**Frederik T M *et al.* A Danish cohort of monozygotic IBD twins**

Authors:

**Frederik Trier Moller, Lina Knudsen , Marcus Harbord , Jack Satsangi, Hannah Gordon, Lene Christiansen, Kaare Christensen, Tine Jess, Vibeke Andersen.**

**Frederik Trier Moller**, **Lina Knudsen, Vibeke Andersen** Research Unit for Molecular Diagnostics, Hospital of Southern Jutland, Aabenraa, Denmark

**Frederik Trier Moller**, **Lina Knudsen, , Lene Christiansen, Kaare Christensen, Vibeke Andersen** Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark

**Frederik Trier Moller**, **Tine Jess** Division of Epidemiology and Disease Surveillance, Statens Serum Institut,

**Hannah Gordon, Marcus Harbord** Department of Gastroenterology, Chelsea and Westminster Hospital, London

**Jack Satsangi,** Centre for Molecular Medicine, University of Edinburgh, Western General Hospital

**Lene Christiansen, Kaare Christensen** Danish Ageing Research Centre, Institute of Public Health, University of Southern Denmark, Odense, Denmark

**Lene Christiansen, Kaare Christensen** Danish Ageing Research Centre, Institute of Public Health, University of Southern Denmark, Odense, Denmark

**Tine Jess** Department of Clinical Medicine, University of Aalborg, Aalborg, Denmark

Author contributions:

**Specific author contributions:**

**Frederik Trier Moller,** Conception and design of the study, generation, analysis and interpretation of data, drafting the manuscript and approval of the final version of the manuscript.

**Vibeke Andersen.** VA: Conception and design of the study, interpretation of data. Critical revision and approval of the final version of the manuscript.

**Tine Jess,** Conception and design of the study, interpretation of data, drafting and critical revision of the manuscript and approval of the final version of the manuscript.

**Lina Knudsen Marcus Harbord , Jack Satsangi, Hannah Gordon, Lene Christiansen, Kaare Christensen,** Conception and design of the study, critical revision of the manuscript and approval of the final version of the manuscript.

**Supported by** the Lundbeck foundation, Region of Southern Denmark, University of Southern Denmark, Hospital of Southern Jutland.

**Institutional review board statement:**

The study was approved by the ethics committee of the region of southern Denmark (approval no: S20120176). Further, the study is included in the regional application to The Data Protection Agency (Institutional Southern Region of Denmark J.nr. 2008-58-0035). To ensure confidentiality direct paired comparisons between twin pairs are not shown.

**Informed consent statement:**

Verbal as well as written informed consent was obtained from participants. This included consent to contact co-twins of the index twins, even if that included informing the co-twin of the diagnosis of the index twin.

**Conflict-of-interest statement:**

**Vibeke Andersen** is an adviser for MSD/Merck, Jansen, and member of advisory board for

MSD/Merck.

**Tine Jess** has received funding for travel and speakers fee from AbbVie.

**Frederik Trier Moller, Lina Knudsen , Marcus Harbord , Jack Satsangi, Hannah Gordon, Lene Christiansen, Kaare Christensen:** no conflicts of interest

**Data Sharing statement:**

Technical appendix, code are available from the corresponding author at [frtm@ssi.dk]. Additional data are available on request but may require further IRB approval/ approvals from the data protection agency, to be shared outside the research group.

Correspondence to

MD, PhD Frederik Trier Moller

Division of Epidemiology and Disease Surveillance Statens Serum Institut, Artillerivej 5, 2300 Copenhagen, Denmark frtm@ssi.dk +4532688696

Abstract word count: 198 (out 256)

Word count = 2976, excluding title page, abstract, references, figures and tables.

# Abstract

## Aim

To describe the establishment of a Danish IBD twin cohort with focus on concordance of treatment and inflammatory markers.

## Method

We identified MZ twins, likely to be discordant or concordant for IBD, by merging information from the Danish Twin Register and the National Patient Register. The twins were asked to provide biological samples, questionnaires, and data access to patient files and public registries. Biological samples were collected via a mobile laboratory, which allowed for immediate centrifugation, fractionation, and storage of samples. The mean time from collection of samples to storage in the -80° C mobile freezer was less than one hour. The diagnoses where validated using the Copenhagen diagnostic criteria.

## Results

We identified 159 MZ IBD twin pairs, in a total of 62 (39%) pairs both twins agreed to participate. Of the supposed 62 IBD pairs, the IBD diagnosis could be confirmed in 54 pairs. The cohort included 10 concordant pairs, whereof some were discordant for either treatment or surgery. The 10 concordant pairs, where both pairs suffered from IBD, included eight CD/CD pairs, one UC/UC pair and one UC/IBDU pair. The discordant pairs comprised 31 UC, 5 IBDU (IBD unclassified), and 8 CD discordant pairs. In the co-twins not affected by IBD, calprotectin was above 100 µg/g in 2 participants, and above 50 µg/g in a further 5 participants.

## Conclusion

The presented IBD twin cohort are an excellent resource for bioinformatics studies with proper adjustment for disease-associated exposures including medication and inflammatory activity in the co-twins.

Key words (MeSH terms): [Digestive System Diseases](http://www.ncbi.nlm.nih.gov/mesh/68004066), Inflammatory Bowel Diseases, Epidemiologic studies, Twins, Biobank.

Core tip:

Using co-twin study designs to segregate genetic and environmental factors holds promise for future discovery when considering subclinical disease in the co-twins. However, as MZ IBD discordant twin are rare this often-mean insufficient power for planned analyses. Hence, collaboration between twin resources is crucial.

Introduction

The inflammatory bowel diseases (IBD), Crohn’s disease (CD) and ulcerative colitis (UC), affect a large number of Europeans [1, 2]. Despite the introduction of new treatments, CD and UC remain chronic conditions with severe disease morbidity, often complicated by surgery and frequent admissions to hospital [1, 3].

