

Quantitative microbiome profiling links gut community variation to microbial load

Doris Vandeputte^{1,2,3*}, Gunter Kathagen^{1,2*}, Kevin D'hoel^{1,2,3*}, Sara Vieira-Silva^{1,2*}, Mireia Valles-Colomer^{1,2}, João Sabino⁴, Jun Wang^{1,2}, Raul Y. Tito^{1,2,3}, Lindsey De Commer¹, Youssef Darzi^{1,2}, Séverine Vermeire⁴, Gwen Falony^{1,2§} & Jeroen Raes^{1,2§}

Current sequencing-based analyses of faecal microbiota quantify microbial taxa and metabolic pathways as fractions of the sample sequence library generated by each analysis^{1,2}. Although these relative approaches permit detection of disease-associated microbiome variation, they are limited in their ability to reveal the interplay between microbiota and host health^{3,4}. Comparative analyses of relative microbiome data cannot provide information about the extent or directionality of changes in taxa abundance or metabolic potential⁵. If microbial load varies substantially between samples, relative profiling will hamper attempts to link microbiome features to quantitative data such as physiological parameters or metabolite concentrations^{5,6}. Saliently, relative approaches ignore the possibility that altered overall microbiota abundance itself could be a key identifier of a disease-associated ecosystem configuration⁷. To enable genuine characterization of host–microbiota interactions, microbiome research must exchange ratios for counts^{4,8,9}. Here we build a workflow for the quantitative microbiome profiling of faecal material, through parallelization of amplicon sequencing and flow cytometric enumeration of microbial cells. We observe up to tenfold differences in the microbial loads of healthy individuals and relate this variation to enterotype differentiation. We show how microbial abundances underpin both microbiota variation between individuals and covariation with host phenotype. Quantitative profiling bypasses compositionality effects in the reconstruction of gut microbiota interaction networks and reveals that the taxonomic trade-off between *Bacteroides* and *Prevotella* is an artefact of relative microbiome analyses. Finally, we identify microbial load as a key driver of observed microbiota alterations in a cohort of patients with Crohn's disease¹⁰, here associated with a low-cell-count *Bacteroides* enterotype (as defined through relative profiling)^{11,12}.

First, we collected a set of 40 fresh faecal samples (the study cohort), which were processed within one hour of egestion. We compiled an accompanying set of basic matching metadata, with an emphasis on anthropometrics and stool characteristics (Supplementary Table 1). Given expected dietary effect sizes² and cohort limitations, participants were not requested to keep food records. Sample analysis was aligned with Flemish Gut Flora Project (FGFP) protocols². Metadata exploration reaffirmed the previously reported association of stool consistency (Bristol Stool Scale (BSS) score¹³) with moisture¹⁴ (Spearman's $\rho = 0.45$, $P = 5.2 \times 10^{-3}$; Supplementary Table 2). Microbiome analysis of frozen faecal aliquots characterized the sample set as within the bounds of FGFP community space, distributed over four enterotypes that were identified on the basis of Dirichlet multinomial mixtures (DMM) (Fig. 1a). Stool moisture and donor age were identified as the microbiome covariates that displayed the largest non-redundant effect size, jointly explaining 9.3% of inter-individual microbiota variation (stepwise distance-based redundancy analysis (dbRDA);

Supplementary Table 3). Association analyses confirmed several previously reported FGFP genus–metadata associations, including the covariation of stool consistency with *Akkermansia* and *Methanobrevibacter*^{15,16} (Supplementary Table 4).

Next, we determined total microbial cell counts in faecal samples using flow cytometry. Because microbiome analyses often begin with frozen material and freeze–thaw cycles can affect cell integrity¹⁷, we compared counts obtained from both fresh and frozen faecal

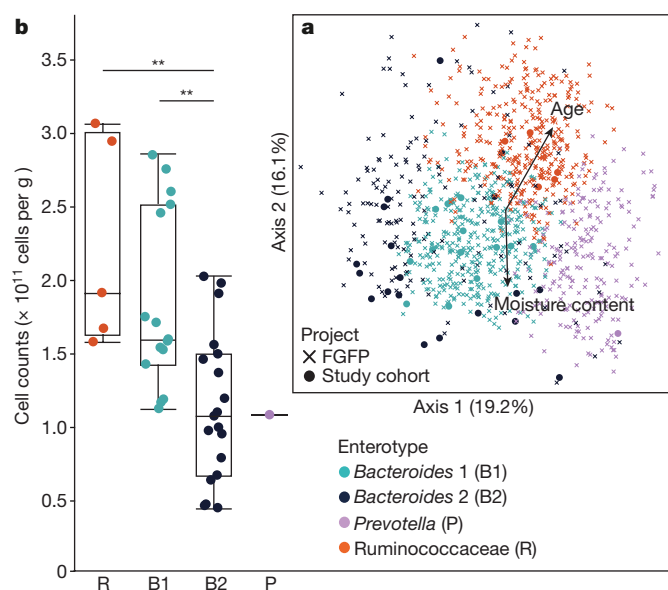


Figure 1 | Faecal microbial loads vary across enterotypes. **a**, Genus-level faecal microbiome community variation, represented by principal coordinates analysis (Bray–Curtis dissimilarity PCoA). Samples from the study cohort (full circles, $n = 40$) and the FGFP cohort (crosses, $n = 1,106$) were enterotyped and coloured accordingly. Stool moisture content and donor age were fitted onto the ordination (arrows scaled to contribution). The percentage of variance explained by the two first PCoA dimensions are reported on the axes. **b**, Microbial load differences between the four enterotypes in the study cohort ($n = 40$). Box plot representation of microbial load (cells per gram of faeces) distribution across the four enterotypes. The body of the box plot represents the first and third quartiles of the distribution and the median line. The whiskers extend from the quartiles to the last data point within $1.5 \times$ interquartile range, with outliers beyond. Two-sided Dunn's adjusted test, $**P < 0.01$. Significant differences in microbial abundance between the *Prevotella* sample and other enterotypes were not assessed. The occurrence of a low-cell-count *Bacteroides* B2 enterotype was confirmed in a disease cohort, with the cell counts of *Prevotella* samples being similar to those of the Ruminococcaceae and *Bacteroides* B1 enterotypes (Extended Data Fig. 4).

¹KU Leuven – University of Leuven, Department of Microbiology and Immunology, Rega Institute, Herestraat 49, B-3000 Leuven, Belgium. ²VIB, Center for Microbiology, Kasteelpark Arenberg 31, B-3000 Leuven, Belgium. ³Research Group of Microbiology, Department of Bioengineering Sciences, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium. ⁴Translational Research Center for Gastrointestinal Disorders (TARGID), KU Leuven, B-3000 Leuven, Belgium.

*These authors contributed equally to this work.

§These authors jointly supervised this work.

