKYSTPM	818
Rattus norvegicus	793	DHLNFMLLPRLVVKSLKEKGIVEPELYEEVTIYFSDIVGFTTICKYSTPM	842
Gallus gallus	791	DRLNFMLLPRPVVKSLKETGLVEPEFFEEVTIYFSDIVGFTTLCKYSTPM	840
Danio rerio	574	DCLNFMLLPGPVVRSLKETGKVEPELFDEVTIYFSDIVGFTTLCHFSTPM	623

The region of the GC-C protein surrounding the amino acid serine (S, red colour) in position 840 in the protein is displayed, with sequences from different species aligned. The high degree of conserved amino acids in the whole region may indicate that this region is important for protein function. Note residue tyrosine (Y, green colour) at position 820, which has been shown to be a site for inhibitory phosphorylation by c-src. [11] The Ser840 (conserved in all compared species) is replaced by isoleucine in Familial Diarrhea Syndrome patients.

Supplementary Figure 3

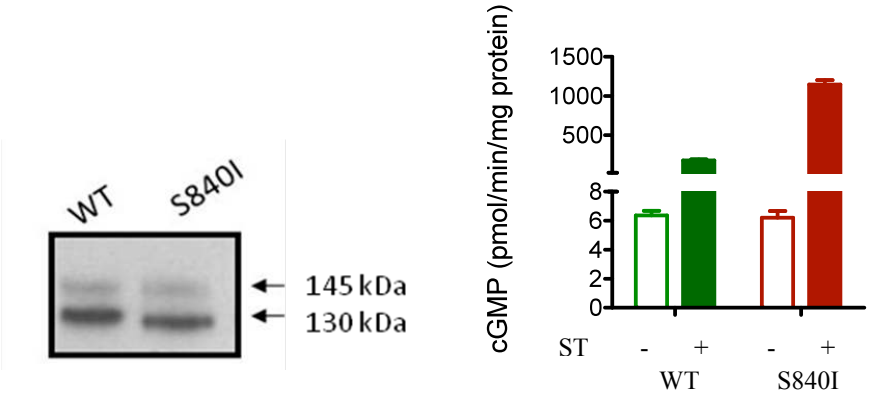
Guanylin-mediated activation of WT GC-C and mutant GC-CS840I in intact cells.



Guanylin-mediated cGMP production by GC-C. HEK293T cells were transfected with plasmids that allowed expression of either WT GC-C or mutant GC-CS840I. Equivalent expression of WT or mutant GC-C was observed on Western blot analysis using a monoclonal antibody to GC-C (left panel). Cells were either treated with medium alone, 1 ∝M or 10 ∝M guanylin. The experiment (n=2) was performed twice and data shown are the mean ± SEM of four determinations (*$p < 0.0001$).

Supplementary Figure 4

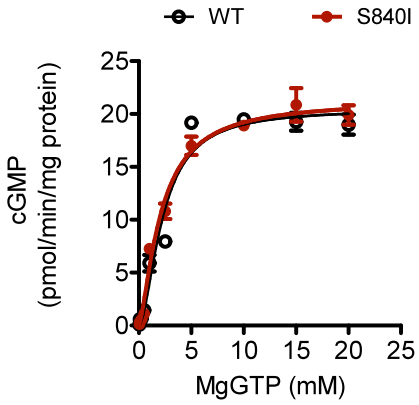
ST-mediated cGMP production in membranes prepared from transfected cells



Membrane fractions from cells transfected with plasmids encoding either WT GC-C or mutant GC-CS840I were prepared. Expression of GC-C was monitored by Western blot analysis (left panel). Guanylate cyclase assays were performed with membrane protein (50 ∝g) in the presence of 10 mM $MgCl_2$, 1 mM GTP, 500 ∝M isobutyl methyl xanthine, with or without the addition of ST (10^{-7} M). Assays were incubated for 10 min at 37 ^{Y}C, reactions stopped and cGMP produced measured by radioimmunoassay as detailed earlier. [8] Data shown are the mean ± SEM of duplicate determinations from a representative experiment, with experiments repeated thrice.

