Supplementary Appendix

The Interplay of Host Genetics and the Gut Microbiota underlying the Onset and Clinical Presentation of Inflammatory Bowel Disease

Floris Imhann[1,2]*, Arnau Vich Vila[1,2]*, Marc Jan Bonder[2], Jingyuan Fu[3], Dirk Gevers[4], Marijn C. Visschedijk[1,2], Lieke M. Spekhorst[1,2], Rudi Alberts[1,2], Lude Franke[2], Hendrik M. van Dullemen[1], Rinze W.F. Ter Steege[1], Curtis Huttenhower[4,6], Gerard Dijkstra[1], Ramnik J. Xavier[4,5], Eleonora A.M. Festen[1,2], Cisca Wijmenga[2], Alexandra Zhernakova[2#], Rinse K. Weersma[1#]

*Shared first authors

#Shared last authors

[1] University of Groningen, University Medical Center Groningen, Department of Gastroenterology and Hepatology, Groningen, the Netherlands

[2] University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, the Netherlands

[3] University of Groningen, University Medical Center Groningen, Department of Pediatrics, Groningen, the Netherlands

[4] Broad Institute of Harvard and MIT, Boston, USA

[5] Massachusetts General Hospital, Boston, USA

[6] Biostatistics Department, Harvard School of Public Health, Boston, USA

Index

1. Supplementary Methods: Phenotypes

All phenotypes in our dataset are presented below:

- Age (also controls)
- Age at diagnosis
- Ethnicity
- Ethnicity Caucasian yes/no
- Gender (also controls)
- Mode of Birth (vaginal or caesarean section)
- Breastfed (yes/no)
- Body Mass Index (also controls), Length, Weight
- First Inflammatory Bowel Disease (IBD) Diagnosis
- Current IBD Diagnosis (Crohn's Disease (CD)/Ulcerative Colitis (UC))
- IBD Disease location: ileum, colon, both
- Combined Diagnosis and Location scores because of overlap location and diagnosis: ileal CD, ileocolonic CD, colonic CD, UC
- IBD Disease activity (Clinical scores)
    - Harvey Bradshaw Index
    - Simple Clinical Colitis Activity Index
- Montreal Classification
    - Montreal A (age at onset)
    - Montreal L1, 2, 3 (location)
    - Montreal L4 yes/no (location)
    - Montreal B1,2,3 (CD patients only) (behaviour)
    - Montreal Bp (peri-anal) (CD patients only) (yes/no)
    - Montreal E1,2,3 (UC patients only) (extent)
    - Montreal S0,1,2,3 (UC patients only) (severity)
- IBD Disease duration (time sampling – time diagnosis)
- Lab Chemistry
    - C-reactive protein
    - Faecal calprotectin
- Serology
    - Anti-neutrophil cytoplasmic antibodies (pos/neg)

    - Anti-*Saccharomyces cerevisiae* antibodies (pos/neg)
- IBD Medication:
    - Mesalazines (yes/no)
    - Mesalazines Local (yes/no)
    - Steroids (yes/no)
    - Steroids Local (yes/no)
    - Mesalazines combined with Steroids local (yes/no)
    - Thiopurines (yes/no)
    - Methotrexate (yes/no)
    - Anti-TNF (yes/no)
- Other medication:
    - Antibiotics (also controls) (yes/no)
    - Proton Pump Inhibitors (also controls) (yes/no)
    - Anti-diarrheal medication (loperamide) (yes/no)
    - Iron (yes/no)
    - Minerals (yes/no)
    - Vitamins (yes/no)
    - Vitamins and Minerals (yes/no)
    - Bile salts (yes/no)
    - Osteoporosis medication (yes/no)
- Self-reported Diets:
    - Vegetarian (yes/no)
    - Vegan (yes/no)
    - On a diet (yes/no)
    - Cholesterol lowering diet (yes/no)
    - Energy limiting diet (weight losing diet) (yes/no)
    - Fat limiting diet (yes/no)
    - Fiber rich diet (yes/no)
    - Salt limiting diet (yes/no)
    - Macrobiotic diet (yes/no)
    - Anthroposophic diet (yes/no)
    - Diabetes diet (sugar limiting diet) (yes/no)
    - Other diet