Although studies of the genome have found 200 loci associated with IBD, the variation in the IBD phenotype explained by these finding is still below 25-30% [4, 5] suggesting a role of environmental factors in IBD pathogenesis. Several studies indicate environmental impact on IBD pathogenesis including; exposure to pathogens [6], disease associated dysbiosis [7], metabolic disequilibrium [8], or epigenetic modifications [9]. More comprehensive studies, addressing these and other potential causes of IBD, could provide invaluable new insight into the pathogenesis of IBD [10], though studies using unrelated individuals would require large populations to overcome genetic variation between unrelated subjects [11].

Monozygotic (MZ) twins share common genotypes and epigenetic profiles at conception [12]. While some epigenetic differences arise during the lifetime of MZ twins [13], the inter-individual variation in relation to e.g. the epigenome and the gut microbiome remain lower between twin pairs than between unrelated persons [14]. Consequently, comprehensive studies of the exposome using IBD discordant MZ twin study designs could prove a powerful tool to assess the combined effects of environmental and endogenous factors, and identify targets for treatment and prevention [15].

A major challenge in such discordant twin pair studies is that quiescent or subclinical disease may blur the boundaries between cases and their co-twin controls [16, 17]. Another major challenge is that IBD discordant twin pairs are also treatment discordant, hence observed differences may derive from differential medication rather than disease discordance. Given enough power, studies using concordant twin pairs in addition to discordant twin pairs could allow researchers to adjust for disease-associated exposures such as medication, as both twins have IBD but may be discordant for some of the applied treatments. Further, calprotectin correlates with intestinal inflammation [18-20], and could reflect quiescent or subclinical disease in the unaffected co-twins.

We describe the establishment of a Danish IBD twin cohort including sampling of biological material, and illustrate the importance of treatment discordance and measurement of inflammatory markers for future bioinformatics studies using IBD affected twins.

# Methods

**Institutional review board statement:**

The study was approved by the ethics committee of the region of southern Denmark (approval no: S20120176). Further, the study is included in the regional application to The Data Protection Agency (Institutional Southern Region of Denmark J.nr. 2008-58-0035). To ensure confidentiality direct paired comparisons between twin pairs are not shown.

We identified MZ twins, likely to be discordant or concordant for IBD, by merging information from the Danish Twin Register and the National Patient Register [21, 22].

## Danish Twin Register

The Danish Twin Register enabled identification of MZ twins living in Denmark at the time of inclusion, with assessment of zygosity correct in 96% of cases [23]. The Register contains 72% of all twin pairs born between 1931–1968, with complete ascertainment of all live born twins since 1968 [21].

## National Patient Register

The National Patient Register is a nationwide register of all hospital discharge diagnoses, including surgical and other procedures recorded in Danish hospitals since 1977 [22]. The Register provides outpatient data from 1994 and surgical procedures since 1996. Diagnoses of UC and CD were identified using the international classification of diseases (ICD) 8th and 10th revision codes for CD (563.00–563.09 and K50) and UC (563.19, 569.04 and K51).

The diagnosis of IBD has previously been found to be accurate in over 90% of IBD cases in the national patient register, using a pathology register as reference [24].

## Cohort recruitment

Merging the Danish Twin Register and the National Patient Register, identified 159 MZ twin pairs in which at least one twin had a diagnosis of either CD or UC according to the National Patient Register as of May 1st 2013. Of these, 113 index twins (the first twin to contract IBD according to the register) responded to the invitational letter of whom 42 twins declined to participate. Of the 71 positive index twin responders, nine co-twins did not wish to participate, leaving 62 pairs for inclusion, Figure 1.

## Data collection

The participants filled out a questionnaire including age, sex, smoking status, medication, dietary patterns including a food frequency questionnaire, a 48-hour dietary recall, time of last meal or exercise, travel history, and pregnancies and disease activity at time of sampling, either Harvey Bradshaw Index (CD) (33) or Simple Clinical Colitis Index (UC) (34),

Data collected from the patient record included disease staging using the Montreal classification (32), any IBD complications, extra intestinal manifestations, and gastrointestinal operations as well as prior IBD medication.

The register diagnosis of CD, UC or IBDU, was validated by hospital records and pathology descriptions using the Copenhagen criteria [25]:

Copenhagen Diagnostic Criteria for CD (at least two of the criteria present) [26, 27]

1. History of abdominal pain, weight loss and/or diarrhoea for more than three months
2. Characteristic endoscopic findings of ulceration (aphthous lesions, snail track ulceration) or cobble stoning or radiological features of stricture or cobble stoning
3. Histopathology consistent with Crohn’s disease (epitheloid granuloma of Langerhans type or transmural discontinuous focal or patchy inflammation)
4. Fistula and/or abscess in relation to affected bowel segments.

Copenhagen Diagnostic Criteria for UC (all three of the criteria present) [26, 28]:

1. History of diarrhoea and/or rectal bleeding and pus for more than one week or repeated episodes
2. Characteristic endoscopic findings of continuous ulceration, vulnerability or granulated mucosa
3. Histopathology consistent with ulcerative colitis (neutrophils within epithelial structures, cryptitis, crypt distortion, crypt abscesses).

Inter-observer variation has previously been found with regards to the Montreal classification [29]. To avoid potential inter-observer variation one researcher validated the diagnoses and assessed the Montreal classification (FTM). Furthermore, to improve validity of diagnoses and phenotypes, complicated cases were reviewed by a gastroenterological specialist and senior physician (VAN). In daily clinical practice, the diagnosis may remain difficult; therefore, we included cases, which were perceived to have IBD by the treating physician although not fulfilling the Copenhagen criteria as IBD cases according to available information from the files. Cases where the diagnosis of IBD was unlikely were designated “Gastrointestinal (GI) symptoms not IBD” for future reference.

## Biological samples

Due to the geographical challenges in sampling a nationwide cohort, a mobile lab was set up using a camper previously fitted for a similar purpose. The camper was equipped with a small lab bench, heating, refrigeration, -20°C freezer, a mobile -80° C freezer, as well as a swinging bucket centrifuge.