Supplementary Figure 5

Basic guanylate cyclase activity in membranes prepared from transfected cells

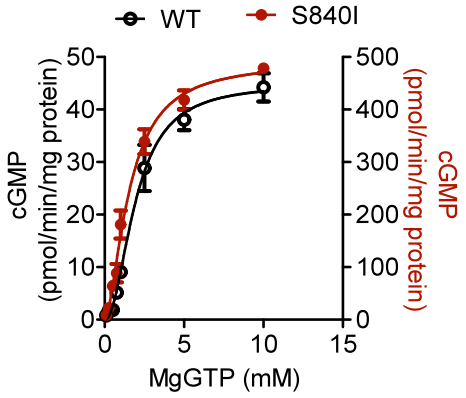


	V_{max} (mM)	K' (mM)	Hill
WT	19.3 ± 0.7	4.6 ± 0.1	1.8 ± 0.1
S840I	21.1 ± 0.9	5.2± 0.7	1.7+ 0.1

Membrane fractions from cells transfected with plasmids encoding either WT GC-C or mutant GC-CS840I were prepared. Membrane protein (50 ∝g) was used for guanylate cyclase assays performed in the presence of 10 mM $MgCl_2$, 1 mM GTP, 500 ∝M isobutyl methyl xanthine without the addition of ST, which would represent basal guanylate cyclase activity. Assays were incubated for 10 min at 37 ϒC, reactions stopped and cGMP produced measured by radioimmunoassay as detailed earlier.[8] Activity seen in untransfected cells was less than 10 % of that seen following transfection of GC-C (data not shown). Data shown in the graph are the mean ± SEM of duplicate determinations from a representative experiment. However, experiments were repeated twice and the table shows the mean ± SEM from kinetic analyses of these data.

Supplementary Figure 6

ST-mediated activation of WT GC-C and mutant GC-CS840I in cell membranes

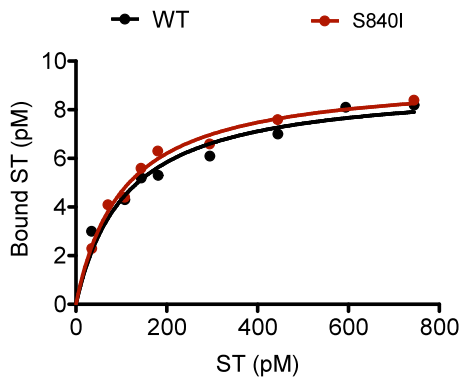


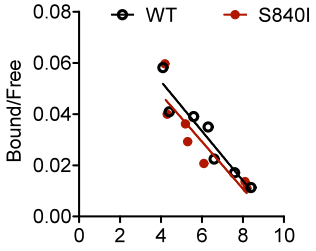
	V_{max} (mM)	K' (mM)	Hill
WT	50.7 ± 3.2	2.7 ± 0.7	1.8 ± 0.1
S840I	462.7 ± 19	1.5 ± 0.3	1.5 ± 0.1

Assays were performed with membrane protein (50 ∝g) prepared from transfected cells, in the presence of 10 mM $MgCl_2$, 1mM GTP, 500 ∝M isobutyl methyl xanthine and ST (10^{-7} M). Incubations were for 10 min at 37 ºC and cGMP produced was measured by radioimmunoassay as described earlier. [8] Data shown in the graph are the mean ± SEM of duplicate determinations from a representative experiment. Note different scales on the y-axes. Experiments were repeated twice and the table shows the mean ± SEM from kinetic analyses of these data.

Supplementary Figure 7

Saturation binding and Scatchard analyses of WT GC-C and mutant GC-CS840I.