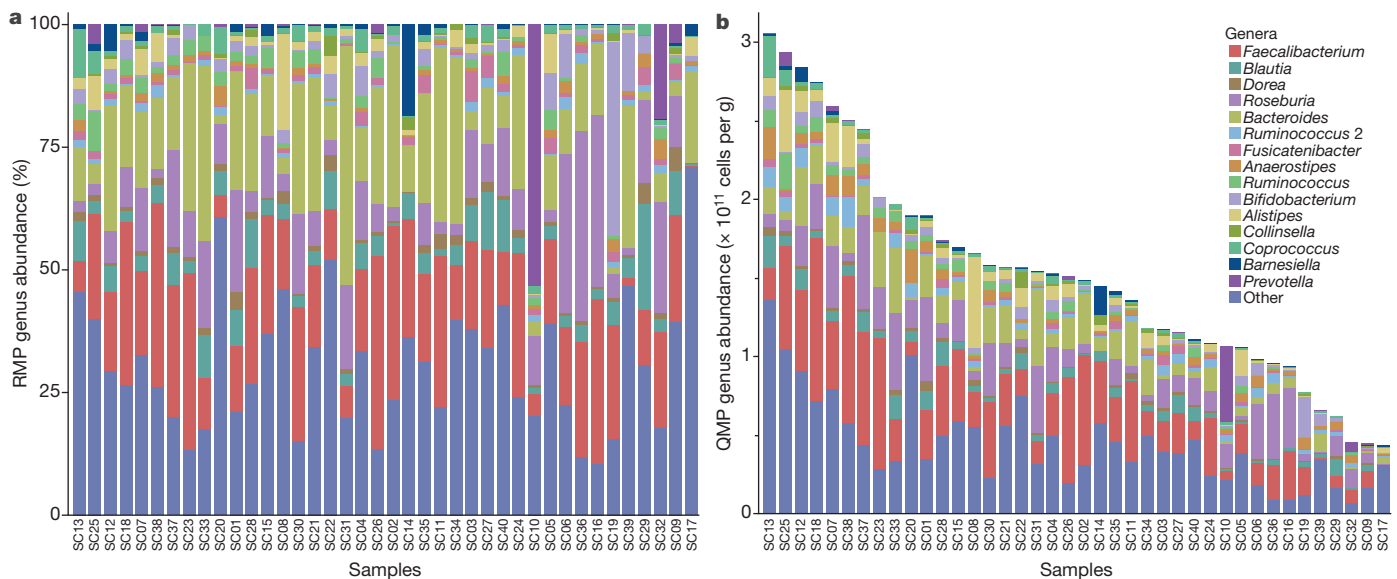


Figure 2 | Relative versus quantitative microbiome profiling. Genus-level faecal microbiome composition of study cohort participants ($n = 40$). **a**, Relative microbiota profiles deduced from standard microbiome sequencing protocols. **b**, Quantitative microbiome profiles deduced from

complementing sequencing with microbial cell counts (cells per gram of faeces). Samples are ordered according to decreasing microbial load. The top 15 most abundant genera are depicted, with all others pooled into 'Other'.

aliquots and found them to be strongly correlated (Pearson's $r = 0.91$, $P = 4.9 \times 10^{-16}$; Extended Data Fig. 1a). Although method-specific technical biases affect the outcomes of both cell-based and molecular microbial enumeration workflows¹⁸, a comparison between quantitative PCR (qPCR) and flow cytometric load assessment yielded comparable abundance profiles (Pearson's $r = -0.53$, $P = 4.7 \times 10^{-4}$; Extended Data Fig. 1b). Focusing our analyses on frozen samples, we observed up to a tenfold variation in cell counts between individuals. Microbial loads were shown to vary between 4.3×10^{10} and 3.1×10^{11} cell counts per gram of faecal material (median 1.5×10^{11} cell counts per gram; Supplementary Table 5), in agreement with previous reports that additionally characterized half of these cells as damaged or dead¹⁹. To assess longitudinal variation in abundance profiles, we quantified cell counts in stool samples collected from 20 healthy individuals (10 women and 10 men) over the course of a week (Supplementary Table 1). Individual microbial load profiles varied substantially (intraclass rank correlation coefficient (ICC) = 0.46). Single daily fluctuations (load $\text{day}_{x+1} - \text{load day}_x$) ranged in magnitude between 1.1×10^8 (participant LC02) and 1.6×10^{11} cell counts per gram (participant LC16; average daily fluctuation 3.8×10^{10} cell counts per gram; Extended Data Fig. 2), emphasizing the need to integrate longitudinal elements in microbiome study designs. Time-series cell counts decreased with stool moisture, although in the study cohort this association was not significant; this decrease was subsequently confirmed in an independent validation dataset (Supplementary Tables 1, 6; Extended Data Fig. 3a).

To integrate cell counts in microbiota analyses, we assessed their associations with microbiome features. We found cell counts to correlate mildly with observed genus richness (Spearman's $\rho = 0.36$, $P = 2.3 \times 10^{-2}$; Extended Data Fig. 3b). In addition, we observed an association between microbial loads and the enterotypes identified on the basis of DMM^{11,12} (Kruskal–Wallis test, $P = 1.23 \times 10^{-3}$; Supplementary Table 5). Higher cell counts were observed in samples that contained large fractions of Ruminococcaceae, and cell densities differentiated two *Bacteroides* clusters (Fig. 1b; Extended Data Fig. 4). The observed genus richness and enterotype associations indicate that microbial load variation does not merely reflect cell concentration as a result of absorption processes along the intestinal tract, as might be suggested by correlations with stool moisture. Instead, cell count dynamics appear to be linked with the previously reported ecosystem differentiation associated with transit time²⁰ and the accompanying

reduction in water content^{2,15}. Cell counts did not correlate with magnitudes of microbiome dissimilarity between individuals (dbRDA, $R^2 = 3.0\%$, $P = 0.29$); although microbial loads differ between enterotypes, relative microbiome composition cannot be used to infer overall microbiota abundance. Indeed, when screening for taxa that were significantly associated with cell counts, only a single positive correlation between *Ruminococcus* and microbial load was retained after correcting for multiple testing (Spearman's $\rho = 0.51$, false discovery rate (FDR) = 4.4×10^{-2} ; Supplementary Table 7).

Next, we used cell counts to transform sequencing data into an absolute microbiome abundance matrix that allowed quantitative microbiome profiling (QMP; in contrast to relative microbiome profiling, RMP), by modifying sequencing depth rarefying procedures. Despite criticism²¹, rarefying sequencing output to an equal number of reads per sample remains a common practice in microbiome research²². However, the observed variation in the number of reads produced in a sequencing process is a technical artefact; for example, it may be the result of equimolar pooling of sample DNA libraries to enhance sequencing effectiveness. In the current dataset, this is confirmed by the fact that sequencing depths did not reflect the microbial loads of the samples (Spearman's $\rho = 0.17$, $P = 0.28$; Extended Data Fig. 5). Although rarefying to equal sample size or sequencing depth is essential for comparing biological diversity between samples, it is inadequate if samples are drawn from ecosystems with markedly different population sizes and species abundance distributions²³ (Extended Data Fig. 6). Given the observed tenfold variation in faecal microbial loads between individuals (Supplementary Table 5), we propose correcting for sampling intensity by rarefying to an even sampling depth (rather than to an even sequencing depth), calculated as 16S rRNA gene copy-number-corrected sequencing depth divided by sample cell count. For each sample, the resulting rarefied genus abundances are proportional to cell counts and can be extrapolated to the total microbial load of each sample. This extrapolation generates quantitative microbiome profiles expressed as the number of cells per gram (Fig. 2; see Extended Data Fig. 6). For interclass comparisons, fitting a negative binomial distribution that accounts for sampling depth rather than sequencing depth²¹ could provide an alternative to the quantitative rarefying approach described above.

