Gut microbial co-abundance networks show specificity in inflammatory bowel disease and obesity

Lianmin Chen, Valerie Collij, Martin Jaeger, Inge C. L. van den Munckhof, Arnau Vich Vila, Alexander Kurilshikov, Ranko Gacesa, Trishla Sinha, Marije Oosting, Leo A. B. Joosten, Joost H. W. Rutten, Niels P. Riksen, Ramnik J. Xavier, Folkert Kuipers, Cisca Wijmenga, Alexandra Zhernakova, Mihai G. Netea, Rinse K. Weersma, Jingyuan Fu.

Nature Communications 2020;11:4018
Abstract

The gut microbiome is an ecosystem that involves complex interactions. Currently, our knowledge about the role of the gut microbiome in health and disease relies mainly on differential microbial abundance, and little is known about the role of microbial interactions in the context of human disease. Here we construct and compare microbial co-abundance networks using 2,379 metagenomes from four human cohorts: an inflammatory bowel disease (IBD) cohort, an obese cohort and two population-based cohorts. We find that the strengths of 38.6% of species co-abundances and 64.3% of pathway co-abundances vary significantly between cohorts, with 113 species and 1,050 pathway co-abundances showing IBD-specific effects and 281 pathway co-abundances showing obesity-specific effects. We can also replicate these IBD microbial co-abundances in longitudinal data from the IBD cohort of the integrative human microbiome (iHMP-IBD) project. Our study identifies several key species and pathways in IBD and obesity and provides evidence that altered microbial abundances in disease can reflect their co-abundance relationship, which expands our current knowledge regarding microbial dysbiosis in disease.
Introduction

The human gut harbours a diverse community of microorganisms that interact closely with both the host and each other. Gut microorganisms are involved in digestion and degradation of nutrients, maintenance of digestive tract integrity, stimulation of the host immune system and modulation of the host metabolism. In recent years, associations have been identified between gut microbiome composition and the development of certain human diseases including diabetes, cardiovascular disorders, obesity and chronic gastrointestinal disorders like inflammatory bowel disease (IBD). Most associations to human diseases have been linked to lower microbial diversity, altered microbial composition and differing abundances of certain microorganisms and pathways. However, the gut microbiome is an ecosystem in which microbes can exchange or compete for nutrients, signalling molecules, or immune-evasion mechanisms through complicated ecological interactions that are far from fully understood. Enthusiasm has thus been rising to decipher these microbial interactions in order to detect key microbes in health and disease. One way of doing this is to create co-abundance networks based on correlations, a method that has the potential to study interactions between microbes and thereby generate hypotheses for experimental validation at a later stage.

Various network inference tools have been developed and applied to infer microbial taxonomic networks in healthy individuals and in individuals with extreme longevity, gestational diabetes, Crohn’s disease and colorectal cancer. These studies have identified microbial genera that are potentially key in health and disease, e.g. Porphyromonas and Bacteroides in gestational diabetes. However, these previous studies were either based on 16S rRNA sequencing data, which yields limited information on microbial species and pathways, or carried out in small cohorts. A further limitation of 16S sequencing is that it can only identify microbial networks up to genus level. As different bacterial species can have very different functional properties, analysis at genus level cannot fully capture the biochemical interactions between microbes. In consequence, the importance of metabolic network construction from metagenomics data has recently been highlighted.

Here we present a metagenomics-based network analysis for bacterial species and metabolic pathways in 2,379 individuals from four cohorts from the Netherlands (Figure S1): an IBD cohort (n=496), an obesity cohort (300OB, n=298) and two population-based cohorts (Lifelines-DEEP (LLD, n=1,135) and 500FG (n=450)). We compare the microbial taxonomic and functional networks under different host health conditions and identify potential key species and pathways that shape host-associated microbial networks (Figure 1). We find that the microbial species and pathway co-abundances vary significantly
between cohorts and report IBD- and obesity-specific co-abundance networks, which expands our current knowledge regarding microbial dysbiosis in disease. The network key species and pathways identified in IBD and obesity highlight their potential roles in regulating the microbial ecosystem in disease.

Results

Construction of gut microbial co-abundance networks
Metagenomic data of the 2,379 participants from the four cohorts was processed using the same pipeline (Figure 1). Principle coordinate analysis showed that microbial composition and functional profiles are largely overlapped, although we observed a significant shift in species composition in the IBD cohort (Figure S2). A total of 134 bacterial species and 343 microbial pathways that were present in more than 20% of the samples in at least one cohort were included for microbial network inference. We established microbial co-abundance relationships by combining the SparCC and SpiecEasi methods. For species networks, we identified 2,604 co-abundances in the LLD cohort, 1,591 in the 500FG cohort, 1,107 in the 300OB cohort and 2,554 in the IBD cohort, yielding 3,454 unique species co-abundances in total (FDR<0.05, Figure 2a, Data S1). Notably, 82.1% of the species co-abundances also exhibited co-occurrence (Figure S3, Data S1 & S2). For pathways, the numbers of co-abundances ranged from 37,279 in 500FG to 40,699 in LLD, yielding a total of 43,355 unique pathway co-abundances (FDR<0.05, Figure 2b, Data S3). Since absence rate of bacterial pathways is much less than in bacterial species, only 29.6% of pathway co-abundances showed co-occurrence (Figure S3, Data S3 & S4). The co-occurrence results are summarized and further discussed in Note S1.

Microbial co-abundance strength varies between cohorts
We hypothesized that co-abundance strengths could be different depending on host physiological status. We thus assessed to what extent the correlation coefficients were variable across cohorts and characterized variable co-abundance relationships for 38.6% of the species co-abundances and 64.3% of the pathway co-abundances (Cochran-Q test, FDR<0.05, Data S1 & S3).

Differential microbial co-abundances are reflected in abundance levels
When zooming in on the 100 species and 304 pathways that were involved in variable co-abundances, 76% of these species and 84% of these pathways also showed significant differences in their abundance levels among cohorts (ANOVA test FDR<0.05, Data S5 & S6). This implies that the variable co-abundance relationship is largely reflected by differential microbial abundance. We summarized the number of differential co-abundances between species from the same genus or from different genera (Figure 3a). The genus with the
Figure 1. Analysis workflow of present study.
most heterogeneous co-abundances was *Streptococcus*, and a large number of variable co-abundances were observed not only between different *Streptococcus* species, but also between *Streptococcus* species and species from other genera such as *Eubacterium* and *Veillonella* (Figure 3a). In particular, *Streptococcus* species were higher in the IBD cohort, consistent with the results of previous studies.14,39 A similar observation was found for the pathway co-abundances, particularly for amino acid biosynthesis pathways, which showed variability not only within themselves but also with respect to various pathways related to nucleoside and nucleotide biosynthesis (Figure 3b).