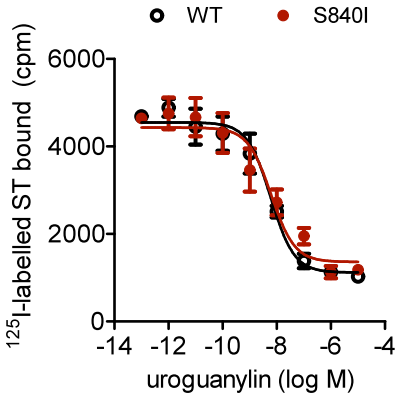
	WT	*S840I*
B_{max} (pM)	9.1	9.4
K_d (pM)	98.6 ± 10	103.7 ± 8

Membrane protein (50-100 ∝g) prepared from transfected cells was incubated with varying concentrations of ^{125}I-labeled ST peptide[8] for 1 h at 37^{o}C. Bound ST peptide was measured following filtration through GF/C filters (left panel). Data analysis was performed using GraphPad Prism 5. Representation of the data in a Scatchard Plot (right panel) allows the determination of the dissociation constant (K_d) of ST binding to GC-C, from the slopes of the lines. Data shown in the graph is from a representative experiment, as is the B_{max} value in the table. The experiment was repeated thrice to obtain the K_d with values shown representing the mean ± SEM. As seen, the binding affinity of ST to either WT or mutant GC-C was similar.

Supplementary Figure 8

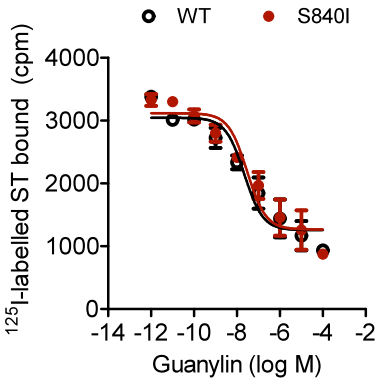
Affinity of binding of uroguanylin and guanylin to WT GC-C and mutant GC-CS840I

A



	WT	*S840I*
K_i (nM)	1.9 ± 0.4	2.0 ± 0.5

B



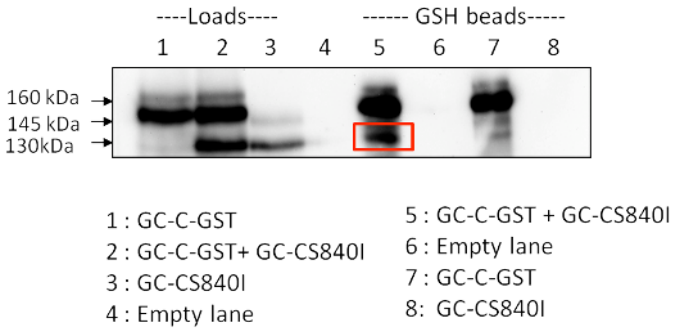
	WT	*S840I*
K_i (nM)	7.3 ± 1.6	9.9 ± 1.6

Membrane protein (50-100 ∝g) was incubated with ~ 100,000 cpm of ^{125}I-labeled ST peptide (~ 2000 Ci/mmol) in the presence of the indicated concentration of uroguanylin (A) or guanylin (B). Incubation was continued for 1 h at 37 °C following which samples were filtered through a GF/C filter and radioactivity associated with the membrane measured. Data was analyzed by GraphPad Prism 5, and the K_i (concentration of ligand required to reduce the binding of radiolabelled ligand (ST) by 50 %) was estimated from non-linear regression

analysis. The inverse of K_i is indicative of the affinity of the ligands for the receptor. Data shown is the mean ± SEM of experiments repeated three times, and the data indicate that there is no significant difference in the affinities of either uroguanylin or guanylin for the WT and mutant receptor