- Environmental factors:
    - Smoking (yes/no)
- Extra-intestinal manifestations and/or complications
    - Any complications (yes/no)
    - Osteoporosis
    - Primary Sclerosing Cholangitis
    - Venous Thrombosis
    - Metastatic Crohn's disease
    - Eye complications/manifestations
        - Any eye complications (yes/no)
        - Episcleritis
        - Uveitis
        - Other eye complications
    - Mouth complications/manifestations
        - Any mouth complications (yes/no)
        - Ulcers
        - Ulcers due to Crohn's disease
        - Erosions in the mouth
        - Other mouth complications
    - Skin complications/manifestations
        - Any skin complications (yes/no)
        - Erythema nodosum
        - Hidradinitis
        - Palmoplantary pustulosis
        - Psoriasis
        - Pyoderma gangrenosum
        - Skin cancer
        - Other skin complications
    - Joints complications/manifestations
        - Any joint complications (yes/no)
        - Ankylosing spondilytis
        - Arthralgia
        - Arthritis

- Dactylitis
- Enthesitis
- Inflammatory back pain

2. Supplementary Methods: Selection of Healthy Controls

582 healthy controls were selected from the 1174 participants of the population cohort LifeLines-DEEP.[1] Based on self-reported questionnaires, LifeLines-DEEP participants who fulfilled the criteria of any of the ROME-III diagnoses (IBS-Constipation, IBS-Diarrhea, IBS-Mixed, IBS-Undetermined, functional bloating, functional constipation or functional diarrhea) were excluded. LifeLines-DEEP participants with celiac disease, gallstones, liver cirrhosis and/or diabetes (type 1 and type 2) were also excluded.[2–4] Further, participants who were on a vegan diet, diagnosed with an eating disorder or on a weight loss diet were excluded.[5] To exclude potential effects of concurrent drug-use on the gut microbiota, we excluded participants using any of the following systemic medications (topical lotions, droplets or creams were allowed): 1. Antibiotics, 2. Anti-fungal medication, 3. Anti-malaria medication, 4. Steroids, 5. Other immunosuppressants, 6. Cytostatics (chemotherapy), 7. Anti-hormones (for breast/prostate cancer or sex-change), 8. Anti-diabetics, 9. Opiates (as they might cause severe constipation), 10. Laxatives (e.g. movicolon), 11. Ursochol (changes bile acids) 12. HIV-inhibitors, 13. Magnesium (can cause diarrhoea). After applying the exclusion criteria, 313 IBD patients and 582 healthy controls were included in the study.

3. Supplementary Methods: Function Analysis

In order to investigate the functional implications of the gut microbiota of IBD patients, we inferred the metagenome composition of our samples. Using PICRUSt, the number of observations from the OTU-table were normalized considering the copy number of the 16S gene per OTU.[6] PICRUSt was also used to predict the gene composition of each sample. The abundance of 6910 genes was imputed. Using the HUMAnN software package, the predicted genes were combined into 181 microbial KEGG pathways.[7]

4. Supplementary Results: Taxonomy analysis of disease location

The differences in the gut microbiota that are related to disease location seen in the primary component analyses can also be seen in the analysis of the taxonomy. For example, the decreases of the genera *Faecalibacterium* and *Bifidobacterium* are only present in patients with ileal CD (FDR = 1.82×10^{-11} and FDR = 0.0046, respectively) or ileocolonic CD (FDR = 0.000037 and FDR = 0.030, respectively) but not in patients with colonic CD or UC when compared to healthy controls. All results of the disease localization analyses can be found in Supplementary Table S3. The relation between the gut microbiota and the extent of colonic disease in UC patients as defined by the Montreal classification (proctitis, left-sided colitis, pancolitis) was also analyzed. A decrease of the phylum Tenericutes was associated with more extensive UC disease (FDR = 0.032) (Supplementary Table S9).

5. Supplementary Results: Analysis of IBD subphenotypes

Serology: ANCA and ASCA

Serological measurements for Anti-neutrophil cytoplasmic antibodies (ANCA) and Anti-*Saccharomyces cerevisiae* antibodies (ASCA) were determined in IBD patients, but did not show any relation with the gut microbiota composition. (Supplementary Tables S10 and S11)

Stricturing and fistulizing disease behavior in CD patients

The family Mogibacteriaceae (FDR = 0.0096) and the genus *Faecalibacterium* (FDR = 0.032) were decreased in more stricturing and fistulizing disease behavior. (Supplementary Table S12)