The mobile lab setup allowed researchers to visit the twins in their home or another private location. The samples were collected adhering to the Sample PRE-analytical Code (SPREC) and Biospecimen Reporting for Improved Study Quality (BRISQ) guidelines, logging primary container, pre- and post-centrifugation conditions, centrifugation parameters and storage conditions, see supplementary materials Table s1 [30, 31]. Faecal specimens were sampled by participants up until 48 hours before the visit and stored in their own freezer at -20°C [32]. Samples were then transferred to a -80°C freezer at the visit, under which conditions faecal samples have been found to be stable in composition [33]. Oral samples were collected with a cytobrush (Cytotak™ Transwab® Labelled Tube MW148) from the dorsum of the tongue, suspended in a buffer medium, and immediately frozen at -80° C. Paraffin was used to collect sputum samples that were either suspended in RNA later or frozen directly at -80°C. One researcher conducted the collection of all samples. All samples were analysed using standard methods centrally to avoid sampling variation between different centers.

The mean time from collection of samples to storage in the -80° C mobile freezer was less than one hour, except for blood samples, which were 60 minutes and 15 seconds please see Supplementary materials Table s2. Records where kept to ensure identification of any deviations from protocol in future analysis.

## Statistics

The study included only basic descriptive statistics using R version 3.2.0. In order to ensure confidentiality, no grouping of the twins below five pairs was presented.

# Results

## Study cohort

Out of 62 MZ twin pairs, after scrutinizing patient records, register data, and questionnaires, we found the index case of eight pairs unlikely to have IBD. The 8 cases were afflicted by the following diagnoses: lymphocytic colitis, irritable bowel syndrome, *Clostridium* *Difficile* infection, ischemic bowel changes and abscesses without pathologic CD features and grouped as GI symptoms not IBD. At least one twin suffered from IBD in all remaining 54 pairs according to patient records before verification of diagnostic criteria. Forty-four were discordant for IBD, of whom 24 out of 31 UC pairs, four out of five IBDU pairs, and six out of eight CD pairs fulfilled the Copenhagen diagnostic criteria. Of the 10 concordant pairs, there were eight CD/CD pairs, one UC/UC pair, and one UC/IBDU pair, where all but one CD index twin fulfilled the Copenhagen criteria. Both verified and suspected cases were included in the cohort, to reflect clinical practice.

## Age at diagnosis

The mean age at diagnosis was lower in the CD concordant than CD discordant pairs (24.75 years versus 31.75 years). The timespan between the diagnosis of an index twin and the IBD co-twin was 6 years on average, ranging from 94 days to 14 years. The mean disease duration at sampling was 15 years on average, ranging from 295 days to 37 years.

## Clinical characteristics, complications, medication and smoking.

Table 1 shows clinical characteristics of the discordant twin pairs. Nine extra intestinal manifestations were present among co-twins, most often arthropathy. Table 2 shows complications, medication and smoking. Though numbers are small, 25% of CD index twins and 63% of CD co-twins received surgery after their IBD diagnosis. Conversely, 50% received azathioprine among the CD index twins versus 13% among the CD co-twins.

## Assessment of inflammatory activity in concordant and discordant twin pairs

Figure 3 shows inflammatory activity in discordant co-twin pairs as measured by calprotectin, at the time of sample collection. There was evidence of gut inflammation in the apparently non-affected co-twin, with faecal calprotectin > 100 µg/g in two individuals and > 50 µg/g in a further five (Figure 3). In two of the index twins whose IBD diagnosis could not be verified, faecal calprotectin was > 100 µg/g, see supplementary materials.

# Discussion

We have established a nationwide cohort of 62 affected or suspected IBD monozygotic twin pairs, which allow assessment of a range of disease- and treatment- associated and phenotypical traits amongst both discordant and concordant MZ IBD twins. Validation of the CD, UC, and IBDU diagnoses resulted in 8 pairs where the diagnosis was unlikely, and 11 pairs where the diagnosis was likely, but the clinical information was too sparse to validate this. Therefore, 43 twin pairs fulfilled the Copenhagen diagnostic criteria. The cohort included 10 concordant pairs, and several of these IBD concordant pairs were discordant for either treatment or surgery. The 44 IBD discordant pairs comprised 31 UC pairs, five IBDU pairs, and eight CD pairs. Inflammatory activity was above the normal range in 7 of the co-twins not affected by IBD, with calprotectin above 100 µg/g in two co-twin pairs and above 50 µg/g in a further five pairs.

The strength of the presented twin cohort lies in the wide range of data collected, from questionnaire data, patient file and public register data, to multiple biological samples. Our mobile laboratory enabled uniform collection of biological material with few deviations from existing guidelines regarding sample collection, storage, and handling. The average time from sampling to storage at -80° C was 1 h or less for all samples. Our uniform sample collection method using a mobile laboratory reduced aberrant data handling normally affecting nationwide multicentre studies, and allowed for a single-site analysis of disease activity data. A drawback of this approach is that more advanced laboratory handling, like cell separation and preserving viable cells, was not performed. Instead, the study used CPT tubes, a commercial cell preservative, and gradual freezing of cells using a “mister frosty”, which has previously been shown to preserve viable cells and cell integrity [34].

While the collection of biological material in this study is more uniform than previous twin studies [14, 35-40], we were unable to perform invasive manoeuvres such as endoscopy with our mobile setup. Though we expect to achieve access to some biopsy material taken from routine endoscopies, a large proportion of the healthy twins had not recently undergone endoscopy, thus limiting opportunities for comparison. Our assessment of clinical characteristics and IBD medication use aggregated data from patient files and questionnaire data. Consequently, treatment not documented by hospital-based physicians or recalled by patients may remain unaccounted for, but this potential bias should be similar between concordant and discordant pairs.

Our inclusion rate was lower than expected at 62 pairs out of the contacted 159, perhaps due to the extent of collected samples, and the need for including both twins. Indeed, some selection bias favouring the inclusion of concordant pairs over discordant pairs could not be ruled out, based on the proportion of concordant pairs invited to the proportion of concordant pairs accepting to participate.