**Specific microbial co-abundances are enriched in disease cohorts**

Next, we analysed whether the variable co-abundance relationships were driven by a particular cohort, i.e. whether the co-abundance strength in one cohort was very different from those in the other three cohorts. After correcting for the age and sex differences between cohorts, 120 species co-abundances (Figure S4) and 1,448 pathway co-abundances (Figure S5) still showed cohort-specificity with an FDR of 7.6%, as estimated by permutation (Data S1 & S3). Interestingly, cohort-specific co-abundances were significantly enriched in the disease cohorts compared to the population-based cohorts: 113 (94%) species co-abundances and 1,050 (72%) pathway co-abundances were specifically related to the IBD cohort (Fisher’s test $P=1.2\times10^{-56}$ and $P<10^{-260}$, respectively, Figure 3c & d) and 281 (19.4%) pathway co-abundances were specifically related to the 300OB cohort (Fisher’s test $P=2.9\times10^{-29}$), as compared to only 3 species and 117 pathway co-abundance relationships specific to the population-based cohorts LLD and 500FG.

---

**Figure 2.** Microbial co-abundance networks in each cohort. a. Venn diagram of the numbers of species co-abundances detected in each cohort. In total, we identified 3,454 co-abundance relationships significant at FDR<0.05 in at least one cohort by combing SparCC and SpiecEasi. b. Venn diagram of numbers of species co-abundances detected in each cohort. Similarly, at microbial metabolic pathway level, 43,355 co-abundance relationships were detected at FDR<0.05.
Figure 3. Differential and cohort-specific microbial co-abundances. a. Differential species co-abundances involved in 45 microbial genera. b. Differential pathway co-abundances involved in 41 microbial metabolic categories. Each dot indicates one microbial genus or metabolic category. Each line represents differential species or pathway co-abundances between species or pathways from either the same or different genera or metabolic categories. The width and darkness of the lines represent the relative number of differential co-abundances. c. Pie chart of 120 cohort-specific species co-abundances showing the proportion of specific co-abundances detected in each cohort. d. Pie chart of 1,448 cohort-specific pathway co-abundances showing the proportion of specific co-abundances detected in each cohort.

(Figure 3c & d). Our results highlight that microbial co-abundances are dependent on host health and disease status. Below we discuss the microbial co-abundance networks in IBD and 300OB in more detail, further replicate our findings in independent cohorts, and assess the relevance of disease subtypes, disease characteristics and medication usage.
CHAPTER 3

The microbial co-abundance network in IBD

Replication of the IBD network in the iHMP-IBD cohort

Of the 2,554 species and 37,699 pathway co-abundances established in our IBD cohort, we were able to assess 2,090 species co-abundances and 37,106 pathway co-abundances in 77 IBD individuals from the integrative Human Microbiome Project (iHMP-IBD). In the baseline samples of the iHMP-IBD cohort, 531 species co-abundances (25.4%) and 21,882 (59.0%) pathway co-abundance could be replicated at P<0.05 (Data S7 & S8). The relatively low replication rate in species co-abundances is largely a power issue, as we also observed that 1,705 (81.6%) species co-abundances and 24,165 (65.1%) pathway co-abundances showed no significant difference in their co-abundance strengths between our IBD cohort and the iHMP-IBD cohort (Cochran-Q test, P>0.05, Figure S6, Data S7 & S8). We then compared the IBD networks between the first and last time points of the iHMP-IBD cohort (~1 year apart) and replicated 90.6% of species co-abundances and 99.6% of pathway co-abundances (Cochran-Q test, P>0.05, Figure S6, Data S7 & S8). This suggests that our estimation of co-abundance strengths in IBD was largely replicable in a different cohort and was stable across time.

Microbial networks of IBD in relation to disease characteristics

Previous studies have shown that observed microbial abundance differences could be explained by certain disease characteristics of IBD. We therefore hypothesized that this could also be the case for co-abundance relationships. We assessed whether IBD co-abundances (including IBD co-abundances at FDR<0.05 and IBD-specific co-abundances) could be related to the disease subtypes [ulcerative colitis (UC, n=189) vs. Crohn’s disease (CD, n=276)], disease location [ileum (n=212) vs. colon (n=286)] and disease activity [inflammation (n=121) vs. no inflammation (n=377)] (Table S1). Most of the co-abundance relationships were comparable between disease characteristics, and only a few showed significant differences at FDR<0.05 (Figure S7, Data S9 & S10), namely 16 species co-abundances related to disease subtypes and 8 species co-abundances related to location. For the pathway co-abundances, 91 were related to disease subtypes, 24 to location and 3 to activity (Cochran-Q test FDR<0.05, Figure S7). Out of these, five co-abundance relationships were related to an important butyrate producer, Faecalibacterium prausnitzii, which showed stronger co-abundance relationships in UC compared to CD. One example here was the negative co-abundance relationship of Faecalibacterium prausnitzii with Haemophilus parainfluenzae, a species known to have pathogenic properties.

Microbial networks of IBD in relation to medication

We further tested whether drug usage can affect microbial co-abundance, as usage of antibiotics (20.0%) and proton pump inhibitors (PPIs, 26.5%) was higher in patients with IBD than in the general population cohorts (1.1% and 8.4%) (Table S1). Here we detected
no significant difference in species co-abundances between antibiotic users and non-users (Cochran-Q test FDR>0.05, Figure S7), while 1,049 out of 37,959 (3.7%) pathway co-abundance relationships showed statistically significant differences between PPI users and non-users, in particular related to the isoprene biosynthesis and methylerythritol phosphate pathways (Cochran-Q test FDR<0.05, Figure S7, Data S10).

Key species and pathways in IBD

We identified 113 species co-abundances and 1,050 pathway co-abundances that showed significantly different effects compared to the other three cohorts. We then assessed whether these IBD-specific co-abundances were highly connected to a specific pathway or species which may be disease-relevant, and our analysis identified three key species and four key pathways for IBD (Figure 4). Key species included *Escherichia coli*, *Oxalobacter formigenes* and *Actinomyces graevenitzii*. *Escherichia coli* and *Oxalobacte formigenes* have previously been associated to IBD (Figure 4a, Data S5).14,24,45-49 Interestingly, *Escherichia coli* shows positive co-abundance relationships with species with pro-inflammatory properties, like *Streptococcus mutans*, and negative co-abundance relationships with species with anti-inflammatory properties, like *Faecalibacterium prausnitzii* (Data S1). The key species we identified for IBD, *Actinomyces graevenitzii*, is a microbe that is most often identified in the oral cavity or respiratory tract.41 Key IBD pathways included a C1 compound utilization and assimilation pathway (P23-PWY: reductive TCA cycle I), two vitamin biosynthesis pathways (FOLSYN-PWY: superpathway of tetrahydrofolate biosynthesis and salvage and PWY-6612: superpathway of tetrahydrofolate biosynthesis) and an amino acid biosynthesis pathway (PWY-5505: L-glutamate and L-glutamine biosynthesis) (Figure 4b, Data S6). The top key functional pathway in IBD was the reductive TCA cycle pathway (P23-PWY), which had 76 IBD-specific co-abundances, and 94.7% of these were replicated in the iHMP-IBD cohort (Data S3 & S6). The reductive TCA cycle has been recognized as a primordial pathway for the production of organic molecules for the biosynthesis of sugars, lipids, amino acids, pyrimidines and menaquinone (Figure 5a).42 For instance, one IBD-specific co-abundance relationship was related to the biosynthesis of menaquinone (PWY-5837), which is also known as vitamin K2. The co-abundance relationship for this pathway in IBD (r=0.1) was weaker than in other cohorts (r=0.3) (Figure 5b), despite the higher abundance of this pathway in IBD (FDR<0.05, Figure 5C, Data S6). *Escherichia coli* is known to be an important species for the biosynthesis of menaquinone, a growth-promoting factor for a variety of microorganisms in the gut microbiota.43 In line with this, we found that 18.8% of the menaquinone biosynthesis pathway in IBD patients was contributed by *Escherichia coli*, two times higher than in the two population-based cohorts (Wilcoxon-test P<3.0x10⁻¹¹) (Data S11). This finding suggests *Escherichia coli* as an important contributor to menaquinone biosynthesis in IBD that may promote the growth of other microorganisms. Indeed, our study also revealed *Escherichia coli* as a key IBD species, exerting IBD-specific co-abundance relationships with
15 species (Data S5). Of these, strong positive correlation was observed for inflammation-related *Streptococcus* species, including *Streptococcus mutans*, *Streptococcus vestibularis* and *Streptococcus infantis* (Figure 5D). Accordingly, higher correlations were observed between menaquinone biosynthesis and *Streptococcus* species in IBD than in the other cohorts (Figure 5e, Data S12).