Supplementary Figure 9

Heterodimerization of WT and mutant GC-C



HEK293T cells were transfected with plasmids that express WT GC-C tagged with GST (GC-C-GST) or mutant GC-CS840I, either singly, or in a 1:1 ratio. 72 h following transfection, membranes were prepared from transfected cells and solubilised as described earlier.[8] Lysates were interacted with glutathione beads, and then subjected to Western blot analysis using a GC-C monoclonal antibody. Lanes 1-3 are samples prior to interaction with the beads. Lanes 5,7,8 indicate the various complexes pulled down by GSH beads. The red box outlines the band corresponding to GC-CS840I that is pulled down along with GC-C-GST, indicating heterodimerization of the wild type and mutant receptors. Note that no GC-CS840I interacts non-specifically with GSH beads (lane 8). Data is representative of experiments performed twice.

References

1. Bentley DR, Balasubramanian S, Swerdlow HP, et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature 2008;456:53-9.

2. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754-60.

3. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res;20:1297-303.

4. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078-9.

5. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res;38:e164.

6. Gudbjartsson DF, Thorvaldsson T, Kong A, Gunnarsson G, Ingolfsdottir A. Allegro version 2. Nat Genet 2005;37:1015-6.

7. Dwarakanath P, Visweswariah SS, Subrahmanyam YV, Shanthi G, Jagannatha HM, Balganesh TS. Cloning and hyperexpression of a gene encoding the heat-stable toxin of Escherichia coli. Gene 1989;81:219-26.

8. Saha S, Biswas KH, Kondapalli C, Isloor N, Visweswariah SS. The linker region in receptor guanylyl cyclases is a key regulatory module: mutational analysis of guanylyl cyclase C. J Biol Chem 2009;284:27135-45.

9. Ghanekar Y, Chandrashaker A, Tatu U, Visweswariah SS. Glycosylation of the receptor guanylate cyclase C: role in ligand binding and catalytic activity. Biochem J 2004;379:653-63.

10. Zor T, Selinger Z. Linearization of the Bradford protein assay increases its sensitivity: theoretical and experimental studies. Anal Biochem 1996;236:302-8.

11. Basu N, Bhandari R, Natarajan VT, Visweswariah SS. Cross talk between receptor guanylyl cyclase C and c-src tyrosine kinase regulates colon cancer cell cytostasis. Mol Cell Biol 2009;29:5277-89.

Supplementary Table 1. Clinical characteristics of affected family members.

All patients with exception of B-VI9 and B-VII1 reported diarrhea or loose stools from neonatal age or childhood. This was combined with meteorism, frequently with very gaseous stools, and some had recurrent abdominal pain. Individuals marked * had more pronounced diarrhea in childhood which improved in adult age. ND: No data.