To assess the influence of QMP on the outcome of microbiome analyses, we investigated the effect of QMP on genus abundance

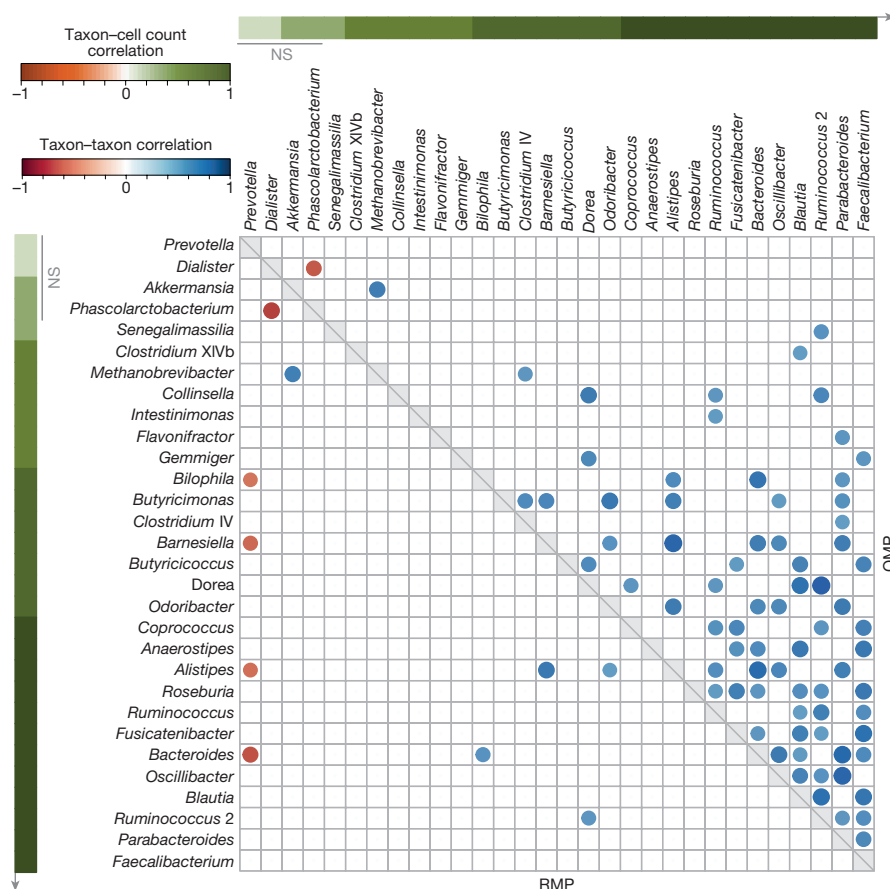


Figure 3 | Relative versus quantitative microbiota network reconstruction. Taxon–taxon interactions in 66 healthy controls from the disease cohort. Pairwise correlations between taxon abundances were calculated from microbiome profiles produced using RMP (lower triangle) and QMP (upper triangle). For significant interactions (two-sided adjusted test, $P < 0.05$), the correlation coefficient (Spearman's ρ) is represented by

the colour and size of the circles. The taxa are ordered by the significance of the correlation between their QMP abundance and the individuals' microbial cell counts (cells per gram of faeces); the correlation coefficient is represented by the colour gradients on the matrix axes (Spearman's two-sided test, $FDR < 0.05$; NS, non-significant; Supplementary Table 10).

profiles in healthy individuals. Because associations between microbiomes and metadata are generally investigated using non-parametric methods (often on the basis of ranks), we analysed whether the sample rank order for each genus was conserved between RMP and QMP analyses. We observed significant rank order position shifts for multiple genera and found that the extent of these shifts was dependent on genus. Rank order concordance varied widely even within the top ten most abundant genera, ranging from *Faecalibacterium* (lowest concordance) to *Prevotella* (highest concordance; Kendall's rank correlation test, τ_B range = 0.60–0.88; Supplementary Table 8), confirming our hypothesis that absolute abundance profiles differ significantly from those generated by relative approaches. As expected, rank changes affected the outcomes of association analyses. Only three of eight FGFP genus–metadata correlations confirmed by RMP could be validated using the QMP approach, all of which were linked to variation in stool consistency (Supplementary Table 4). When assessing non-redundant effect sizes of the metadata in quantitative microbiome variation, we no longer detected an independent contribution of age. The effect of stool moisture on QMP remained significant and accounted for 4.3% of quantitative inter-individual microbiota variation (dbRDA, 7.3% in RMP; Supplementary Table 3).

In a comparative analysis, it is impossible to establish absolute growths or declines of particular genera on the basis of relative taxon abundances. Such genera abundance shifts are particularly relevant to deducing and interpreting species interaction networks that are based on co-occurrence, as these analyses have been shown to be susceptible to the compositionality effects that result from relative abundance

measures^{5,24}. To assess the impact of QMP on genus co-occurrence patterns, we reanalysed 66 samples that had been selected from the FGFP² as healthy controls for a recent disease-focused microbiome study¹⁰. Genus co-occurrence networks were reconstructed using both RMP and QMP data matrices (Fig. 3; Supplementary Table 9). A far larger number of significantly co-varying genus pairs were detected in the QMP network (76, versus 10 in the RMP network), most of which were part of a network module associated with total faecal microbial load. Most of the RMP network pairs were recovered in QMP analyses, with the noteworthy exception of *Prevotella* co-exclusions. These latter include the often reported²⁵ trade-off between *Prevotella* and *Bacteroides* (RMP, Spearman's $\rho = -0.59$, $FDR = 1.6 \times 10^{-4}$), which we found to be no longer significant after integrating sample total cell count (QMP, Spearman's $\rho = -0.33$, $FDR = 1$; Extended Data Fig. 7). In contrast to *Prevotella*, *Bacteroides* QMP abundances correlated strongly with total microbial loads, resulting in the loss of any significant association between these two genera in a quantitative context. Benchmarking a relative versus quantitative approach using simulated data (based on negative binomial distributions) revealed both increased detection of true associations (sensitivity) and a decreased false discovery rate in QMP, compared with RMP (t -tests, $P < 10^{-15}$; Extended Data Fig. 8). Notably, widening simulated inter-individual microbial load ranges from mild (twofold) to average (up to eightfold) maximum differences doubled the performance gap between RMP and QMP.

To demonstrate the potential effect of QMP on clinical research, we expanded the dataset of healthy controls by performing QMP

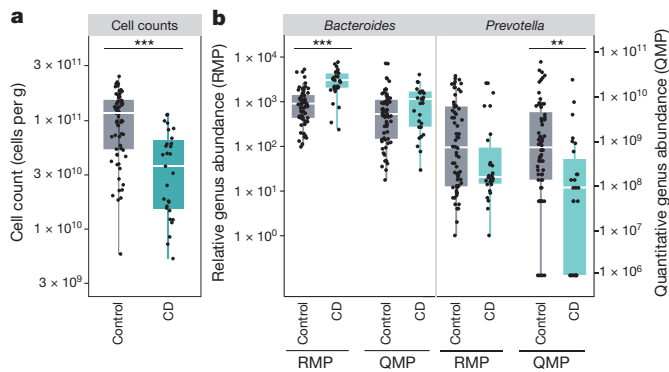


Figure 4 | Quantitative microbiome alterations in Crohn's disease.