**Figure 4.** Cohort-specific species and pathway co-abundances. a. Cohort-specific co-abundances identified for three key species in the IBD cohort, involving 33 IBD-specific co-abundances. Each dot indicates one species. Red indicates IBD key species. Each line represents one IBD-specific co-abundance relationship. b. Cohort-specific co-abundances identified for four key pathways in IBD and one key pathway in 300OB, involving 385 cohort-specific co-abundances. Each line represents a cohort-specific correlation between two pathways. Yellow lines represent obesity-specific co-abundances. Grey lines represent IBD-specific co-abundances. Each dot indicates one pathway. Pathways belonging to the same metabolic category have the same colour and are clustered as sub-circles. Colour legends are shown in the plot.
GUT MICROBIAL CO-ABUNDANCE NETWORKS SHOW SPECIFICITY IN INFLAMMATORY BOWEL DISEASE AND OBESITY

Figure 5. Menaquinone biosynthesis related to *Streptococcus* overgrowth in IBD. a. Menaquinone biosynthesis (PWY-5837) from the reductive TCA cycle (P23-PWY) in bacteria. b. The menaquinone biosynthesis pathway shows IBD-specific interaction with the reductive TCA cycle pathway. c. Both menaquinone biosynthesis and reductive TCA cycle pathway abundance are significantly higher (ANOVA test, FDR<0.05) in the IBD cohort than in the two population-based cohorts. Box plots show medians and the first and third quartiles (the 25th and 75th percentiles) of abundance after correcting for age and sex, respectively. The upper and lower whiskers extend the largest and smallest value no further than 1.5*IQR, respectively. Outliers are plotted individually. (Source data is provided as a Source Data file) d. Three *Streptococcus* species show IBD-specific co-abundance with *Escherichia coli*. e. The menaquinone biosynthesis pathway shows strong positive correlation with three *Streptococcus* species in IBD. N=2379 independent samples are involved (N_{LLD}=1135, N_{500FG}=450, N_{300OB}=298, N_{IBD}=496). The forest plots show co-abundance strength and direction in each cohort, with square dot for the correlation coefficient and the bar for the 95% confidence interval.
CHAPTER 3

The microbial co-abundance network in 300OB

Replication of 300OB network in LLD obese individuals. 1,107 species and 37,886 pathway co-abundances were detected in the 300OB cohort (Figure 2). These estimated co-abundance strengths were largely replicable in 134 obese individuals with matched age and BMI from the LLD cohort, with 991 (89.5%) species co-abundances and 32,963 (87.0%) pathway co-abundances showing no difference (Cochran-Q test P>0.05, Figure S8, Data S13 & S14).

Microbial networks in relation to obesity-related diseases

The 300OB cohort was set up to study cardiovascular disease in obese individuals, including 139 patients with atherosclerotic plaque and 159 obese controls (Table S1). In addition, 35 300OB participants had diabetes. Here we observed only three species co-abundances related to cardiovascular disease, with all three showing stronger co-abundances in patients with plaque than in patients without (Cochran-Q test FDR<0.05, Figure S9, Data S13 & S14). These were positive co-abundances between Dorea longticatena and Dorea formicigenerans and negative co-abundances of Lachnospiraceae bacterium 9.1.43BF AA with Coprococcus comes and Dorea longicatena.

Key pathways in obesity

When we compared microbial co-abundances in the 300OB to the other three cohorts, we identified 281 pathway co-abundances that showed a significantly different effect, i.e. obesity-specific co-abundances. One key pathway in obesity was degradation of allantoin (PWY0-41, Figure 4b, Data S6), which showed obesity-specific co-abundance relationships with 85 pathways. Allantoin is one of the active principles in various plants, e.g. yams, and is found to enhance insulin secretion and lower plasma glucose. Its degradation product, oxamate, plays an inhibitory role in oxaloacetate/aspartate amino acids. In line with this, we found that the allantoin degradation pathway showed stronger negative correlations with the biosynthesis pathways of oxaloacetate/aspartate amino acids (including lysine, homoserine, methionine, threonine and isoleucine) and the biosynthesis pathway of aspartate (PWY0-781, Figure 6), which were both positively associated with fasting glucose level and negatively associated with fasting insulin level (P<0.05, Table S2).
Figure 6. Allantoin degradation pathway links to glycaemia in obesity. The allantoin degradation pathway shows stronger negative correlation with 14 amino acid biosynthesis pathways in the obesity cohort compared to the other cohorts. These pathways represent the biosynthesis of six amino acids: a. isoleucine (PWY-5103, PWY-3001, BRANCHED-CHAIN-AA-SYN-PWY and ILEUSYN-PWY), b. methionine (PWY-6151, PWY-5347, MET-SAM-PWY and HOMOSER-METSYN-PWY), c. threonine (THRESYN-PWY and PWY-724), d. lysine (PWY-5097 and PWY-2942), e. aspartate (PWY-781) and f. homoserine (METSYN-PWY). All six amino acids are involved in the oxaloacetate/aspartate amino acids biosynthesis pathway. Lines with arrows represent metabolic relationships. Lines with a circle represent an inhibitory role in a metabolic pathway. N=2379 independent samples are involved (N_{LLD}=1135, N_{500FG}=450, N_{300OB}=298, N_{IBD}=496). The forest plots show co-abundance strength and direction in each cohort, with square dot for the correlation coefficient and the bar for the 95% confidence interval.
Discussion

This study is a microbial co-abundance network analysis based on metagenomics data, involving 2,379 participants from two population-based cohorts (LifeLines-DEEP and 500FG) and two disease cohorts (IBD and 300OB). We report 3,454 species and 43,355 pathway co-abundance relationships that were significant in at least one cohort. Among them, the effect sizes of 38.6% of species co-abundances and 64.3% of pathway co-abundances were significantly different between cohorts. In particular, 113 species co-abundances and 1,050 pathway co-abundances showed IBD cohort-specific effects and 281 pathway co-abundances had specific effects in the 300OB cohort. Our study provides evidence that microbial dysbiosis can be reflected in alterations in microbial co-abundance.