Patient	Sex/ age	Stools Per day	Stool consistency	Meteorism	Recurrent abdominal pain	Hospitalisation due to electrolyte disturbances	Colonoscopy	Biopsies Colon/ ileum	Crohn's disease	Small bowel obstruction	Resection of intestine	Dilated intestine
A-IV1	F 89	4-6	Watery	ND	ND	Age 87	Normal	Normal	-	-	-	Age 87
A-IV4	F 82	5-6	Watery-explosive	ND	-	-	Normal	Normal	-	-	-	-
A-V1	F 59	2-3*	Normal Intermittent diarrhea	+	-	Several times related to gastroenteritis in adult life	Normal	Normal	-	-	-	-
A-V3	M 63	2-3	Semisolid - watery	+	-	-	Normal	Normal	-	-	-	-
A-V5	F 59	3-4	Semisolid - watery	+	-	-	Normal	Normal	-	-	-	-
A-V6	M 49	4-10	Watery	ND	ND	-	5 mm low grade adenoma, coecum	Normal	-	Suspected at age 48	-	Age 48
A-V8	M 60	1-2*	Normal	+	-	-	ND	ND	-	-	-	-
A-V10	M 67	2*	Semisolid	+	-	-	Diverticulum in terminal ileum	Normal	-	-	-	-
A-VI2	F 28	5-10	Watery	++	-	-	Ulceration at the ileocecal valve	Granulation polyp	-	-	-	-
A-VI6	M 28	1-3	Normal-semisolid Intermittent diarrhea	+	+	-	Eosinophilic polyps	Eosinophilic cells	-	-	-	-
A-VI7	M 45	1-4	Normal-semisolid Intermittent diarrhea	+	-	-	Normal	Normal	-	-	-	-
B-V2	F 62	3-4	Semisolid	+++	+	Age 58	Capsule endoscopy at age 60: Signs of inflammatation in ileum, increasing in distal direction	Normal colon biopsies but biopsies of ileum not done at age 60	Suspected at age 60	2x at age 20	25 cm of terminal ileum	Grossly dilated ileum (10 cm) at surgery age 20 Dilated ileum 4.4 cm at age 40
B-V3	F 65	3	Watery-semisolid	+++	+	ND	Normal	ND	-	Volvulus of caecum/colon ascendens age 27. Adhesional band obstruction 4 months later	Resection of terminal ileum	Distended ileum at surgeries (10cm)
B-V5	F 69	1-2*	Normal	+	Earlier not now	-	ND	ND	-	-	-	ND
B-V7	M 73	1-3	Normal to semisolid	+	+	-	ND	ND	-	ND	-	ND
B-VI1	F 31	2-3	Watery	+++	-	4 days old	Normal 1986	ND	-	-	-	Dilated ileum 4.5 cm at 4 years
B-VI3	F 40	4-12	Watery	+++	+++	2 weeks old	ND	ND	-	-	-	Distended abdomen and distended colon neonatal age
B-VI5	M 45	4-6	Watery	+	+	Neonatally	ND from colono-scopy at age 22	Transmural ileal inflammation, focal active colitis at age 31	Age 22	Age 31[a]	15 cm ileum and 5 cm coecum, then 170 cm of ileum	Gross dilatation at surgery

B-VI7	M 43	6-10	Watery	+++	+++	-	ND	Transmural ileal inflammation, no granulomas at age 38	Age 38	Age 38 and 39	Ileal resection both at age 38 and 39[b]	Now permanent gross dilatation (10 cm) in distal 1m of ileum
B-VI8	M 38	0-4	Watery-semisolid	+	+	Age 17 months, convulsions	Colitis and apthous lesions all segments, terminal ileitis age 37	Active colitis all segments, mild ileitis age 37	Age 37	-	-	Moderately dilated ileum on MRI
B-VI9	M 41	2-5	Diarrhea, watery-semisolid, started at age 15[c]	-	-	After gastroenteritis at age 35	Ileal inflammation first noticed at age 30, also at age 34 and 41	Ileal inflammation No granulomas	Age 30	Age 35 Stenotic terminal ileum	Resection of 50 cm of terminal ileum	Dilated ileum proximal to a narrow terminal ileum at age 30
B-VI10	M 44	1-2	Intermittent diarrhea	+	-	ND	ND	ND	ND	ND	ND	ND
B-VII1	M 2	1-2 *per week*	Neonatal diarrhea, *but now constipation*	ND	ND	3 weeks old	ND	Rectal biopsies normal	-	-	-	Dilated intestinal loops neonatally
B-VII2	M 3	2-3	Watery-semisolid	+	ND	6months 2 ½ years	Normal	Normal	-	Age 6 months: Suspected sigmoideum volvulus related to rotavirus infection	-	Dilated colon at surgery aged 6 months
B-VII3	M 17	5	Watery	+++	+++	1 month	ND	ND	-	Age 10: Distal ileum rotated and herniated through mesenterial slit	-	Dilated 30-40 cm of distal ileum at surgery
B-VII4	M 16	1-3	Normal-semisolid	+++	+++	Neonatally Several times later related to gastroenteritis	ND	ND	-	-	-	Dilated rectum, 9 cm
B-VII5	M 12	2-5	Watery-semisolid	+++	+	Neonatally	ND	ND	-	-	-	-
C-V1	F 65	10-20	Watery	++	++	-	ND	ND	Suspected at age 65, high calprotectin[d]	-	-	Dilated distal ileum (4 cm), air-fluid-levels on X-ray
C-VI2	F 44	4-5	Watery-semisolid	+++	+++	-	Ileal apths, sigmoideal inflammation aged 41	Sigmoidal inflammation	Suspected at age 41 At present untreated (no inflammation)	Suspected at age 32	-	Atonic dilated ileum at explorative laparotomy age 32
C-VI3	F 41	4-12	Watery-semisolid	+++	+++	-	ND	ND	-	Age 41[e]	-	Grossly dilated ileum at laparotomy
C-VII2	F 4	>3	Watery-semisolid	+	+	Neonatally	ND	ND	-	-	-	-
C-VII3	M 17	3-4	Watery, occasional formed stools	+++	+++	In school age	Normal	ND	-	-	-	-