The IBD twins were identified using nationwide registers, reducing bias often bestowed upon twin studies relying on advertising for recruitment. Given sufficient power, concordant pairs may play a crucial role in discerning the effects of disease-associated traits, such as medical therapy, from the effects of IBD e.g. on the methylome or metagenome. In addition, though we did not have the power to test this formally, the mean age at diagnosis seemed lower in the CD concordant pairs at 25 years versus 32 years among CD discordant pairs. Results from previous twin studies are conflicting on this point [41-43]. If indeed such a difference exists, one possible explanation might be that concordant pairs carry a larger genetic liability to disease, with a lower threshold for disease throughout life, increasing the risk of both twins contracting IBD, resulting in twin concordance. A previous Swedish twin study found the total allele frequency of *Nod/Card* mutations to be 4.4 times higher among concordant twin pairs compared to discordant twin pairs contributing to, but not explaining concordance [41].

Phenotypical characteristics differentiate CD and UC from each other and from other conditions such as IBS, microscopic colitis and infections. A combination of clinical evaluation, endoscopic, histological, radiological, and/or bio-chemical investigations provides the diagnostic foundation for CD and UC [44, 45]. The correct classification of twin discordance is paramount. Of note, classification is not only dependent on the correct diagnosis but also on time interval following the diagnosis of the index twin, with the risk of contracting IBD declining with time for the co-twin. Although the maximum time-span between concordant pairs in this cohort was below the mean disease duration of 15 years in the present cohort, final verification of twin discordance can only be assessed at the end of the lifespan of both twins. Methods do however exist that take time to event into account [46]. While future disease may not be a problem in a discordant twin study design, this is only true if the exposures causing this disease are not already present. Quiescent and subclinical disease may complicate the distinction between cases and their co-twin controls. Indeed, a newly published study indicates the presence of latent or emerging disease in family members of affected IBD cases [16]. In addition, family members of affected IBD cases have increased calprotectin levels as compared to the background population [17]. Though a normal calprotectin level does not exclude IBD, due to the dynamic nature of this condition, and a low level in well treated IBD patients, calprotectin levels correlate with intestinal inflammation [18-20], and could thus reflect an increased liability to IBD in familial members of IBD-affected cases [47, 48]. One [48] twin study published in the past 10 years has reported on increased levels of intestinal inflammatory activity biomarkers such as calprotectin among the unaffected co-twins, while the majority previous twin studies have not [14, 35-40, 49]. We found two co-twins with no history of IBD with calprotectin values exceeding 100 µg/g, and a further 5 with values above the normal range. This may be important, as increased calprotectin may reflect many of the exposures leading to disease may already be present in a co-twin, if subclinical disease is not already present in the co-twin. As a result, inter-individual differences with impact on disease pathogenesis within pairs may be harder to assess, suggesting that calprotectin levels should be considered in analysis.

Disease discordant IBD twins remain rare and precious to research [14, 35-40, 49]. Though providing a powerful model for research, this will often mean insufficient power for planned analyses. Hence, collaboration between twin resources is crucial. Collaboration with the Nixon Twin and Multiplex (TAM) UK IBD cohort, analysing epigenetic data within similar biological material, has already been established [50]. Thus, both the Danish and the British IBD twin cohorts will include a range of clinical, epidemiological and biological data enabling researchers to study a cross section of the IBD exposome (Figure 2).

## Conclusion

The present cohort demonstrates the importance of assessing inflammatory biomarkers reflecting subclinical inflammatory activity among otherwise healthy co-twins in discordant twin studies. The present cohort will be part of international collaborations, thereby increasing the power to detect disease- associated factors, and allow sufficient concordant twins to be included in studies to adjust for treatment effects. Consequently, analysis of a range of data from cohorts of monozygotic IBD pairs using bioinformatic methods such as metagenomics, metabolomics, proteomic and epigenetics could provide new insight into the role of the exposome in IBD pathogenesis.

# References

1 **Burisch J**, Jess T, Martinato M, Lakatos PL, -EpiCom E. The burden of inflammatory bowel disease in Europe. *J Crohns Colitis* 2013; **7**(4): 322-337 [PMID: 23395397 PMCID: 23395397 DOI: 10.1016/j.crohns.2013.01.010]

2 **Molodecky NA**, Soon IS, Rabi DM, Ghali WA, Ferris M, Chernoff G, Benchimol EI, Panaccione R, Ghosh S, Barkema HW, Kaplan GG. Increasing Incidence and Prevalence of the Inflammatory Bowel Diseases With Time, Based on Systematic Review. *Gastroenterology* 2012; **142**(1): 46-54.e42 [PMID: 22001864 DOI: 10.1053/j.gastro.2011.10.001]

3 **Baumgart DC**, Carding SR. Inflammatory bowel disease: cause and immunobiology. *The Lancet* 2007; **369**(9573): 1627-1640 [PMID: 17499605 DOI: [http://dx.doi.org/10.1016/S0140-6736(07)60750-8](http://dx.doi.org/10.1016/S0140-6736%2807%2960750-8)]

4 Gordon H, Moller FT, Andersen V, Harbord M. Heritability in Inflammatory Bowel Disease: From the First Twin Study to Genome-Wide Association Studies. *Inflammatory bowel diseases* 2015 [PMID: 25895112 DOI: <http://dx.doi.org/10.1097/MIB.0000000000000393>]

5 **Jostins L**, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, Lee JC, Schumm LP, Sharma Y, Anderson CA, Essers J, Mitrovic M, Ning K, Cleynen I, Theatre E, Spain SL, Raychaudhuri S, Goyette P, Wei Z, Abraham C, Achkar JP, Ahmad T, Amininejad L, Ananthakrishnan AN, Andersen V, Andrews JM, Baidoo L, Balschun T, Bampton PA, Bitton A, Boucher G, Brand S, Buning C, Cohain A, Cichon S, D'Amato M, de JD, Devaney KL, Dubinsky M, Edwards C, Ellinghaus D, Ferguson LR, Franchimont D, Fransen K, Gearry R, Georges M, Gieger C, Glas J, Haritunians T, Hart A, Hawkey C, Hedl M, Hu X, Karlsen TH, Kupcinskas L, Kugathasan S, Latiano A, Laukens D, Lawrance IC, Lees CW, Louis E, Mahy G, Mansfield J, Morgan AR, Mowat C, Newman W, Palmieri O, Ponsioen CY, Potocnik U, Prescott NJ, Regueiro M, Rotter JI, Russell RK, Sanderson JD, Sans M, Satsangi J, Schreiber S, Simms LA, Sventoraityte J, Targan SR, Taylor KD, Tremelling M, Verspaget HW, De VM, Wijmenga C, Wilson DC, Winkelmann J, Xavier RJ, Zeissig S, Zhang B, Zhang CK, Zhao H, Silverberg MS, Annese V, Hakonarson H, Brant SR, Radford-Smith G, Mathew CG, Rioux JD, Schadt EE, Daly MJ, Franke A, Parkes M, Vermeire S, Barrett JC, Cho JH. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012; **491**(7422): 119-124 [PMID: 23128233 PMCID: PMC3491803 DOI: <http://dx.doi.org/10.1038/nature11582>]