Our study yielded several findings. We identified three species and four pathways in IBD and one pathway in 300OB that served as key players in disease-specific co-abundance networks. Key IBD-associated species included *Escherichia coli* and *Oxalobacter formigenes*. A higher abundance of the pathogenic species *Escherichia coli* has previously been associated to IBD, likely due to an increased release of oxidized haemoglobin into the intestinal lumen as a result of chronic inflammation of the gastrointestinal walls. Consistent with this, we replicated high abundances of *Escherichia coli* and low abundances of anaerobic metabolism pathways in IBD. *Escherichia coli* also showed strong positive co-abundance with other inflammation-inducing species in IBD, including streptococcus species such as *Streptococcus mutans*, *Streptococcus vestibularis* and *Streptococcus infantis*. In contrast, these co-abundances were either weak or negative in our population-based and obesity cohorts. We further identified *Actinomyces graevenitzii* as a key species in IBD. Although no evidence supports a direct role for *Actinomyces* in the pathogenesis of IBD, *Actinomyces graevenitzii* has been associated to celiac disease in children and can induce actinomycosis, with both conditions sharing similar abdominal pathologies with IBD. Two case reports have also suggested that *Actinomyces* may aggravate the intestinal injuries caused by inflammation.

The top key functional pathway in IBD was the reductive TCA cycle pathway (P23-PWY), which had 76 IBD-specific co-abundances. Interestingly, the key IBD species *Escherichia coli* is known to be an important species for the biosynthesis of menaquinone, a growth-promoting factor for a variety of microorganisms in the gut microbiota. In line with this, we found that 18.8% of the menaquinone biosynthesis pathway in IBD patients was attributed to *Escherichia coli*, which is two-times higher than in the two population-based cohorts (Wilcoxon-test $P<3.0\times10^{-11}$). Another notable IBD key pathway is the tetrahydrofolate pathway, which is responsible for folic acid derivative biosynthesis.
GUT MICROBIAL CO-ABUNDANCE NETWORKS SHOW SPECIFICITY IN INFLAMMATORY BOWEL DISEASE AND OBESITY

and supplementation of folic acid. This pathway has been shown to reduce the risk of colorectal cancer in IBD patients.\(^6^0\) Interestingly, previous research has shown that oral intake of L-glutamine attenuates the colitis induced by dextran sulfate sodium in mice.\(^6^1\)

We identified a negative co-abundance with L-glutamine biosynthesis and biosynthesis of other amino acids like L-isoleucine and L-methionine. Previous research showed that both these amino acids play an important role in the immune system.\(^6^2,6^3\) L-glutamine has been tested as supplement in patients with IBD, but did not show improvements in clinical outcomes like disease activity scores.\(^6^4\) Our results show large numbers of connections for L-glutamine with other pathways such as the biosynthesis of other amino acids. These pathways might also be of interest when exploring L-glutamine as an intervention for IBD.

In obesity, we identified the allantoin degradation pathway as a key pathway, showing obesity-specific co-abundance relationship to 85 pathways, mainly negative correlations with biosynthesis of oxaloacetate/aspartate amino acids. These pathways are related to insulin secretion and glucose metabolism. However, their co-abundance relationships did not show significant differences between patients and non-diabetic individuals, which is likely due to a power issue as there were only 35 diabetic patients in 300OB. Instead, we found three species co-abundances related to presence of atherosclerotic plaque, involving *Dorea longticatena*, *Dorea formicigenerans*, *Lachnospiraceae bacterium 9.1.43BFAA* and *Coprococcus comes*. Notably, *D. longticatena* and *Lachnospiraceae* species have been linked to atherosclerotic cardiovascular disease.\(^9\)

Altogether, our analyses show that microbial dysbiosis in disease may not be driven solely by differences in abundance level, it may also reflect shifts in microbial interactions that are mirrored in co-abundance analyses. Particularly when applied to metagenomics sequence data, pathway-based co-abundance networks provide further insights into functional dysbiosis in IBD and obesity. However, we also acknowledge several limitations of our study. This is an *in-silico* network analysis based on correlation in bacterial abundance levels. Even with the large sample size, our study is still undersized for making comparisons to the number of interactions assessed. In recent years, many different network tools have been developed to tackle the statistical challenges in inferring networks for compositional data. In this study, we applied two independent methods, SparCC and SpiecEasi, to establish microbial co-abundance networks based on MetaPhlan and HUMAnN2 annotation. Our analysis can thus be biased due to these annotation tools. Other annotation tools, e.g. mcSEED , may yield different pictures of microbial community and functional profile, thereby identifying different co-abundance networks.\(^6^5\) Thus, such *in-silico*–based network inferences require further functional validation. Although bacterial genes are believed to be expressed uniformly, previous studies have also shown that meta-transcription can exert dynamic changes in response
to environmental perturbations that cannot be detected at the metagenome level.\textsuperscript{66–68} Thus, in order to understand the microbial ecosystem in terms of functional interaction in diseases, we need complementary approaches like meta-proteomics and meta-metabolomics that provide a more direct readout of the functional properties of the gut microbiome. Furthermore, the cross-sectional design of this study makes it hard to assess the stability of our findings over time. However, we did observe similar findings for the iHMP-IBD cohort for 98.2\% of species co-abundances and 99.4\% of pathway co-abundances between two time points spanning one year (Cochran-Q test \( P > 0.05 \)). This implies that co-abundance relationships are largely consistent over time.

Additionally, due to our study design, we cannot disentangle cause from consequence. Longitudinal studies are therefore warranted and should be combined with functional validation. Moreover, especially in the context of IBD, which is a heterogeneous disease, we had limited ability to pinpoint co-abundance networks to specific disease characteristics like the subtypes CD and UC. This is probably due to the lack of power to detect this by subgrouping our cohorts. Larger cohorts with well-documented disease characteristics are needed in the future.

This study presents the microbial network analysis to examine both microbial species and functional pathways based on metagenomics sequencing. Our data show that dysbiosis of the gut microbial ecosystem in disease can be assessed by the altered abundance level, but can also be seen at the level of microbial interaction, at least in terms of co-abundances. We have also identified IBD-specific and obesity-specific species and pathways that potentially play important roles in regulating the microbial ecosystem in disease, and these disease-specific microbial interactions extend our current knowledge about the role of the microbiome in disease.