Microbiota alterations in a cohort of patients with Crohn's disease, using a relative or quantitative microbiome profiling approach. **a**, Differences in microbial load (cells per gram of faeces, log-transformed y axis) between healthy controls (Control, $n = 66$) and patients with Crohn's disease (CD, $n = 29$). Wilcoxon rank-sum test, *** $P < 0.001$. **b**, Example of discordant genera when microbiota alterations in samples from patients with Crohn's disease are assessed using RMP or QMP (log-transformed y axis): *Bacteroides* is significantly increased in Crohn's disease patients using RMP, but the signal is lost using QMP. Conversely, *Prevotella* is detected as decreased in patients with Crohn's disease only when using QMP. The body of the box plot represents the first and third quartiles of the distribution and the median line. The whiskers extend from the quartiles to the last data point within $1.5 \times$ interquartile range, with outliers beyond. Kruskal–Wallis test, *** $FDR < 0.001$, ** $FDR < 0.01$.

on samples from 29 patients with Crohn's disease, also described in the previous analysis¹⁰. Flow cytometric analysis of microbial loads revealed cell counts that were three times lower in the stool samples of patients with Crohn's disease as compared to healthy controls (Kruskal–Wallis test, $P < 10^{-6}$; Fig. 4 and Extended Data Fig. 2). When projected onto the FGFP backbone, enterotyping on the basis of RMP caused most of the samples from patients with Crohn's disease to stratify into the low-cell-count *Bacteroides* cluster previously identified in the study cohort (Extended Data Fig. 4). Although RMP did therefore permit discrimination between patients and controls on the basis of microbiomes, only QMP analyses identified reduced microbial abundance as a key feature of the microbiome alterations associated with Crohn's disease. Detailed assessment of disease-associated microbiome signals produced by both RMP and QMP yielded signatures for Crohn's disease that overlapped substantially; however, these signatures also exhibited noteworthy discrepancies (Supplementary Table 11). Among these discrepancies, we highlight the fact that *Bacteroides* abundances are associated with Crohn's disease only when applying RMP, whereas *Prevotella* abundances are found to be decreased in patients with Crohn's disease only when using QMP (Fig. 4). These observations emphasize the limitations of RMP data analysis: an erroneous interpretation derived from the results of proportional profiling might suggest a putative causal role for *Bacteroides* in the onset or development of Crohn's disease. Overall, QMP revealed that more than half of the genera eligible for analysis were reduced in the stool samples of patients with Crohn's disease (37 of 64, versus 20 of 64 in RMP). Accordingly, RMP underestimated the well-documented decrease in microbiota richness associated with Crohn's disease²⁶. Using a quantitative approach, we identified that this key microbiome signal is driven by an overall reduction in microbiota cell counts rather than by the blooming of disease-associated taxa, with the effect size of Crohn's disease on observed species richness being estimated at 51.5% using QMP (Wilcoxon rank-sum test, $P = 1.2 \times 10^{-7}$) compared to 32.3% using RMP ($P = 1.4 \times 10^{-3}$, Extended Data Fig. 9).

By integrating microbial cell counting into a sequencing workflow, we demonstrated that microbial loads differ significantly between individuals and that they are associated with specific types of faecal

microbiota. Next, we used count data to transform sequencing-generated relative genus abundance data into quantitative microbiome profiles. Applying QMP to samples taken from healthy controls and from a cohort of patients with Crohn's disease, we demonstrated that QMP has a substantial effect on co-occurrence analyses and the characterization of disease-associated microbiota alterations. Although our workflow does not address all of the current biases in microbiome research, it presents a method that is straightforward and easy to implement and that permits the quantitative assessment of microbiota variation. This method can be readily applied to investigate the variation in microbiota that occurs between individuals or over time, is associated with diseased or healthy states, or that results from disease treatment. Given its throughput and technical feasibility, QMP should permit us to assess key microbiome findings founded on relative profiling, in a quantitative context. With minor modifications, QMP can be translated for application in any meta-omic workflow, enabling clinical microbiomics to move towards becoming a quantitative diagnostic science.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 13 March; accepted 22 September 2017.

Published online 15 November 2017.

- Zhernakova, A. *et al.* Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science* **352**, 565–569 (2016).
- Falony, G. *et al.* Population-level analysis of gut microbiome variation. *Science* **352**, 560–564 (2016).
- Valles-Colomer, M. *et al.* Meta-omics in inflammatory bowel disease research: applications, challenges, and guidelines. *J. Crohn's Colitis* **10**, 735–746 (2016).
- Satinsky, B. M., Gifford, S. M., Crump, B. C. & Moran, M. A. Use of internal standards for quantitative metatranscriptome and metagenome analysis. *Methods Enzymol.* **531**, 237–250 (2013).
- Morton, J. T. *et al.* Balance trees reveal microbial niche differentiation. *mSystems* **2**, e00162–16 (2017).
- Gloor, G. B., Wu, J. R., Pawlowsky-Glahn, V. & Egozcue, J. J. It's all relative: analyzing microbiome data as compositions. *Ann. Epidemiol.* **26**, 322–329 (2016).
- Harmesen, H. J. M., Pouwels, S. D., Funke, A., Bos, N. A. & Dijkstra, G. Crohn's disease patients have more IgG-binding fecal bacteria than controls. *Clin. Vaccine Immunol.* **19**, 515–521 (2012).
- Prots, R. *et al.* Absolute quantification of microbial taxon abundances. *ISME J.* **11**, 584–587 (2017).
- Stamm, F. *et al.* Adjusting microbiome profiles for differences in microbial load by spike-in bacteria. *Microbiome* **4**, 28 (2016).
- Sabino, J. *et al.* Primary sclerosing cholangitis is characterised by intestinal dysbiosis independent from IBD. *Gut* **65**, 1681–1689 (2016).
- Holmes, I., Harris, K. & Quince, C. Dirichlet multinomial mixtures: generative models for microbial metagenomics. *PLoS One* **7**, e30126 (2012).
- Arumugam, M. *et al.* Enterotypes of the human gut microbiome. *Nature* **473**, 174–180 (2011).
- Lewis, S. J. & Heaton, K. W. Stool form scale as a useful guide to intestinal transit time. *Scand. J. Gastroenterol.* **32**, 920–924 (1997).
- Vandeputte, D., Falony, G., D'hoel, K., Vieira-Silva, S. & Raes, J. Water activity does not shape the microbiota in the human colon. *Gut* **66**, 1865–1866 (2017).
- Vandeputte, D. *et al.* Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. *Gut* **65**, 57–62 (2016).
- Tighe, E. F. *et al.* Gut microbiota composition associated with stool consistency. *Gut* **65**, 540–542 (2016).
- Fuller, B. J. Cryoprotectants: the essential antifreezes to protect life in the frozen state. *Cryo Lett.* **25**, 375–388 (2004).
- Hugon, P. *et al.* Molecular studies neglect apparently gram-negative populations in the human gut microbiota. *J. Clin. Microbiol.* **51**, 3286–3293 (2013).
- Ben-Amor, K. *et al.* Genetic diversity of viable, injured, and dead fecal bacteria assessed by fluorescence-activated cell sorting and 16S rRNA gene analysis. *Appl. Environ. Microbiol.* **71**, 4679–4689 (2005).
- Roager, H. M. *et al.* Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. *Nat. Microbiol.* **1**, 16093 (2016).
- McMurdie, P. J. & Holmes, S. Waste not, want not: why rarefying microbiome data is inadmissible. *PLOS Comput. Biol.* **10**, e1003531 (2014).
- Goodrich, J. K. *et al.* Conducting a microbiome study. *Cell* **158**, 250–262 (2014).
- Magurran, A. E. *Measuring Biological Diversity* (Blackwell, 2004).
- Faust, K. & Raes, J. Microbial interactions: from networks to models. *Nat. Rev. Microbiol.* **10**, 538–550 (2012).

25. Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K. & Knight, R. Diversity, stability and resilience of the human gut microbiota. *Nature* **489**, 220–230 (2012).
26. Pascal, V. *et al.* A microbial signature for Crohn's disease. *Gut* **66**, 813–822 (2017).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank all study participants, F. Giraldo for enabling sample collection at the PXL Hasselt, L. Rymenans and C. Verspecht for faecal DNA extraction and library preparation, K. Verbeke for facilitating moisture content determinations, and P. Goncalves for advice on simulating microbial data for benchmarking the QMP and RMP approach. The main funding for this study comes from a KU Leuven CREA grant. D.V. is supported by the Agency for Innovation by Science and Technology (IWT). G.K., K.D., M.V.-C., S.V.-S., and J.W. are funded by the Research Foundation Flanders (FWO-Vlaanderen). This work is further supported through funding by VIB, the Rega Institute for Medical Research, KU Leuven, FP7 METACARDIS (HEALTH-F4-2012-305312), and H2020 SYSCID (grant agreement 733100).

Author Contributions This study was conceived by G.F. Experiments were designed by D.V., S.V., G.F., and J.R. Sampling of cohorts was set up and carried out by D.V., G.K., K.D., S.V.-S., M.V.-C., J.S., J.W., R.Y.T., L.D.C., and G.F. Optimization of sequencing protocols was performed by R.Y.T.; data pre-processing by D.V., M.V.-C., J.S., J.W., and Y.D.; flow cytometry analyses by G.K. and K.D.; statistical analyses by D.V., G.K., K.D., S.V.-S., M.V.-C., J.S., J.W., and G.F.; network analyses by S.V.-S.; and simulation experiments by D.V. and S.V.-S. G.F. developed the QMP protocol. S.V.-S., G.F., and J.R. drafted the manuscript. All authors revised the article and approved the final version for publication.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to J.R. (jeroen.raes@kuleuven.be).

Reviewer Information *Nature* thanks W. M. de Vos and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Sample collection. *Ethical compliance.* All experimental protocols were approved by the Commissie Medische Ethiek, UZ KU Leuven. Study design complied with all relevant ethical regulations, aligning with the Declaration of Helsinki and in accordance with Belgian privacy laws. All participants gave their informed consent. *Study cohort.* The study cohort consisted of forty volunteers recruited among the staff of the KU Leuven research facilities. Ethical approval of the study protocol was obtained (B322201629015). No inclusion or exclusion criteria were imposed. A limited set of anthropometric metadata was compiled at enrolment, including gender, age, height, and weight (Supplementary Table 1). Participants were asked to collect a maximal amount of faeces (single defecation) in a plastic receptacle with lid and to deposit the sample in a labelled non-transparent zip-lock bag at the research facility immediately after defecation. On receiving the sample, the research team scored stool consistency using the BSS score¹³. All faecal samples were processed within an hour of the reported time of egestion.

Longitudinal and validation cohort. To assess longitudinal variation in faecal cell counts, we analysed faecal samples provided by 20 healthy individuals (10 women and 10 men) over a one-week sampling effort (B322201524875; Supplementary Table 1). Participants were asked to collect a faecal sample at each defecation, with a maximum of one sample per day. Non-microbiome findings were replicated in a validation cohort of fifty-four volunteers recruited among staff and students of the Hogeschool PXL (Supplementary Table 1). Sampling protocols were aligned with FGFP procedures. No inclusion or exclusion criteria were imposed. For both the validation dataset (duplicate) and the longitudinal cohort (single), faecal cell counts and moisture content were determined starting from frozen aliquots of non-homogenized faecal material.

Disease cohort. The full description of the sub-cohort of patients with Crohn's disease has previously been published¹⁰. The cohort comprised 29 patients with Crohn's disease and 66 healthy controls who had been selected from the FGFP database². Microbiota phylogenetic profiling was performed as described below for the study cohort. However, for the disease cohort, the V4 region of the 16S rRNA gene was amplified with the primer pair 515F and 806R with single multiplex identifier (MID) and adaptors, in accordance with previously published methods²⁷. Cell counts (triplicate) were determined using frozen aliquots of non-homogenized faecal material.

Cell counts and stool moisture. After mechanical homogenization (5 min, 150 r.p.m.; Stomacher 3500), fresh study cohort faecal samples were divided into 0.2-g aliquots. Two aliquots were used immediately for cell count determinations and the remaining ones were frozen (−80 °C) for further analyses (including cell counts of frozen samples, in triplicate). For cell counting, 0.2-g aliquots were diluted 100,000 times in physiological solution (8.5 g l^{−1} NaCl; VWR International). In order to remove debris from the faecal solutions, samples were filtered using a sterile syringe filter (pore size 5 µm; Sartorius Stedim Biotech GmbH). Next, 1 ml of the microbial cell suspension obtained was stained with 1 µl SYBR Green I (1:100 dilution in dimethylsulfoxide; shaded 15 min incubation at 37 °C; 10,000 concentrate, Thermo Fisher Scientific). The flow cytometry analysis of the microbial cells present in the suspension was performed using a C6 Accuri flow cytometer (BD Biosciences), according to previously published methods²⁸. Fluorescence events were monitored using the FL1 533/30 nm and FL3 >670 nm optical detectors. Forward and sideways-scattered light was also collected. The BD Accuri CFlow software was used to gate and separate the microbial fluorescence events on the FL1–FL3 density plot from the faecal sample background. A threshold value of 2,000 was applied on the FL1 channel. The gated fluorescence events were evaluated on the forward–sideways density plot, to exclude remaining background events and to obtain an accurate microbial cell count. Instrument and gating settings were identical for all samples (fixed staining–gating strategy²⁸; Extended Data Fig. 10). Stool moisture content was determined in duplicate on 0.2 g frozen homogenized faecal material (−80 °C) as the percentage of mass loss after lyophilization.

qPCR assessment of bacterial loads. Frozen aliquots of homogenized faeces were thawed in physiological solution. Next, DNA was extracted directly from a volume corresponding to 4 mg of faecal material using the MoBio PowerMicrobiome RNA isolation kit with an additional incubation step of 10 min at 95 °C to maximize lysis. Extracted DNA was quantified using a Qubit 2.0 Fluorometer (Thermo Fisher Scientific) and served as a template for qPCR amplification of bacterial 16S rRNA genes (primer pair Uni16SF and Uni16SR; CCATGAAGTCGGAATCGCTAG and GCTTGACGGGCGGTGT, respectively²⁹) using an Applied Biosystems 7500 Fast Real-Time PCR System (Thermo Fisher Scientific). PCR assay mixtures consisted of 10 µl PowerUp SYBR Green Master Mix (Thermo Fisher Scientific), two times 2 µl primer solution (300 nM), 4 µl sterile nuclease-free water, and 2 µl template DNA

solution. The PCR amplification program encompassed two initial denaturation steps at 50 °C and 95 °C for 120 s, followed by 40 two-step cycles at 95 °C for 3 s and at 60 °C for 30 s. In each run, negative (without DNA), extraction (without pellet), and positive controls (with extracted genomic DNA from *Faecalibacterium prausnitzii* DSM 17677¹) were included. At the end of each run, a melting curve analysis was performed. Cycle threshold values were determined using the 7500 software v2.0.6 (Thermo Fisher Scientific). All qPCR assays were performed in triplicate. Finally, cycle threshold values were normalized using an inter-plate calibrator in order to account for differences among qPCR runs.