6 **Jess T**, Simonsen J, Nielsen NM, Jorgensen KT, Bager P, Ethelberg S, Frisch M. Enteric Salmonella or Campylobacter infections and the risk of inflammatory bowel disease. *Gut* 2011; **60**(3): 318-324 [PMID: 21193449 DOI: 10.1136/gut.2010.223396]

7 **Gevers D**, Kugathasan S, Denson Lee A, Vázquez-Baeza Y, Van Treuren W, Ren B, Schwager E, Knights D, Song Se J, Yassour M, Morgan Xochitl C, Kostic Aleksandar D, Luo C, González A, McDonald D, Haberman Y, Walters T, Baker S, Rosh J, Stephens M, Heyman M, Markowitz J, Baldassano R, Griffiths A, Sylvester F, Mack D, Kim S, Crandall W, Hyams J, Huttenhower C, Knight R, Xavier Ramnik J. The Treatment-Naive Microbiome in New-Onset Crohn s Disease. *Cell host & microbe* 2014; **15**(3): 382-392 [PMID: 24629344 DOI: 10.1016/j.chom.2014.02.005]

8 **Lin H-M**, Helsby NA, Rowan DD, Ferguson LR. Using metabolomic analysis to understand inflammatory bowel diseases. *Inflammatory Bowel Diseases* 2011; **17**(4): 1021-1029 [PMID: 20629098 DOI: 10.1002/ibd.21426]

9 **Nimmo ER**, Prendergast JG, Aldhous MC, Kennedy NA, Henderson P, Drummond HE, Ramsahoye BH, Wilson DC, Semple CA, Satsangi J. Genome-wide methylation profiling in Crohnʼs disease identifies altered epigenetic regulation of key host defense mechanisms including the Th17 pathway. *Inflammatory Bowel Diseases* 2012; **18**(5): 889-899 [PMID: 22021194 DOI: 10.1002/ibd.21912]

10 **Huang H**, Vangay P, McKinlay CE, Knights D. Multi-omics analysis of inflammatory bowel disease. *Immunology letters* 2014; **162**(2 Pt A): 62-68 [PMID: 25131220 DOI: 10.1016/j.imlet.2014.07.014]

11 **1000 Genomes Project Consortium AG**, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012; **491**(7422): 56-65 [PMID: 23128226 DOI: <http://dx.doi.org/10.1038/nature11632>]

12 **Bell JT**, Spector TD. A twin approach to unraveling epigenetics. *Trends in Genetics* 2011 [PMID: 21257220 DOI: 10.1016/j.tig.2010.12.005]

13 **Fraga MF**, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, Heine-Suner D, Cigudosa JC, Urioste M, Benitez J, Boix-Chornet M, Sanchez-Aguilera A, Ling C, Carlsson E, Poulsen P, Vaag A, Stephan Z, Spector TD, Wu YZ, Plass C, Esteller M. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A* 2005; **102**(30): 10604-10609 [PMID: 16009939 PMCID: 1174919 DOI: 10.1073/pnas.0500398102]

14 **Willing BP**, Dicksved J, Halfvarson J, Andersson AF, Lucio M, Zheng Z, Järnerot G, Tysk C, Jansson JK, Engstrand L. A pyrosequencing study in twins shows that gastrointestinal microbial profiles vary with inflammatory bowel disease phenotypes. *Gastroenterology* 2010; **139**(6): 1844-1854.e1841 [PMID: 20816835 DOI: 10.1053/j.gastro.2010.08.049]

15 **van Dongen J**, Slagboom PE, Draisma HHM, Martin NG, Boomsma DI. The continuing value of twin studies in the omics era. *Nat Rev Genet* 2012; **13**(9): 640-653 [PMID: 22847273 DOI: 10.1038/nrg3243]

16 **Biancone L**, Calabrese E, Petruzziello C, Capanna A, Zorzi F, Onali S, Condino G, Lolli E, Ciccacci C, Borgiani P, Pallone F. A family study of asymptomatic small bowel Crohn's disease. *Dig Liver Dis* 2014; **46**(3): 276-278 [PMID: 24360029 PMCID: 24360029 DOI: 10.1016/j.dld.2013.11.003]

17 **Thjodleifsson B**, Sigthorsson G, Cariglia N, Reynisdottir I, Gudbjartsson DF, Kristjansson K, Meddings JB, Gudnason V, Wandall JH, Andersen LP, Sherwood R, Kjeld M, Oddsson E, Gudjonsson H, Bjarnason I. Subclinical intestinal inflammation: an inherited abnormality in Crohn’s disease relatives? *Gastroenterology* 2003; **124**(7): 1728-1737 [PMID: 12806605 DOI: [http://dx.doi.org/10.1016/S0016-5085(03)00383-4](http://dx.doi.org/10.1016/S0016-5085%2803%2900383-4)]

18 **van Rheenen PF**, Van de Vijver E, Fidler V. Faecal calprotectin for screening of patients with suspected inflammatory bowel disease: diagnostic meta-analysis. *BMJ* 2010; **341**: c3369 [PMID: 20634346 PMCID: PMC2904879 DOI: 10.1136/bmj.c3369]