Materials and methods

Study cohorts

All four cohorts used in this study have been described before.\textsuperscript{3,14,69,70} In short, the Lifelines Deep cohort (LLD) is a large prospective cohort study from the north of the Netherlands.\textsuperscript{71} LLD contains 58.20\% females and 41.80\% males, the mean age (SD) of participants is 45.04 (13.60) years and their mean BMI is 25.26 (4.18) (Figure S1). In this study, we included 1,135 LLD individuals for whom there is metagenomics and phenotype data.\textsuperscript{3} The 500 Functional Genomics (500FG) cohort consists of 534 healthy adult volunteers from the Netherlands.\textsuperscript{69,70} In 500FG, 56.50\% are women and 43.50\% are men, the mean age of participants is 27.43 (12.35) years and their mean BMI is 22.70 (2.72) (Figure S1). In this study, we included 450 500FG participants for whom metagenomics
GUT MICROBIAL CO-ABUNDANCE NETWORKS SHOW SPECIFICITY IN INFLAMMATORY BOWEL DISEASE AND OBESITY

Data is available. The 300 Obesity cohort (300OB) is a part of the Functional Genomics project and consists of 302 individuals from the Netherlands with a BMI higher than 27. These individuals have been clinically screened for obesity-related comorbidities. Around half of participants are clinically diagnosed with metabolic syndrome. 300OB is 55.70% male and 44.30% female, the mean age of participants is 67.07 (5.39) years and their mean BMI is 30.73 (3.48) (Figure S1). In this study, we included 298 participants from 300OB for whom metagenomics data is available. The 1000 Inflammatory Bowel Disease (1000IBD) cohort consists of patients with IBD recruited at the specialized IBD outpatient clinic of the University Medical Center Groningen in the Netherlands. IBD diagnosis was made based on accepted radiological, endoscopic and histopathological evaluation. The 1000IBD cohort is 60.70% female and 39.30% male, the mean age of participants is 43.45 (14.52) years and their mean BMI is 25.55 (5.17) (Figure S1). In this study, we included 496 IBD participants for whom metagenomics data is available.

Ethical approval
All participants signed an informed consent form prior to sample collection. Institutional ethics review board (IRB) approval was available for all four cohorts: the Lifelines DEEP (ref. M12.113965) and the IBD (IRB-number 2008.338) cohorts were approved by the UMCG IRB and the 500FG study (NL42561.091.12, 2012/550) and 300OB (NL46846.091.13) cohorts were approved by the Ethical Committee of Radboud University Nijmegen.

Metagenomic data generation and pre-processing
All participants from the four cohorts were asked to collect faecal samples at home and to place them in their home freezer (-20°C) within 15 minutes after production. Subsequently, a nurse visited the participant to pick up the faecal samples on dry ice and transfer them to the laboratory. Aliquots were then made and stored at -80°C until further processing. The same protocol for faecal DNA isolation and metagenomics sequencing was used for all four cohorts. Faecal DNA isolation was performed using the AllPrep DNA/RNA Mini Kit (Qiagen, cat. 80204). After DNA extraction, faecal DNA was sent to the Broad Institute of Harvard and MIT in Cambridge, Massachusetts, USA, where library preparation and whole genome shotgun sequencing were performed on the Illumina HiSeq platform. From the raw metagenomic sequencing data, low-quality reads were discarded by the sequencing facility and reads belonging to the human genome were removed by mapping the data to the human reference genome (version NCBI37) with Bowtie2 (v2.1.0).

The relative abundance of gut microbial taxonomic units was determined using MetaPhlan (v2.7.2), and the relative abundances of metabolic pathways were determined using the HUMAnN2 pipeline (v0.10.0), which maps DNA/RNA reads to a customized database of functionally annotated pan-genomes. HUMAnN2 reported the abundances of gene families from the UniProt Reference Clusters (UniRef90), which were further mapped...
to microbial pathways from the MetaCyc metabolic pathway database. In total, we detected 698 species and 489 pathways present in at least one of the four cohorts. To deal with sparse microbial data in the network analysis, we focused on species/pathways present in at least 20% of samples in at least one cohort. This provided a confined list of 134 species and 343 pathways for use in the network analysis. Together these accounted for, on average, 86.9% and 99.9% of taxonomic and functional compositions, respectively.

**Statistical analysis**

**Co-abundance network inference**

Co-abundance analysis on compositional data is challenging because it is likely to exhibit spurious correlations due to the dependency of fractions (i.e. relative abundance sums to 1). In particular, the problem can be more serious in a microbial community with low compositionality. We therefore first assessed the inverse Simpson index of microbial composition for the effective number of species ($n_{eff}$). Our analysis showed high compositionality in both functional pathway composition (2.09, 2.10, 2.11 and 2.08 in LLD, 500FG, 300OB and IBD, respectively) and species composition (10.74, 11.87, 12.30 and 8.80 in LLD, 500FG, 300OB and IBD, respectively). Following the suggestion of Weiss et al. based on their assessment of the performance of eight different methods (Bray-Curtis, Pearson, Spearman, CoNet, LSA, MIC, RMT and SparCC), we decided to use the SparCC method because it has been proven to be able to infer linear relationships with high precision for high diversity compositions with $n_{eff}$ lower than 13. Species composition data from MetaPhlan was converted to predicted read counts by multiplying relative abundances by the total sequence counts, and then subjected to a Python-based SparCC tool. For pathway analysis, the read counts from HUMAnN2 were directly used for SparCC. Significant co-abundance was controlled at FDR 0.05 level using 100x permutation. In each permutation, the abundance of each microbial factor was randomly shuffled across samples. To reduce indirect associations, we further applied SpiecEasi (v1.0.6), which infers the microbial network underlying graphical model using the concept of conditional independence. In this way, we obtained 3,454 species and 43,355 pathway co-abundances that were detectable by both methods (Figure 1).

**Co-occurrence network inference**

Presence and absence of each bacterial species and metabolic pathway were treated as binary traits. The pair-wise co-occurrence relationship between two microbial factors (species or pathway) in each cohort was assessed using Pearson’s Chi-squared test. If the number of co-occurrence pairs was greater than the number of co-exclusion pairs, the two microbial factors were considered to be a co-occurrence. If the number of co-occurrence pairs was less than the number co-exclusion pairs, the two factors were considered to be a co-exclusion. Permutation (100x) was conducted to determine significance at an
FDR<0.05. In each permutation, the presence and absence of each microbial factor was randomly shuffled across samples. At species level, we detected 6,015 co-occurrence relationships that were found in at least one cohort, with 3,423 found in LLD, 1,845 in 500FG, 1,199 in 300OB and 4,701 in IBD (Data S2). At pathway level, we detected 19,903 co-occurrence relationships that appeared in at least one cohort, with 13,501 found in LLD, 7,581 in 500FG, 7,580 in 300OB and 16,596 in IBD (Data S4).

Network heterogeneity analysis
To assess the variability of networks among the four cohorts, we conducted Cochran-Q tests to assess the heterogeneity of effect sizes and directions across the four cohorts for each co-abundance (correlation coefficient generated by SparCC) and co-occurrence (odds ratio). Here we treated each cohort as one study and conducted the Cochran’s Q test using the metagen() function from the package meta (v4.9.5) in R, which calculates the squared difference between individual study effects and the pooled effect using inverse variance weighting.85 For each co-abundance, the P-values from the Cochran-Q test were recorded, and co-abundances with significant heterogeneity were controlled at FDR 0.05 level determined by permutation (100x). In this case, all samples from the four cohorts were randomly shuffled across cohorts, i.e. shuffling the cohort labels but keeping the correlation structure of species and pathways intact. Co-abundances with a Cochran-Q test FDR<0.05 were considered heterogeneous, while co-abundances with Cochran-Q test P>0.05 were considered stable. We also summarized species co-abundance co-abundances based on microbial genus and pathway co-abundance co-abundances based on metabolic category.