Microbiota phylogenetic profiling. *Sequencing data pre-processing.* Faecal microbiome profiling of the study cohort was performed as previously described². In brief, DNA was extracted from faecal material using the MoBio PowerMicrobiome RNA isolation kit. The V4 region of the 16S rRNA gene was amplified with the primer pair 515F and 806R (GTGYCAGCMGCCGCGGTAA and GGACTACNVGGGTWTCTAAT, respectively), modified to contain a barcode sequence between each primer and the Illumina adaptor sequences to produce dual-barcoded libraries³⁰. Sequencing was performed on the Illumina MiSeq platform (MiSeq Reagent Kit v2, 500 cycles, 20% PhiX) according to the manufacturer's specifications to generate paired-end reads of 250 bases in length in each direction. After de-multiplexing with Je³¹ (strict mode, zero mismatch), fastq sequences were merged using FLASH³² software with default parameters, except for min-overlap and max-overlap, which were set to 140 and 230, respectively. Successfully combined reads were filtered on the basis of quality using seqtk trimfq with default parameters (<https://github.com/lh3/seqtk>). This procedure removed between 15% and 32% of the total number of reads. The remaining reads varied between 22426 and 79902 reads per sample (study cohort). Chimeras were removed with the uchime2_ref algorithm of USEARCH (version 9.2.64)³³. Pre-processing of the disease cohort samples was performed as described¹⁰.

Relative microbiome profiling (RMP). For relative microbiome analyses, each sample depth was downsized to 10,000 reads by randomly selecting reads. The taxonomy of reads was assigned using RDP classifier 2.12³⁴.

Quantitative microbiome profiling (QMP). For quantitative microbiome analyses, samples were downsized to an even sampling depth, defined as the ratio between sample size (16S rRNA gene copy-number-corrected sequencing depth) and microbial load (average total cell count per gram of frozen faecal material; Supplementary Table 5). 16S rRNA gene copy numbers were retrieved from the ribosomal RNA operon copy number database rrnDB³⁵. The copy-number-corrected sequencing depth of each sample was rarefied to the level necessary to equal the minimum observed sampling depth in the cohort.

Statistical analyses. No statistical methods were used to predetermine sample size. Statistical analyses were performed in R using the packages phyloseq³⁶, vegan³⁷, FSA³⁸, coin³⁹, DirichletMultinomial⁴⁰, ICC⁴¹, and fitdistrplus⁴². All statistical tests used were two-sided.

Faecal microbiome derived features and visualization. Observed richness was calculated with the R package phyloseq³⁶. Enterotyping (or community typing) using the DMM approach was performed in R as previously described¹¹. To increase accuracy, enterotyping was performed on a combined genus–abundance matrix that included study and disease cohort samples, complemented with 1,106 samples from the FGFP². Microbiome variation between individuals was visualized by PCoA using Bray–Curtis dissimilarity on the genus–abundance matrix. Metadata was fitted on the PCoA ordination using the vegan package in R. *Features and metadata associations.* Taxa unclassified at the genus level or present in less than 20% of the samples were excluded from the statistical analyses. Correlations between continuous variables (genus abundances or metadata) were analysed using non-parametric Spearman tests. Associations with two-level categorical metadata (gender) were performed using Wilcoxon rank-sum tests. Corrections for multiple testing (Benjamini–Hochberg method⁴³, FDR) were performed where applicable. Metadata differences between enterotypes (four-level categorical data) were assessed using non-parametric ANOVA (Kruskal–Wallis) and post hoc Dunn's test for all pairs of comparisons between groups, with Benjamini–Hochberg adjustment for multiple testing (adjusted *P* value).

Intraclass rank correlation coefficient. The ICC was calculated for the rank-transformed longitudinal dataset, defining each individual as a distinct class, using a one-way ANOVA fixed effects model. The ICC estimation uses the variance components from a one-way ANOVA (among-group variance and within-group variance; $ICC = \text{var}_{\text{among}} / (\text{var}_{\text{among}} + \text{var}_{\text{within}})$).

Ordinal association between genera in RMP and QMP matrices. Kendall rank correlation coefficient (assessing the degree to which samples maintain the same relative position to each other; τ_B test) was used to test the concordance of ranking for each genus between the RMP and the QMP matrices.

Estimation of the community variation explained by cumulative contribution of metadata variables. Variation partitioning by stepwise dbRDA was performed to

discern the degree to which variation in microbial community profiles (Bray–Curtis dissimilarity from RMP or QMP) could be explained by the cumulative contribution of metadata variables.

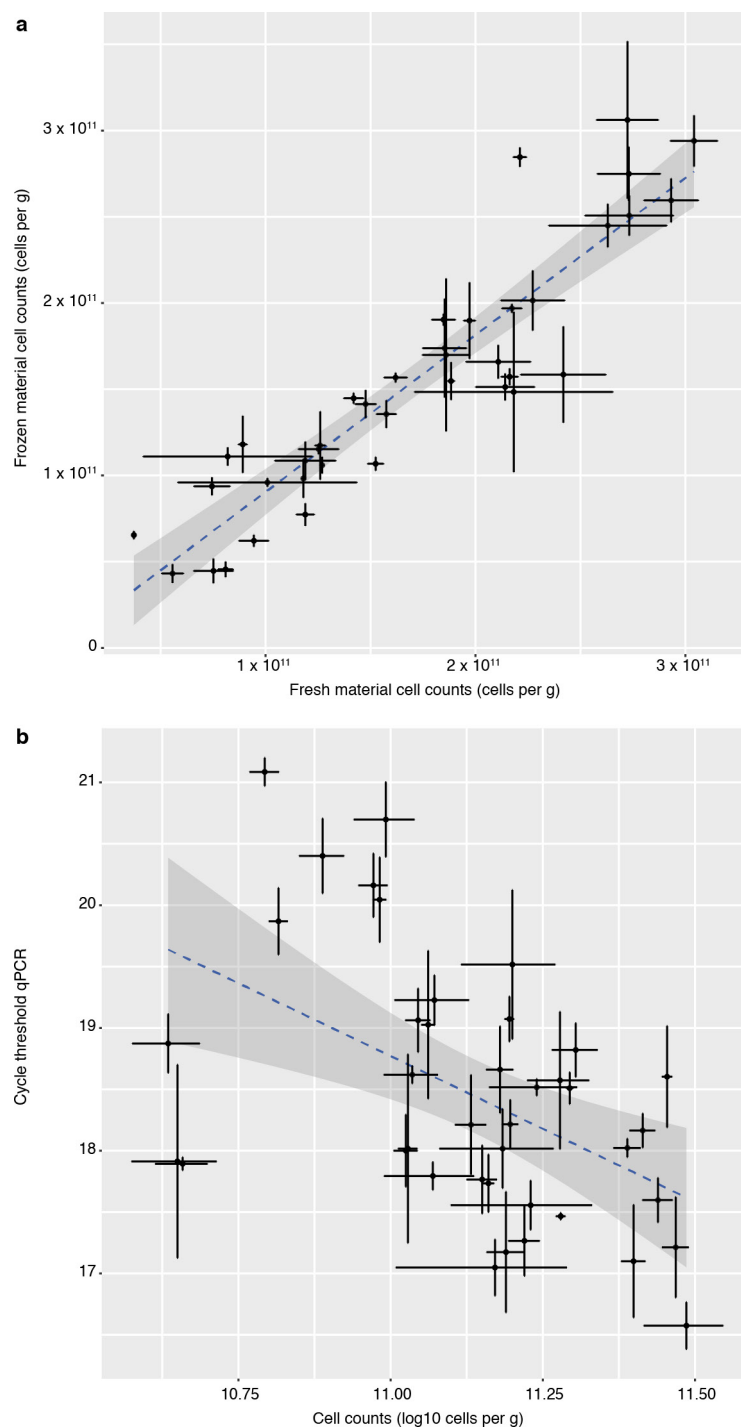
Reconstruction of co-abundance networks in RMP and QMP matrices. Taxon–taxon associations (edges) were defined by Spearman correlations between pairs of taxa with multiple testing correction (Bonferroni). Taxa were pre-filtered to exclude taxa unclassified at the genus level and taxa that were present in less than 50% of the samples.