19 **Limburg PJ**, Ahlquist DA, Sandborn WJ, Mahoney DW, Devens ME, Harrington JJ, Zinsmeister AR. Fecal calprotectin levels predict colorectal inflammation among patients with chronic diarrhea referred for colonoscopy. *Am J Gastroenterol* 2000; **95**(10): 2831-2837 [PMID: 11051356 DOI: 10.1111/j.1572-0241.2000.03194.x]

20 **Summerton CB**, Longlands MG, Wiener K, Shreeve DR. Faecal calprotectin: a marker of inflammation throughout the intestinal tract. *Eur J Gastroenterol Hepatol* 2002; **14**(8): 841-845 [PMID: 12172403 DOI: <http://dx.doi.org/10.1097/00042737-200208000-00005>]

21 **Skytthe A**, Kyvik KO, Holm NV, Christensen K. The Danish Twin Registry. *Scand J Public Health* 2011; **39**(7 Suppl): 75-78 [PMID: 21775358 DOI: <http://dx.doi.org/10.1177/1403494810387966>]

22 **Lynge E**, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scand J Public Health* 2011; **39**(7 Suppl): 30-33 [PMID: 21775347 DOI: 10.1177/1403494811401482]

23 **Christiansen L**, Frederiksen H, Schousboe K, Skytthe A, von Wurmb-Schwark N, Christensen K, Kyvik K. Age- and Sex-differences in the Validity of Questionnaire-based Zygosity in Twins. *Twin Research and Human Genetics* 2003; **6**(04): 275-278 [PMID: 14511432 DOI: doi:10.1375/twin.6.4.275]

24 **Fonager K**, Sorensen HT, Rasmussen SN, Moller-Petersen J, Vyberg M. Assessment of the diagnoses of Crohn's disease and ulcerative colitis in a Danish hospital information system. *ScandJGastroenterol* 1996; **31**(2): 154-159 [PMID: 8658038 DOI: <http://dx.doi.org/10.3109/00365529609031980>]

25 **Burisch J**, Cukovic-Cavka S, Kaimakliotis I, Shonová O, Andersen V, Dahlerup JF, Elkjaer M, Langholz E, Pedersen N, Salupere R, Kolho K-L, Manninen P, Lakatos PL, Shuhaibar M, Odes S, Martinato M, Mihu I, Magro F, Belousova E, Fernandez A, Almer S, Halfvarson J, Hart A, Munkholm P. Construction and validation of a web-based epidemiological database for inflammatory bowel diseases in Europe An EpiCom study. *J Crohns Colitis* 2011; **5**(4): 342-349 [PMID: 21683305 DOI: 10.1016/j.crohns.2011.02.016]

26 **Binder V**, Both H, Hansen PK, Hendriksen C, Kreiner S, Torp-Pedersen K. Incidence and prevalence of ulcerative colitis and Crohn's disease in the County of Copenhagen, 1962 to 1978. *Gastroenterology* 1982; **83**(3): 563-568 [PMID: 6980161]

27 **Munkholm P**. Crohn's disease--occurrence, course and prognosis. An epidemiologic cohort-study. *Dan Med Bull* 1997; **44**(3): 287-302 [PMID: 9233548]

28 **Langholz E**. Ulcerative colitis. An epidemiological study based on a regional inception cohort, with special reference to disease course and prognosis. *Dan Med Bull* 1999; **46**(5): 400-415 [PMID: 10605619]

29 **Krishnaprasad K**, Andrews JM, Lawrance IC, Florin T, Gearry RB, Leong RWL, Mahy G, Bampton P, Prosser R, Leach P, Chitti L, Cock C, Grafton R, Croft AR, Cooke S, Doecke JD, Radford-Smith GL. Inter-observer agreement for Crohn's disease sub-phenotypes using the Montreal Classification: How good are we? A multi-centre Australasian study. *Journal of Crohn's and Colitis* 2012; **6**(3): 287-293 [PMID: 22405164 DOI: <http://dx.doi.org/10.1016/j.crohns.2011.08.016>]

30 **Moore HM**, Kelly A, Jewell SD, McShane LM, Clark DP, Greenspan R, Hainaut P, Hayes DF, Kim P, Mansfield E, Potapova O, Riegman P, Rubinstein Y, Seijo E, Somiari S, Watson P, Weier HU, Zhu C, Vaught J. Biospecimen Reporting for Improved Study Quality. *BiopreservBiobank* 2011; **9**(1): 57-70 [PMID: 21826252 DOI: <http://dx.doi.org/10.1089/bio.2010.0036>]

31 **Betsou F**, Lehmann S, Ashton G, Barnes M, Benson EE, Coppola D, DeSouza Y, Eliason J, Glazer B, Guadagni F, Harding K, Horsfall DJ, Kleeberger C, Nanni U, Prasad A, Shea K, Skubitz A, Somiari S, Gunter E, International Society for B, Environmental Repositories Working Group on Biospecimen S. Standard Preanalytical Coding for Biospecimens: Defining the Sample PREanalytical Code. *Cancer Epidemiology Biomarkers & Prevention* 2010; **19**(4): 1004-1011 [PMID: 20332280 DOI: 10.1158/1055-9965.EPI-09-1268]

32 **Qin J**, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto J-M, Hansen T, Le Paslier D, Linneberg A, Nielsen HB, Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Li S, Jian M, Zhou Y, Li Y, Zhang X, Li S, Qin N, Yang H, Wang J, Brunak S, Doré J, Guarner F, Kristiansen K, Pedersen O, Parkhill J, Weissenbach J, Antolin M, Artiguenave F, Blottiere H, Borruel N, Bruls T, Casellas F, Chervaux C, Cultrone A, Delorme C, Denariaz G, Dervyn R, Forte M, Friss C, van de Guchte M, Guedon E, Haimet F, Jamet A, Juste C, Kaci G, Kleerebezem M, Knol J, Kristensen M, Layec S, Le Roux K, Leclerc M, Maguin E, Melo Minardi R, Oozeer R, Rescigno M, Sanchez N, Tims S, Torrejon T, Varela E, de Vos W, Winogradsky Y, Zoetendal E, Bork P, Ehrlich SD, Wang J. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 2010; **464**(7285): 59-65 [PMID: 20203603 DOI: 10.1038/nature08821]