Cohort-specific co-abundance selection
For heterogeneous co-abundances and co-occurrences (Cochran-Q test FDR<0.05), we further assessed whether these relationships showed cohort-specificity, i.e. whether the effect size of co-abundance/co-occurrence in one cohort was very different from that in the other three. In this analysis, effect size for co-abundance was the SparCC correlation coefficient and the odds ratio for co-occurrence. We adopted interquartile ranges based the outlier detection method (Figure S10).86 We ranked the effect sizes from low to high, say b1, b2, b3, b4, and then calculated the corresponding 25%, 50% and 75% quartile values (Q1, Q2 and Q3, respectively). Interquartile range (IQR) was then calculated, and we assessed whether the smallest or largest effect size fell outside of Q1-2xIQR or Q3+2xIQR. If only one met the condition, we called this co-abundance specific and assigned it to the corresponding cohort (Figure S10). To assess whether cohort-specific co-abundances were enriched for a specific cohort, we conducted Fisher’s exact test. We also calculated the average FDR of cohort-specific co-abundances using 100x permutations as described above for the heterogeneity analysis.
Key microbial species and pathway detection
To assess to what extent cohort-specific microbial relationships were linked to a specific species or pathway, we calculated the number of cohort-specific microbial relationships per species/pathway. To define the key species/pathway, we took the maximum number of false cohort-specific relationships per species/pathway from each permutation and determined the key species/pathway cut-off as the upper range of the 95% of confidence interval based on 100x permutations. At this cut-off, there is a 5% probability that a false enrichment could occur by chance. In this way, a species with at least 13 cohort-specific co-abundances or a pathway with at least 70 cohort-specific co-abundances was recognized as a key species or pathway. For co-occurrence networks, these numbers were 10 for key species and 45 for key pathways. In such a way, we detected 192 cohort-specific species co-abundances and 2,235 cohort-specific pathway co-abundances.

Assessing impact of confounding factors
The age and sex distributions were different between cohorts (Figure S1). To assess the impact of age and sex, we conducted partial correlation analysis (Figure S11). For example, to assess the co-abundance between species A and B, we first assessed the Pearson correlation of A and B to each covariant, say C, respectively. Then, a pairwise correlation matrix of A, B and C was subjected to partial correlation (Figure S11) using the partial correlation function cor2pcor from the R package corpcor (version 1.6.9). This insured that the partial correlation determined between A and B was independent of the covariant C. To assess the impact, we compared the correlation coefficient between SparCC correlation and partial correlation for all co-abundances and found comparable effect size (Figure S12). After regressing out the confounding effects of age and sex on cohort-specific co-abundances, 120 out of 192 (62.5%) species and 1,448 out of 2,235 (64.8%) pathway co-abundances remained cohort-specific.

Replication of microbial networks
To replicate microbial networks in IBD, we used data from 77 IBD patients from the Integrative Human Microbiome Project (iHMP-IBD) as a replication cohort. Given the iHMP-IBD’s longitudinal study design, we could examine metagenomics data from the first and the last sample collection for each individual. In all, 91% of the species (123 out of 134) and 99% of the pathways (340 out of 343) found in our IBD cohort were also detected in the first sample collection in iHMP-IBD. The differences in co-abundance strength between the IBD cohort and the iHMP-IBD cohort were also detected using the Cochran-Q test. A significant P>0.05 was applied to define replicable co-abundances. We also investigated the stability of microbial networks in iHMP-IBD by comparing the microbial co-abundances in the first and last sample collection from the same participants using the same approach.
To replicate microbial networks in 300OB, we selected 134 obese individuals from the LLD cohort with matched age and BMI. Here we considered a co-abundance to be replicable if the Cochran-Q test heterogeneity between the discovery and replication cohorts was not significant at \( P > 0.05 \).

**Assessing the relevance of microbial co-abundances to sub-phenotypes**

Patients in the IBD and 300OB cohorts have different disease subtypes, and both cohorts had higher proportions of drug users than our population cohorts. In particular, the IBD cohort contained 276 patients with CD and 189 with UC. Within the IBD cohort, 126 patients took PPIs and 97 took antibiotics. In the 300OB cohort, 53.4\% (159 out of 298) had an atherosclerotic plaque detected by ultrasound and 35 were diabetic.\(^72\) To assess the co-abundance related to disease sub-phenotypes, we split the cohorts based on disease subtypes or medication use and inferred microbial co-abundance using SparCC. The Cochran-Q test was applied to assess the differential microbial co-abundances at FDR<0.05.

**Species contributions to pathways and species–pathway associations**

Since the pathway abundances reported by HUMAnN2 are computed at both community and individual species level, we further looked into the contribution of species to each pathway and reported the top contributor (species).\(^76\) To show the functional relationship between species and pathways (e.g. whether a given pathway has the potential to promote the growth of a species through its metabolic products), we also checked the correlation (Spearman) between microbial species and pathway abundance after adjusting for age, sex and read depth using a linear regression model.\(^88\) FDR was further calculated based on 100x permutation.

**Network Visualization**

Cohort-specific networks based on cohort-specific co-abundances were visualized using a circle plot or heatmap with hierarchical clustering analysis (ward.D clustering based on Minkowski distance). Both species and pathways networks were visualized using the package igraph (v1.2.4.1) in R.\(^89\) For species networks, species belonging to the same genus were clustered together. For pathway networks, pathways from the same metabolic category were presented in a sub-circle, and categories with a limited number of pathways (fewer than 4) were grouped into the other category. Classification of pathways was based on the MetaCyc metabolic pathway database.\(^78,79\)

**Data availability**

All relevant data supporting the key findings of this study are available within the article and its Supplementary Information files. Data underlying Figure 5C and Figure S2 are provided as a Source Data file. Data underlying all the other Figures are provided.
in Supplementary Data files. The raw metagenomic sequencing data and human phenotypes (i.e. age and sex) used for the analysis presented in this study are available from the European Genome-phenome Archive and National Center for Biotechnology Information data repositories: LifeLines Deep cohort [https://www.ebi.ac.uk/ega/datasets/EGAD00001001991], 1000 IBD cohort [https://www.ebi.ac.uk/ega/datasets/EGAD00001004194], 300OB cohort [https://ega-archive.org/dacs/EGAC00001001143], 500FG cohort [https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA319574]. The iHMP data is available via: [https://ibdmdb.org/tunnel/public/summary.html]. Data access is subject to local rules and regulations.

**Code availability**
For this study the following software was used: kneadData (v0.4.6.1), Bowtie2 (v2.1.0), MetaPhlAn2 (v2.7.2), HUMAnN2 (v0.10.0), SparCC Python package, R (v3.5.2), SpiecEasi R package (v1.0.6) and meta R package (v4.9.5). Code used for generating the microbial abundance profiles is publicly available at: [https://github.com/GRONINGEN-MICROBIOME-CENTRE/Groningen-Microbiome/blob/master/Scripts/Metagenomics_pipeline_v1.md]. Code used for the statistical analyses is publicly available at: [https://github.com/GRONINGEN-MICROBIOME-CENTRE/Groningen-Microbiome/tree/master/Projects/Microbial%20co-abundance%20network].