Extra-intestinal Manifestations (EIM) and Complications of IBD and the gut microbiota

The skin manifestation psoriasis was associated with an increase of the family Enterococcaceae (FDR = 0.016) and the genus *Enterococcus* (FDR = 0.018) of the phylum Firmicutes, as well as an increase of an unclassified genus of the family Enterobacteriaceaea of the phylum Proteobacteria (FDR = 0.018). Other skin, joint, eye or mouth manifestations or complications of IBD were not associated with statistically significant changes in bacterial abundance. (Supplementary Table S13)

Effects of Medication on the gut microbiota of IBD patients

Several types of medication were associated with changes in the gut microbiota of IBD patients. The use of mesalazines was associated with an increase of the genus *cc_115* of the family Erysipelotrichaceae (FDR = 0.000021), as well as with increases of the genera *Anaerofustis* and *Oscillospira* (FDR = 0.013 and FDR = 0.024, respectively). The use of thiopurines (azathioprine, 6-mercaptopurine or 6-thioguanine) was associated with an increase of the family Enterococcaceae (FDR = 0.0096) and an increase of the genus *Enterococcus* (FDR = 0.049). Methotrexate use was associated with an increase of the genus *Atopobium* (FDR = 0.0046). The use of PPIs affected 9 bacterial taxa, including an increase of the genera *Streptococcus* (FDR = 7.38×10^{-14}) and *Rothia* (FDR = 6.85×10^{-10}). The use of steroids, TNF-a inhibitors, antibiotics, anti-diarrheal medication, bile salts, osteoporosis medication, iron, vitamins and/or minerals were not statistically significantly associated with gut microbiota differences in IBD patients. (Supplementary Table S14)

Mode of birth and breastfeeding

There were no statistically significant differences in the composition of the gut microbiota between IBD patients who had a vaginal versus a caesarean birth or were breastfed or not. (Supplementary Table S15)

Smoking in IBD patients

In all, 25% of IBD patients were smokers at the time of stool sample collection. The genus *Paraprevotalla* of the phylum Bacteroidetes was increased in smokers compared to non-smokers in IBD patients (FDR = 0.036). (Supplementary Table S16)

Self-reported diets, vegetarians

Only a limited number of IBD patients reported being on a diet as seen in Table 1. Self-reported diets or being a vegetarian did not have a statistically significant effect on the composition of the gut microbiota of IBD patients. (Supplementary Table S17)

6. Supplementary Results and Discussion: Imputed Function

Multiple metabolic pathways including butyrate metabolism, endotoxin metabolism and antibiotics resistance pathways were differentially expressed between IBD patients, UC patients, CD patients, ileal CD, ileocolonic CD and colonic CD as compared to healthy controls. The metabolic alterations are more profound in ileal disease (ileal CD) compared to colonic disease (colonic CD and ulcerative colitis). All altered KEGG pathways are presented in Supplementary Figure S3 and Supplementary Table S7 (FDR < 0.05). The top three altered pathways in the gut microbiota of CD patients were increased Sphingolipid metabolism (ko00600; FDR = 3.35×10^{-27}), decreased Aminoacyl tRNA biosynthesis (ko00970; FDR = 1.34×10^{-23}) and increased Other glycan degradation (ko00511; 4.13×10^{-23}). In UC patients, the increase of Other glycan degradation (ko00511; FDR = 9.10×10^{-11}), the increase of Cyanoamino acid metabolism (ko00460; FDR = 1.89×10^{-8}) and the increased Sphingolipid metabolism (ko00600; FDR = 3.64×10^{-8}) were the top three altered pathways of the gut microbiota.

Function of the gut microbiota affected by IBD: implications for intestinal inflammation

The decreased abundance of the butyrate-producing genera *Faecalibacterium* and *Ruminococcus* is reflected in the decreased butyrate metabolism in CD patients in the functional analysis. Together with the decreased abundance of *Bifidobacterium spp,* these gut microbiota alterations lead to a decrease of T_{reg}-induction and IL-10 production.[8–13] The increase of the phylum Proteobacteria, the class Gammaproteobacteria and the family Enterobacteriaceae is reflected in the increase in endotoxin metabolism in CD patients. These increased endotoxin levels enhance intestinal inflammation. Both the taxonomic analyses and the functional imputation show that IBD patients have a pro-inflammatory gut microbiota.