33 **Wu GD**, Lewis JD, Hoffmann C, Chen YY, Knight R, Bittinger K, Hwang J, Chen J, Berkowsky R, Nessel L, Li H, Bushman FD. Sampling and pyrosequencing methods for characterizing bacterial communities in the human gut using 16S sequence tags. *BMC Microbiol* 2010; **10**: 206 [PMID: 20673359 PMCID: 2921404 DOI: 10.1186/1471-2180-10-206]

34 **Clarke DM**, Yadock DJ, Nicoud IB, Mathew AJ, Heimfeld S. Improved post-thaw recovery of peripheral blood stem/progenitor cells using a novel intracellular-like cryopreservation solution. *Cytotherapy* 2009; **11**(4): 472-479 [PMID: 19499402 PMCID: PMC3021966 DOI: 10.1080/14653240902887242]

35 **Petersen B-S**, Spehlmann ME, Raedler A, Stade B, Thomsen I, Rabionet R, Rosenstiel P, Schreiber S, Franke A. Whole genome and exome sequencing of monozygotic twins discordant for Crohn's disease. *BMC Genomics* 2014; **15**: 564 [PMID: 24996980 PMCID: PMC4102722 DOI: 10.1186/1471-2164-15-564]

36 **Spehlmann ME**, Begun AZ, Saroglou E, Hinrichs F, Tiemann U, Raedler A, Schreiber S. Risk factors in German twins with inflammatory bowel disease: results of a questionnaire-based survey. *J Crohns Colitis* 2012; **6**(1): 29-42 [PMID: 22261525 DOI: 10.1016/j.crohns.2011.06.007]

37 **Lepage P**, Häsler R, Spehlmann ME, Rehman A, Zvirbliene A, Begun A, Ott S, Kupcinskas L, Doré J, Raedler A, Schreiber S. Twin study indicates loss of interaction between microbiota and mucosa of patients with ulcerative colitis. *Gastroenterology* 2011; **141**(1): 227-236 [PMID: 21621540 DOI: 10.1053/j.gastro.2011.04.011]

38 **Halfvarson J**. Genetics in twins with Crohn's disease: less pronounced than previously believed? *Inflammatory bowel diseases* 2011; **17**(1): 6-12 [PMID: 20848478 DOI: 10.1002/ibd.21295]

39 **Bengtson M-B**, Aamodt G, Vatn MH, Harris JR. Concordance for IBD among twins compared to ordinary siblings--a Norwegian population-based study. *J Crohns Colitis* 2010; **4**(3): 312-318 [PMID: 21122520 DOI: 10.1016/j.crohns.2009.12.008]

40 **Willing B**, Halfvarson J, Dicksved J, Rosenquist M, Jarnerot G, Engstrand L, Tysk C, Jansson JK. Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease. *InflammBowelDis* 2009; **15**(5): 653-660 [PMID: 19023901 DOI: 10.1002/ibd.20783]

41 **Halfvarson J**, Bresso F, D'Amato M, Järnerot G, Pettersson S, Tysk C. CARD15/NOD2 polymorphisms do not explain concordance of Crohn's disease in Swedish monozygotic twins. *Dig Liver Dis* 2005; **37**(10): 768-772 [PMID: 16002353 DOI: 10.1016/j.dld.2005.05.005]

42 **Spehlmann ME**, Begun AZ, Burghardt J, Lepage P, Raedler A, Schreiber S. Epidemiology of inflammatory bowel disease in a German twin cohort: results of a nationwide study. *Inflamm Bowel Dis* 2008; **14**(7): 968-976 [PMID: 18253950 DOI: 10.1002/ibd.20380]

43 **Tysk C**, Lindberg E, Järnerot G, Floderus-Myrhed B. 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**(7): 990-996 [PMID: 3396969 DOI: <http://dx.doi.org/10.1136/gut.29.7.990>]

44 **Dignass A**, Eliakim R, Magro F, Maaser C, Chowers Y, Geboes K, Mantzaris G, Reinisch W, Colombel J-F, Vermeire S, Travis S, Lindsay JO, Van Assche G. Second European evidence-based consensus on the diagnosis and management of ulcerative colitis Part 1: Definitions and diagnosis. *Journal of Crohn's and Colitis* 2012; **6**(10): 965-990 [PMID: 23040452 DOI: <http://dx.doi.org/10.1016/j.crohns.2012.09.003>]

45 **Van Assche G**, Dignass A, Panes J, Beaugerie L, Karagiannis J, Allez M, Ochsenkühn T, Orchard T, Rogler G, Louis E, Kupcinskas L, Mantzaris G, Travis S, Stange E. The second European evidence-based Consensus on the diagnosis and management of Crohn's disease: Definitions and diagnosis, 2010

46 **Scheike TH**, Holst KK, Hjelmborg JB. Estimating twin concordance for bivariate competing risks twin data. *Stat Med* 2014; **33**(7): 1193-1204 [PMID: 24132877 PMCID: 24132877 DOI: 10.1002/sim.6016]

47 **Moller FT**, Andersen V, Wohlfahrt J, Jess T. Familial Risk of Inflammatory Bowel Disease: A Population-Based Cohort Study 1977-2011. *Am J Gastroenterol* 2015; **110**(4): 564-571 [PMID: 25803400 DOI: 10.1038/ajg.2015.50]

48 **Zhulina Y**, Hahn-Strömberg V, Shamikh A, Peterson CGB, Gustavsson A, Nyhlin N, Wickbom A, Bohr J, Bodin L, Tysk C, Carlson M, Halfvarson J. Subclinical inflammation with increased neutrophil activity in healthy twin siblings reflect environmental influence in the pathogenesis of inflammatory bowel disease. *Inflamm Bowel Dis* 2013; **19**(8): 1725-1731 [PMID: 23669399 DOI: 10.1097/MIB.0b013e318281f2d3]

49 **Jess T**, Riis L, Jespersgaard C, Hougs L, Andersen PS, Orholm MK, Binder V, Munkholm P. Disease concordance, zygosity, and NOD2/CARD15 status: follow-up of a population-based cohort of Danish twins with inflammatory bowel disease. *The American journal of gastroenterology* 2005; **100**(11): 2486-2492 [PMID: 16279904 DOI: 10.1111/j.1572-0241.2005.00224.x]

50 <http://www.ibdtam.org.uk/>

# Figure legends

## Figure 1 Legend

The collection of twin pairs.