**Supplementary materials**
https://drive.google.com/drive/u/0/folders/1c7ZRKSLptzaR3yMqmo1gTb-ATyjZTHXS
References

1. Chen LM, Garmaeva S, Zhernakova A, et al. A system biology perspective on environment-host-microbe interactions. Hum Mol Genet. 2018;27:eR187-R194
2. Falony G, Joossens M, Vieira-Silva S, et al. Population-level analysis of gut microbiome variation. Science. 2016;352:560-564
3. Zhernakova A, Kurilshikov A, Bonder MJ, et al. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. Science. 2016;352:565-569
4. Lloyd-Price J, Mahurkar A, Rahnavardet G, et al. Strains, functions and dynamics in the expanded Human Microbiome Project. Nature. 2017;550:61-66
5. Kurilshikov A, Wijmenga C, Fu J, et al. Host genetics and gut microbiome: Challenges and perspectives. Trends Immunol. 2017;38:633-647
6. Wu H, Esteve E, Tremaroli V, et al. Metformin alters the gut microbiome of individuals with treatment-naive type 2 diabetes, contributing to the therapeutic effects of the drug. Nat Med. 2017;23:850-858
7. Grasset E, Puel A, Charpentier J, et al. A specific gut microbiota dysbiosis of type 2 diabetic mice induces GLP-1 resistance through an enteric NO-dependent and gut-brain axis mechanism. Cell Metab. 2017;26:278
8. Fu J, Bonder MJ, Cenit MC, et al. The gut microbiome contributes to a substantial proportion of the variation in blood lipids. Circ Res. 2015;117:817-824
9. Jie Z, Xia H, Zhong SL, et al. The gut microbiome in atherosclerotic cardiovascular disease. Nat Commun. 2017;8:845
10. Liu R, Hong J, Xu X, et al. Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. Nat Med. 2017;23:859-868
11. Bouter KE, van Raalte DH, Groen AK, et al. Role of the gut microbiome in the pathogenesis of obesity and obesity-related metabolic dysfunction. Gastroenterology. 2017;152:1671-1678
12. Halfvarson J, Brislawn CJ, Lamendella R, et al. Dynamics of the human gut microbiome in inflammatory bowel disease. Nat Microbiol. 2017;2:17004
13. 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;67:108-119
14. Vich Vila A, Imhann F, Collilj V, et al. Gut microbiota composition and functional changes in inflammatory bowel disease and irritable bowel syndrome. Sci Transl Med. 2018;10:eaap8914
15. Imhann F, Bonder MJ, Vich Vila A, et al. Proton pump inhibitors affect the gut microbiome. Gut. 2016;65:740-748
16. Mangalam A, Shahi SK, Luckey D, et al. Human gut-derived commensal bacteria suppress CNS inflammatory and demyelinating disease. Cell Rep. 2017;20:1269-1277
17. Petriz BA, Franco OL. Metaproteomics as a complementary approach to gut microbiota in health and disease. Front Chem. 2017;5:4
18. Rosshart SP, Vassallo BG, Angeletti D, et al. Wild mouse gut microbiota promotes host fitness and improves disease resistance. Cell. 2017;171:1015-1028
19. Surana NK, Kasper DL. Moving beyond microbiome-wide associations to causal microbe identification. Nature. 2017;552:244-247
20. Baumler AJ, Sperandio, V. Interactions between the microbiota and pathogenic bacteria in the gut. Nature. 2016;535:85-93
21. Eickhoff MJ, Bassler BL. SnapShot: Bacterial Quorum Sensing. Cel. 2018;174:1328
22. Schirmer M, Smeekens SP, Vlamakis H, et al. Linking the human gut microbiome to inflammatory cytokine production capacity. Cell. 2016;167:1897
23. Berry D, Widder S. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. Front Microbiol. 2014;5:219
24. Banerjee S, Schlaeppi K, van der Heijden MGA. Keystone taxa as drivers of microbiome structure and functioning. Nat Rev Microbiol. 2018;16:567-576
25. Faust K, Sathirapongsasuti JF, Izard J, et al. Microbial co-occurrence relationships in the human microbiome. PLoS Comput Biol. 2012;8:e1002606
26 Xia LC, Ai D, Cram J, et al. Efficient statistical significance approximation for local similarity analysis of high-throughput time series data. Bioinformatics. 2013;29:230-237
27 Reshef DN, Reshef YA, Finucane HK, et al. Detecting associations in large data sets. Science. 2011;334:1518-1524
28 Deng Y, Jiang YH, Yang Y, et al. Molecular ecological network analyses. BMC Bioinf. 2012;13:113
29 Friedman J, Alm E J. Inferring correlation networks from genomic survey data. PLoS Comput Biol. 2012;8:e1002687
30 Faust K, Sathirapongsasuti JF, Izard J, et al. Microbial co-occurrence relationships in the human microbiome. PLoS Comput Biol. 2012;8:e1002606
31 Biagi E, Franceschi C, Rampelli S, et al. Gut microbiota and extreme longevity. Curr Biol. 2016;26:1480-1485
32 Flemer B, Lynch DB, Brown JM, et al. Tumour-associated and non-tumour-associated microbiota in colorectal cancer. Gut. 2017;66:633-643
33 Wang J, Zheng J, Shi W, et al. Dysbiosis of maternal and neonatal microflora associated with gestational diabetes mellitus. Gut. 2018;67:1614-1625
34 Gevers D, Kugathasan S, Denson LA, et al. The treatment-naive microbiome in new-onset Crohn’s disease. Cell Host Microbe. 2014;15:382-392
35 Yilmaz B, Juillerat P, Øyås O, et al. Microbial network disturbances in relapsing refractory Crohn’s disease. Nat Med. 2019;25:323-336
36 Rottgers L, Faust K. Can we predict keystones? Nat Rev Microbiol. 2019;17:193-193
37 Banerjee S, Schlaeppi K, van der Heijden MGA. Reply to ‘Can we predict microbial keystones?’. Nat Rev Microbiol. 2019;17:194-194
38 Kurtz ZD, Müller CL, Miraldi ER, et al. Sparse and compositionally robust inference of microbial ecological networks. PLoS Comput Biol. 2015;11:e1004226
39 Schirmer M, Garner A, Vlamakis H, et al. Microbial genes and pathways in inflammatory bowel disease. Nat Rev Microbiol. 2019;17:497-511
40 Musher DM. Haemophilus Species. Med Microbiol 4th edition, Galveston (TX), Galveston 1996
41 Hall V. Actinomyces--gathering evidence of human colonization and infection. Anaerobe. 2008;14:1-7
42 Smith E, Morowitz HJ. Universality in intermediary metabolism. Proc Natl Acad Sci U S A. 2004;101:13168-13173
43 Fenn K, Strandwitz P, Stewart EJ, et al. Quinones are growth factors for the human gut microbiota. Microbiome. 2017;5:161
44 Kojima A, Nakano K, Wada K, et al. Infection of specific strains of Streptococcus mutans, oral bacteria, confers a risk of ulcerative colitis. Sci Rep. 2012;2:332
45 Doyuk E, Ormerod OJ, Bowler ICJ. W. Native valve endocarditis due to Streptococcus vestibularis and Streptococcus oralis. J Infection. 2002;45:39-41
46 Pimenta F, Gertz Jr RE, Park SH, et al. Streptococcus infantis, Streptococcus mitis, and Streptococcus oralis strains with highly similar cps5 loci and antigenic relatedness to serotype 5 pneumococci. Front Microbiol. 2018;9:3199