7. Supplementary Results: Figure Legends

- Supplementary Figure S1: Alpha diversity (Shannon Index) of IBD patients and healthy controls, depicted in a violin plot.
- Supplementary Figure S2: Imputed function analysis: increased and decreased KEGG-pathways of the gut microbiota in IBD, CD, UC, ileal CD, colonic CD and ileocolonic CD patients versus healthy controls (FDR < 0.05). Bacterial function was imputed using PICRUSt and HUMAnN.

8. Supplementary Results: Table Legends

- Supplementary Table S1: Taxa abundances of CD patients, UC patients and healthy controls on phylum, class, order, family and genus levels.
- Supplementary Table S2: MaAsLin results on genetic risk scores.
- Supplementary Table S3: MaAsLin results on diagnosis and location.
- Supplementary Table S4: MaAsLin results on disease activity in IBD patients.
- Supplementary Table S5: MaAsLin results on Montreal Classification: severity of UC.
- Supplementary Table S6: MaAsLin results on disease duration in IBD patients.
- Supplementary Table S7: MaAsLin results on imputed function (KEGG-pathways) of the gut microbiota of IBD, CD, ileal CD, ileocolonic CD, colonic CD and UC patients versus healthy controls.

- Supplementary Table S8: MaAsLin results on imputed function (KEGG-pathways) of the gut microbiota related to the Harvey Bradshaw Index disease activity metric of Crohn's Disease patients.
- Supplementary Table S9: MaAsLin results on Montreal Classification: extent of UC.
- Supplementary Table S10: MaAsLin results on anti-neutrophil cytoplasmic antibodies.
- Supplementary Table S11: MaAsLin results on anti-*Saccharomyces cerevisiae* antibodies.
- Supplementary Table S12: MaAsLin results on Montreal Classification: stricturing and fistulizing disease in CD patients.
- Supplementary Table S13: MaAsLin results on extra-intestinal manifestations and complications in IBD patients.
- Supplementary Table S14: MaAsLin results on IBD medication in IBD patients.
- Supplementary Table S15: MaAsLin results on mode of birth and breastfeeding of IBD patients.
- Supplementary Table S16: MaAsLin results on smoking behavior in IBD patients.
- Supplementary Table S17: MaAsLin results on self-reported diets of IBD patients.

References Supplementary Appendix

1 Tigchelaar EF, Zhernakova A, Dekens JAM. An introduction to LifeLines DEEP: study design and baseline characteristics. 2014;:0–21.
2 Qin N, Yang F, Li A, *et al.* Alterations of the human gut microbiome in liver cirrhosis. *Nature* Published Online First: 23 July 2014. doi:10.1038/nature13568
3 Qin J, Li Y, Cai Z, *et al.* A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 2012;**490**:55–60. doi:10.1038/nature11450
4 Tjellström B, Stenhammar L, Sundqvist T, *et al.* The effects of oats on the function of gut microflora in children with coeliac disease. *Aliment Pharmacol Ther* 2014;**39**:1156–60. doi:10.1111/apt.12707
5 Faith JJ, Guruge JL, Charbonneau M, *et al.* The long-term stability of the human gut microbiota. *Science* 2013;**341**:1237439. doi:10.1126/science.1237439
6 Langille MGI, Zaneveld J, Caporaso JG, *et al.* Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* 2013;**31**:814–21. doi:10.1038/nbt.2676
7 Abubucker S, Segata N, Goll J, *et al.* Metabolic Reconstruction for Metagenomic Data and Its Application to the Human Microbiome. *PLoS Comput Biol* 2012;**8**:e1002358. doi:10.1371/journal.pcbi.1002358
8 Velasquez-Manoff M. Gut Microbiome: The Peacekeepers. *Nature* 2015;**518**:S3–

11. doi:10.1038/518S3a
9 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. doi:10.1073/pnas.0804812105
10 Morgan XC, Tickle TL, Sokol H, *et al.* Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol* 2012;**13**:R79. doi:10.1186/gb-2012-13-9-r79
11 Gevers D, Kugathasan S, Denson LA, *et al.* The Treatment-Naive Microbiome in New-Onset Crohn's Disease. *Cell Host Microbe* 2014;**15**:382–92. doi:10.1016/j.chom.2014.02.005
12 Furusawa Y, Obata Y, Fukuda S, *et al.* Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* 2013;**504**:446–50. doi:10.1038/nature12721
13 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. doi:10.1038/nature12331