Benchmarking the QMP and RMP approach for co-abundance network reconstruction using simulated RMP and QMP data. Microbial count data matrices (simulated real data) were simulated for three ranges of microbial load difference between samples: twofold, fourfold, and eightfold maximum microbial load difference ($\pm 20\%$). Two hundred and fifty taxa abundances were drawn for 100 samples from negative binomial distributions simulated using parameters (n, p) derived from the study dataset using the R package *fitdistrplus*. Taxon–taxon associations (Spearman's $\rho = 0.25–0.35$) were induced for 20 pairs of taxa present in at least 50% of the samples. Next, simulated real data matrices were rarefied to 80% of the lowest sample total count to obtain simulated RMP matrices using the R package *phyloseq*. Simulated QMP matrices were then obtained by multiplying simulated RMP matrices by the original microbial loads of the simulated real data. Taxon–taxon associations were assessed as in the study dataset (Spearman's correlation in taxa present in at least 50% of the samples). Confusion matrices for the simulated RMP and QMP data were calculated by comparing significant ($P < 0.05$) correlations to expectations from the simulated real data. Associations detected in real simulated data were defined as positives (P). False positive (FP) associations were identified as present in simulated RMP and QMP, but as absent in simulated real data. Likewise, true positive (TP) correlations were defined as present in both simulated real data and simulated RMP and QMP. For a given RMP–QMP matrix, the false discovery rate was calculated as $FP/(FP + TP)$. The true positive rate (sensitivity) was determined by TP/P . Reported values were derived from 50 iterations. Differences in false discovery rate and sensitivity between the QMP and RMP approaches were assessed using a paired *t*-test.

Code availability. An open source QMP R-script is available on <http://www.raeslab.org/software/QMP> and the Github repository (<https://github.com/raeslab/QMP>).

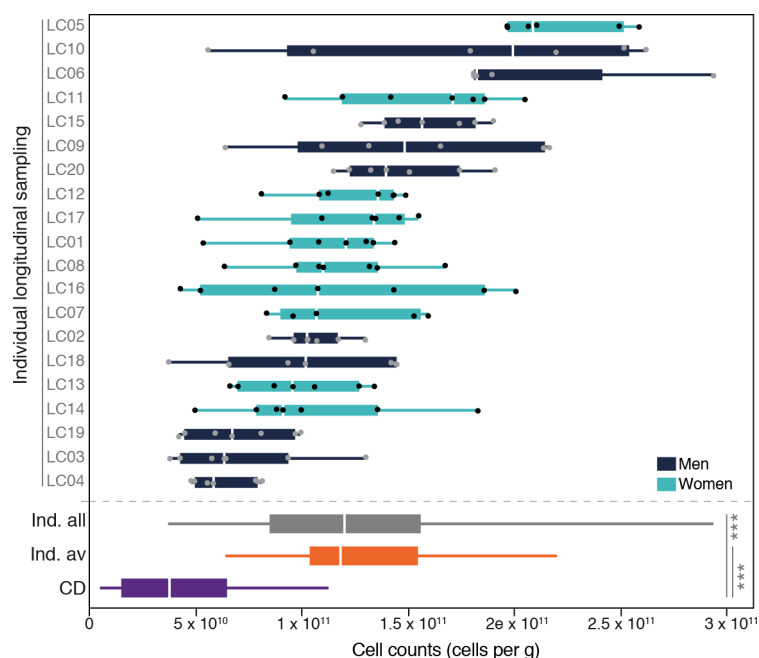
Data availability. Raw amplicon sequencing data that support the findings of this study have been deposited in European Nucleotide Archive with accession codes PRJEB21504 and ERP023761. Source Data for all figures (with the exception of Extended Data Figs 6 and 10) are provided with the paper. All other data are available from the corresponding author upon reasonable request.

27. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* **79**, 5112–5120 (2013).
28. Prest, E. I., Hammes, F., Köttsch, S., van Loosdrecht, M. C. M. & Vrouwenvelder, J. S. Monitoring microbiological changes in drinking water systems using a fast and reproducible flow cytometric method. *Water Res.* **47**, 7131–7142 (2013).
29. Ramseier, C. A. *et al.* Identification of pathogen and host-response markers correlated with periodontal disease. *J. Periodontol.* **80**, 436–446 (2009).
30. Tito, R. Y. *et al.* *Dialister* as a microbial marker of disease activity in spondyloarthritis. *Arthritis Rheumatol.* **69**, 114–121 (2017).
31. Girardot, C., Scholtalbers, J., Sauer, S., Su, S. Y. & Furlong, E. E. Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. *BMC Bioinformatics* **17**, 419 (2016).
32. Magoč, T. & Salzberg, S. L. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–2963 (2011).
33. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. & Knight, R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**, 2194–2200 (2011).
34. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
35. Stoddard, S. F., Smith, B. J., Hein, R., Roller, B. R. K. & Schmidt, T. M. rrnDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Res.* **43**, D593–D598 (2015).
36. McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**, e61217 (2013).
37. Oksanen, J. *et al.* *vegan*: community ecology package. R package version 2.2–1 <https://CRAN.R-project.org/package=vegan> (2015).
38. Ogle, D. H. FSA: fisheries stock analysis. R package version 0.8.13 <https://cran.r-project.org/package=FSA> (2017).
39. Hothorn, T., Hornik, K., van de Wiel, M. A. & Zeileis, A. A Lego system for conditional inference. *Am. Stat.* **60**, 257–263 (2006).
40. Morgan, M. DirichletMultinomial: Dirichlet-multinomial mixture model machine learning for microbiome data. R package version 1.18.0 <https://cran.r-project.org/package=dirmult> (2017).
41. Wolak, M. ICC: facilitating estimation of the intraclass correlation coefficient. R package version 2.3.0 <https://cran.r-project.org/package=ICC> (2016).
42. Delignette-Muller, M. L. & Dutang, C. fitdistrplus: an R package for fitting distributions. *J. Stat. Softw.* **64**, 1–34 (2015).
43. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
44. Vieira-Silva, S. *et al.* Species–function relationships shape ecological properties of the human gut microbiome. *Nat. Microbiol.* **1**, 16088 (2016).
45. Smith, C. J. & Osborn, A. M. Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. *FEMS Microbiol. Ecol.* **67**, 6–20 (2009).
46. Hammes, F. *et al.* Flow-cytometric total bacterial cell counts as a descriptive microbiological parameter for drinking water treatment processes. *Water Res.* **42**, 269–277 (2008).
47. Habtewold, T., Duchateau, L. & Christophides, G. K. Flow cytometry analysis of the microbiota associated with the midguts of vector mosquitoes. *Parasit. Vectors* **9**, 167 (2016).