47 Niu CS, Chen W, Wu HT, et al. Decrease of plasma glucose by allantoin, an active principle of yam (Dioscorea spp.), in streptozotocin-induced diabetic rats. J Agric Food Chem. 2010;58:12031-12035
48 Tsai CC, et al. Allantoin activates imidazoline I-3 receptors to enhance insulin secretion in pancreatic beta-cells. Nutr Metab. 2014;11:41
49 Marlier JF, Cleland WW, Zeczycki TN. Oxamate is an alternative substrate for pyruvate carboxylase from rhizobium etli. Biochemistry. 2013;52:2888-2894
50 Kotlowski R, Bernstein CN, Sepehri S, et al. High prevalence of Escherichia coli belonging to the B2+D phylogenetic group in inflammatory bowel disease. Gut. 2007;56:669-675
51 Mirespasi-Lauridsen HC, Halkjaer SI, Mortensen EM, et al. Extraintestinal pathogenic Escherichia coli are associated with intestinal inflammation in patients with ulcerative colitis. Sci Rep. 2016;6:31152
52 Zhu H, Li YR. Oxidative stress and redox signaling mechanisms of inflammatory bowel disease: updated experimental and clinical evidence. Exp Biol Med. 2012;237:474-480
53 Albenberg L, Esipova TV, Judge CP, et al. Correlation between intraluminal oxygen gradient and radial partitioning of intestinal microbiota. Gastroenterology. 2014;147:1055-1063
54 Ou G, Hedberg M, Hörstedt P, et al. Proximal small intestinal microbiota and identification of rod-shaped bacteria associated with childhood celiac disease. Am J Gastroenterol. 2009;104:3058-3067
55 Nagaoka K, Izumikawa K, Yamamoto Y, et al. Multiple lung abscesses caused by Actinomyces graevenitzii mimicking acute pulmonary coccidiomycosis. J Clin Microbiol. 2012;50:3125-3128
56 Kononen E, Wade, WG. Actinomyces and related organisms in human infections. Clin Microbiol Rev. 2015;28:419-442
57 Wong VK, Turmezei TD, Weston VC. Actinomycosis. BMJ. 2011;343:d6099
58 Lin K, Lin S, Lin AN, et al. A rare thermophilic bug in complicated diverticular abscess. Case Rep Gastroenterol. 2017;11:377-381
59 Burr NE, Hull MA, Subramanian V. Folic acid supplementation may reduce colorectal cancer risk in patients with inflammatory bowel disease. J Clin Gastro. 2017;51:247-253
60 Jeong SY, Im YN, Youm JY, et al. l-glutamine attenuates DSS-induced colitis via induction of MAPK phosphatase-1. Nutrients. 2018;10 pii:E288
61 Gu C, Mao X, Chen D, et al. Isoleucine plays an important role for maintaining immune function. Curr Protein Pept Sci. 2019;20:644-651
62 Martinez Y, Li X, Liu G, et al. The role of methionine on metabolism, oxidative stress, and diseases. Amino Ac. 2017;49:2091-2098
63 Limketkai BN, Wolf A, Parian AM. Nutritional interventions in the patient with inflammatory bowel disease. Gastroenterol Clin North Am. 2018;47:155-177
64 Overbeek R, Olson R, Pusch GD, et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res. 2014;42:D206-D214
65 Franzosa EA, Morgan XC, Segata N, et al. Relating the metatranscriptome and metagenome of the human gut. Proc Natl Acad Sci U S A. 2014;111:E2329-2338
66 Mehta RS, Abu-Ali GS, Drew DA, et al. Stability of the human faecal microbiome in a cohort of adult men. Nat Microbiol. 2018;3:347-355
67 Schirmer M, Franzosa EA, Lloyd-Price J, et al. Dynamics of metatranscription in the inflammatory bowel disease gut microbiome. Nat Microbiol. 2018;3:337-346
68 Schirmer M, Smeekens SP, Vlamakis H, et al. Linking the human gut microbiome to inflammatory cytokine production capacity. Cell. 2016;167:1125-1136
69 Ter Horst R, Jaeger M, Smeekens SP, et al. Host and environmental factors influencing individual human cytokine responses. Cell. 2016;167:1111-1124
70 Tigchelaar EF, Zhernakova A, Dekens JAM, et al. Cohort profile: LifeLines DEEP, a prospective, general population cohort study in the northern Netherlands: study design and baseline characteristics. BMJ Open. 2015;5:e006772
71 Kurilshikov A, van den Munckhof ICL, Chen L, et al. Gut microbial associations to plasma metabolites linked to cardiovascular phenotypes and risk: a cross-sectional study. Circ Res. 2019;124:1808-1820
72 Imhann F, Van der Velde KJ, Barbieri R, et al. The 1000IBD project: multi-omics data of 1000 inflammatory bowel disease patients; data release 1. BMC Gastroenterol. 2019;19:5
73 Langmead B, Trapnell C, Pop M, et al. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10:R25
74 Truong DT, Franzosa EA, Tickle TL, et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat Methods. 2015;12:902-903
75 Franzosa EA, McIver LJ, Rahnavard G, et al. Species-level functional profiling of metagenomes and metatranscriptomes. Nat Methods. 2018;15:962-968
76 UniProt Consortium. UniProt: a hub for protein information. Nucleic Acids Res. 2015;43:D204-D212
77 Caspi R, Billington R, Ferrer L, et al. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res. 2016;44:D471-D480
79 Caspi R, Billington R, Fulcher CA, et al. The MetaCyc database of metabolic pathways and enzymes. Nucleic Acids Res. 2018;46:D633-D639
80 Lark RM. Compositional data analysis in the geosciences: from theory to practice. J Roy Stat Soc a Sta. 2008;171:313-314
81 Filzmoser P, Hron K. Correlation analysis for compositional data. Math Geosci. 2009;41:905-919
82 Gloor GB, Wu JR, Pawlowsky-Glahn V, et al. It's all relative: analyzing microbiome data as compositions. Ann Epidemiol. 2016;26:322-329
83 van den Boogaart KG, Tolosana-Delgado R. "compositions": A unified R package to analyze compositional data. Comput Geosci. 2008;34:320-338
84 Weiss S, Van Treuren W, Lozupone C, et al. Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. ISME J. 2016;10:1669-1681
85 Schwarzer G. meta: An R package for meta-analysis. R News. 2007;7:6
86 Barbato G, Barini EM, Genta G, et al. Features and performance of some outlier detection methods. J Appl Stat. 2011;38:2133-2149
87 Integrative HMP (iHMP) Research Network Consortium. The integrative human microbiome project: dynamic analysis of microbiome-host omics profiles during periods of human health and disease. Cell Host Microbe. 2014;16:276-289
88 Zhernakova DV, Le TH, Kurilshikov A, et al. Individual variations in cardiovascular-disease-related protein levels are driven by genetics and gut microbiome. Nat Genet. 2018;50:1524-1532
89 Csardi GNT. The igraph software package for complex network research. InterJournal, Complex Sys. 2006;1695:doi:http://igraph.org/
GUT MICROBIAL CO-ABUNDANCE NETWORKS SHOW SPECIFICITY IN INFLAMMATORY BOWEL DISEASE AND OBESITY