[a] Termial ileitis with 13 cm long ulcer at age 31. First resection of 15 cm ileum and 5 cm caecum. Reoperation 7 days later due to new obstruction. 5 weeks later again bowel obstruction and 170 cm intestine was resected.

[b] Resection of 12 cm terminal ileum at age 38 and 30 cm of terminal ileum at age 39. At age 39 there was no stenosis, but an atonic segment of ileum.

[c] Diarrhea started at age 15 after a gastroenteritis. Diarrhea subsided in his twenties only to reappear permanently from age 28.

[d] Calprotectin measured recently was 243 mg/kg feces, which may indicate inflammation. Mucus and occasionally blood has been observed In feces. The patient also had hypomagnesemia and hypokalemia.

[e] Bowel obstruction in distal ileum described as caused by adhesions (first laparotomy). New signs of obstruction and new laparotomy 7 days later, gross dilatation of distal ileum (7 cm), but intestine was viable and resection was not performed. Patient experienced large fluid losses

Supplementary Table 2. Additional clinical characteristics of the affected family members.

Four patients have had urolithiasis, six have vitamin B12-deficiency, five have esophagitis ± esophageal hernia, six patients have hypertension, including two patients who developed this at young age. GI: Gastrointestinal.

Patient	Sex/ age	Factors worsening condition	Uro-lithiasis	B12 deficiency	Esophagitis± Esophageal hernia	Hypertension
A-IV1	F 89				Yes + hernia, Barret's esophagus	Yes
A-IV4	F 82			Yes	Yes + hernia	Yes
A-V1	F 59	GI-infections				Yes, aged 42
A-V3	M 63	Fruit and vegetables				
A-V5	F 59	Fruit				
A-V6	M 49	Fruit	Yes			Yes, aged 22
A-V8	M 60	Fruit, grapes				
A-V10	M 67	Red meat				
A-VI2	F 28	Pregnancy GI-infections			Yes	
A-VI6	M 28		Yes x 2			
A-VI7	M 45					
B-V2	F 62	Fat, sweets GI-infections		Yes		
B-V3	F 65	Sweets, dairy products				
B-V5	F 69					Yes
B-V7	M 73					
B-VI1	F 31	Sweets Carbonated soft-drinks GI-infections				
B-VI3	F 40	Sweets GI-infections				
B-VI5	M 45	Unknown				
B-VI7	M 43	Sensitive to certain types of food				
B-VI8	M 38	GI-infections				
B-VI9	M 41	Dark bread, sweets	Yes	Yes		

B-VI10	M 44	Certain types of food				
B-VII1	M 2					
B-VII2	M 3	GI-infections				
B-VII3	M 17	GI-infections Fruit juices				
B-VII4	M 16	Cow's milk Fruits GI-infections				
B-VII5	M 12	Milk Fruits GI-infections				
C-V1	F 65	GI-infections		Yes	Yes	Yes
C-VI2	F 44	GI-infections	Yes	Yes		
C-VI3	F 41	GI-infections		Yes	Yes + hernia	
C-VII2	F 4	GI-infections				
C-VII3	M 17					

Supplementary Table 3. Exome sequencing of the candidate region on chromosome 12.