## Figure 2 legend

The figure shows how the collected twin data may be used in different downstream analysis. The figure illustrates the initial genetic concordance in liability, the progressive discordance for disease liability due to heterogeneous exposures, and the possible future concordance for IBD, but not necessarily for treatment.

## Figure 3 Legend

The figure shows fecal calprotectin measures stratified by phenotype.

# Tables

## Table 1 Clinical characteristics.

|  |  |  |
| --- | --- | --- |
| **Pair type** | **Discordant twin pairs** | **Concordant twin pairs** |
| **Status** | **Co-twin** | **CD** | **IBDU** | **UC** | **non-IBD GI symptoms** | **IBD co-twin** | **IBD index twin** |
| **N** | 52 | 8 | 5 | 31 | 8 | 10 | 10 |
| **Males/ Females** | 23/29 | 2/6 | 1/4 | 16/15 | 4/4 | 4/6 | 4/6 |
| **Age** | 50(26-78) | 47 (26-67) | 57 (34-77) | 49 (32-70) | 55 (27-78) | 49 (28-68) | 49 (28-68) |
| Age at onset  |  | 32 (21-47) | 43 (23-73) | 35 (20-59) | 38 (23-62) | 21 (14-29) | 23 (11-34) |
| Age at diagnosis |  | 34 (25-46) | 41 (29-72) | 34 (17-66) | 48 (18-72) | 24 (11-37) | 31 (21-47) |
| **CPH criteria fulfilled** |  | 6 (75%) | 4 (80%) | 24 (77%) | 0 (0%) | 8 (80%) | 9 (90%) |
| **Disease location** |  |  |  |  |  |  |  |
|  L1 ileal | 2 (25%) | 2 (40%) |  |  | 3 (30%) | 2 (20%) |
|  L2 colonic | 3 (38%) | 0 (0%) |  |  | 1 (10%) | 1 (10%) |
|  L3 ileocolonic | 0 (0%) | 0 (0%) |  |  | 4 (40%) | 4 (40%) |
|  L4 isolated upper disease | 0 (0%) | 0 (0%) |  |  | 0 (0%) | 1 (10%) |
|  B1 non stricturing non penetrating | 4 (50%) | 2 (40%) |  |  | 3 (30%) | 1 (10%) |
|  B2 stricturing | 2 (25%) | 0 (0%) |  |  | 4 (40%) | 5 (50%) |
|  B3 penetrating | 0 (0%) | 0 (0%) |  |  | 3 (30%) | 2 (20%) |
|  P perianal disease | 1 (13%) | 0 (0%) |  |  | 2 (20%) | 0 (0%) |
|  Proctitis |  | 1 20% | 6 19% |  |  |  |
|  Left sided |  | 0 0% | 9 29% |  |  |  |
|  Extensive |  | 1 20% | 10 32% |  |  |  |

## Table 2 complications, medication and smoking.

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **Pair type** | **Discordant twin pairs** | **Concordant twin pairs** |
| **Status** | **Co-twin** | **CD** | **IBDU** | **UC** | **non-IBD GI symptoms** | **IBD co-twin** | **IBD index twin** |
| **N** | 52 | 8 | 5 | 31 | 8 | 10 | 10 |
| **Complications** |  |  |  |  |  |  |  |
| GI complications\* | 3 (6%) | 2 (25%) | 1 (20%) | 2 (6%) | 0 (0%) | 4 (40%) | 5 (50%) |
| Extra intestinal manifestations**∇** | 5 (10%) | 3 (38%) | 2(40%) | 12 (39%) | 2 (25%) | 2 (20%) | 6 (60%) |
| Ever surgery | 0 (0%) | 2 (25%) | 1 (20%) | 12 (39%) | 2 (25%) | 7 (70%) | 5 (50%) |
| colectomy | 0 (0%) | 1 (13%) | 1 (20%) | 7 (23%) | 0 (0%) | 0 (0%) | 0 (0%) |
| **Medication** |  |  |  |  |  |  |  |
| Ever TNF-Inhibitor | 0 (0%) | 2 (25%) | 1 (20%) | 4 (13%) | 0 (0%) | 1 (10%) | 1 (10%) |
| Ever glucocorticoids | 0 (0%) | 3 (38%) | 4 (80%) | 17 (55%) | 2 (25%) | 6 (60%) | 6 (60%) |
| Ever other imunosupressor¤  | 0 (0%) | 4 (50%) | 2 (40%) | 6 (19%) | 0 (0%) | 2 (20%) | 5 (50%) |
| Ever 5-ASA | 1 (2%) | 5 (63%) | 4 (80%) | 22 (71%) | 1 (13%) | 5 (50%) | 6 (60%) |
| **Smoking** |  |  |  |  |  |  |  |
| Never smoker | 28 (54%) | 6 (75%) | 2 (40%) | 21 (68%) | 5 (63%) | 3 (30%) | 2 (20%) |
| Current smoker | 10 (19%) | 1 (13%) | 1 (20%) | 2 (6%) | 2 (25%) | 3 (30%) | 7 (70%) |
| Former smoker | 13 (25%) | 1 (13%) | 2 (40%) | 7 (23%) | 1 (13%) | 4 (40%) | 1 (10%) |

**¤**methotrexate, azathioprine, and cyclosporine

**\***fistula, adherences, strictures, toxic megacolon, abscess, perforation, colorectal\_cancer.

**∇** hepatitis, primary\_sclerosing\_cholangitis , autoimmune\_pancreatitis ,uveitis ,

erythema\_nodosum, pyoderma\_gangrenosum, arthritis , aphthous\_ulcers

## Figure 1



## Figure 2.



## Figure 3