DNA from individuals A-V1, B-VI3 and C-VI3 representing the three branches of the family, was whole exome sequenced to a median exome coverage of 74X and coverage of 95X in the candidate region on chromosome 12. All variants in the coding parts (exons) of the candidate region (14,466,726-17,410,570 bp from start of chromosome 12) in any of the three patients are listed. Positions are given in bp from start of short arm of chromosome 12. Only one common novel coding variant not present in the 1000 genomes or dbSNP was found; c.2519G>T in *GUCY2C* . In the candidate region, more than 99% of bases in the coding CDS were covered at ≥20X.

Position	Type of mutation	Gene	cDNA-change	Protein-change	dbSNP	A-V1	B-VI3	C-VI3
14469159	missense	*ATF7IP*	c.1043A>T	p.N348I	rs2231909	T/T	T/T	A/T
14478568	missense	*ATF7IP*	c.1589A>G	p.K530R	rs3213764			A/G
14548035	missense	*PLBD1*	c.1600C>G	p.P534A	rs1600	G/G	G/G	C/G
14555862	synonymous	*PLBD1*	c.895T>C		rs10846013	T/C		
14579911	missense	*PLBD1*	c.793G>A	p.V265I	rs7957558	A/A	A/A	A/A
14666289	missense	*GUCY2C*	c.2519G>T	p.S840I	NOVEL	G/T	G/T	G/T
14685329	synonymous	*GUCY2C*	c.2022C>T		rs10772800	T/T	T/T	T/T
14721160	missense	*GUCY2C*	c.843T>G	p.F281L	rs1420635	G/G	G/G	G/G
14727344	synonymous	*GUCY2C*	c.510C>T		rs56003832	C/T	C/T	C/T
14867820	synonymous	*C12orf60*	c.684G>C		rs17761825		G/C	
14884706	missense	*ART4*	c.793G>A	p.D265N	rs11276	G/A	A/A	G/A
14884875	synonymous	*ART4*	c.624C>T		rs3088189	C/T	T/T	C/T
14885121	synonymous	*ART4*	c.378T>C		rs1861698	T/C	C/C	T/C
14926348	missense	*MGP*	c.304A>G	p.T102A	rs4236	A/G	G/G	A/G
14986825	synonymous	*ARHGDIB*	c.504G>C		rs4703	G/C		
14994872	synonymous	*ARHGDIB*	c.42T>C		rs2430711	C/C	C/C	C/C
15153446	synonymous	*RERG*	c.408C>T		rs1055151	T/T	T/T	C/T
15561004	synonymous	*PTPRO*	c.1626T>C		rs1050646	T/C	T/C	
15570387	synonymous	*PTPRO*	c.2088C>A		rs6488782	C/A		
15624936	synonymous	*PTPRO*	c.519G>A		rs3748299			G/A
16238575	synonymous	*SLC15A5*	c.1563C>T		rs3942536	C/T		C/T
16261095	missense	*SLC15A5*	c.1482T>A	p.D494E	rs1671511	A/A	A/A	A/A
16288944	missense	*SLC15A5*	c.812C>T	p.P271L	rs1527014	T/T	T/T	T/T
16289001	missense	*SLC15A5*	c.755G>T	p.R252L	rs1852450	G/T		G/T
16321551	synonymous	*SLC15A5*	c.336C>T		rs1671487	T/T	T/T	T/T

Graphic design: Communication Division, UiB / Print: Skipnes Kommunikasjon AS

uib.no

ISBN: 9788230843970 (print)
9788230854426 (PDF)