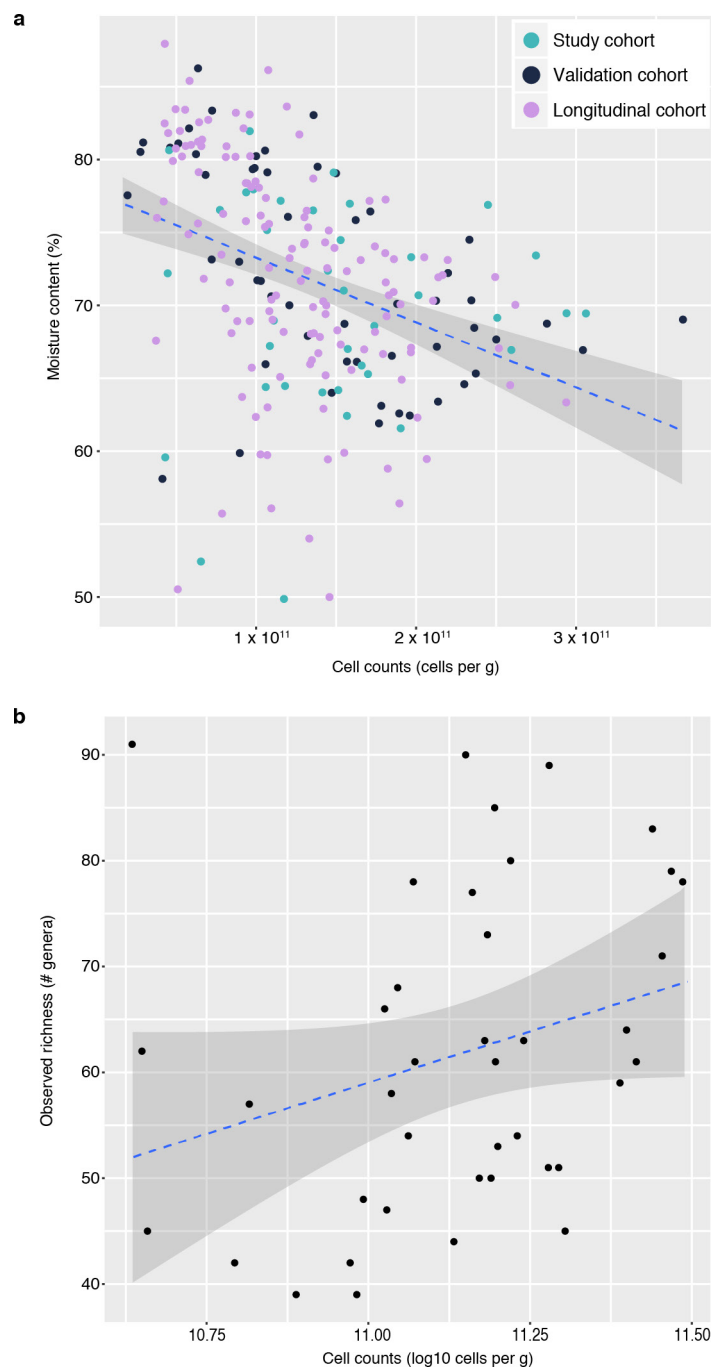


Extended Data Figure 1 | Quantification of microbial loads of frozen faecal samples. a, Microbial cell counts in fresh and frozen faecal aliquots strongly correlate with one another (study cohort, $n = 39$; Pearson's $r = 0.91$, two-sided $P = 4.9 \times 10^{-16}$). Data points represent mean values. Error bars represent the s.d. of duplicate (fresh) and triplicate (frozen) cell counts. **b,** Comparison between flow cytometric assessment of microbial loads and estimation of bacterial abundances on the basis of qPCR (study cohort, $n = 40$). Data points represent mean values. Error bars represent the s.d. of triplicate values (qPCR and cell counts). Although comparing cell-based and molecular enumeration methods is not recommended¹⁸, the measurements were correlated (Pearson's $r = -0.53$,

two-sided $P = 4.7 \times 10^{-4}$). In a complex ecosystem, enumerating bacteria on the basis of qPCR would introduce biases through the extraction, purification, and amplification of DNA, 16S rRNA gene copy number variation, and community replication rate (shown to differ between enterotypes⁴⁴), among others. Moreover, qPCR has been reported to be only sensitive enough to detect twofold changes in gene concentration or microbial load⁴⁵. Flow cytometry is less specific and results might be affected by formation of aggregates⁸. For QMP analyses, we opted for a flow cytometry approach given its technical straightforwardness (limited number of technical manipulations; see Methods), reproducibility^{46,47}, and throughput.

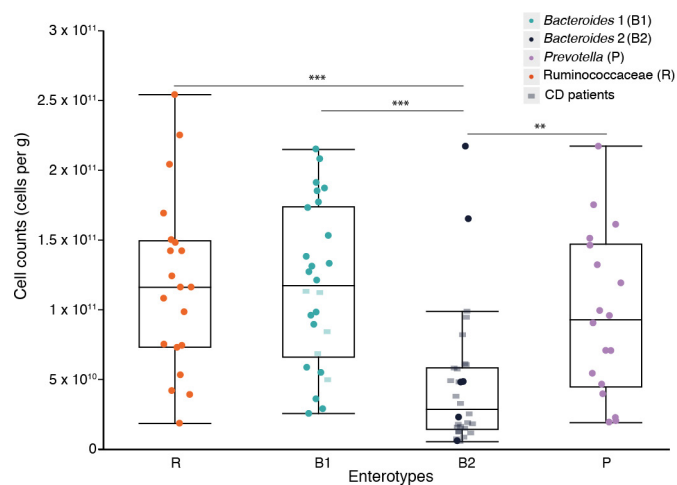


Extended Data Figure 2 | Intra-individual versus inter-individual microbial cell count variation in human faeces. Twenty healthy individuals sampled daily over the course of a week (maximum of one sample per day; longitudinal cohort). Healthy individuals have significantly higher cell counts than patients with Crohn's disease. Patient group disease cohort, $n = 29$ (CD; purple) versus overall daily samples, $n = 132$ (Ind. All; grey) or the average cell count of individuals (Ind. av; orange). Two-sided Wilcoxon rank-sum test, $***P < 0.001$. The body of the box plot represents the first and third quartiles of the distribution and the median line. The whiskers extend from the quartiles to the last data point within $1.5 \times$ interquartile range, with outliers beyond.

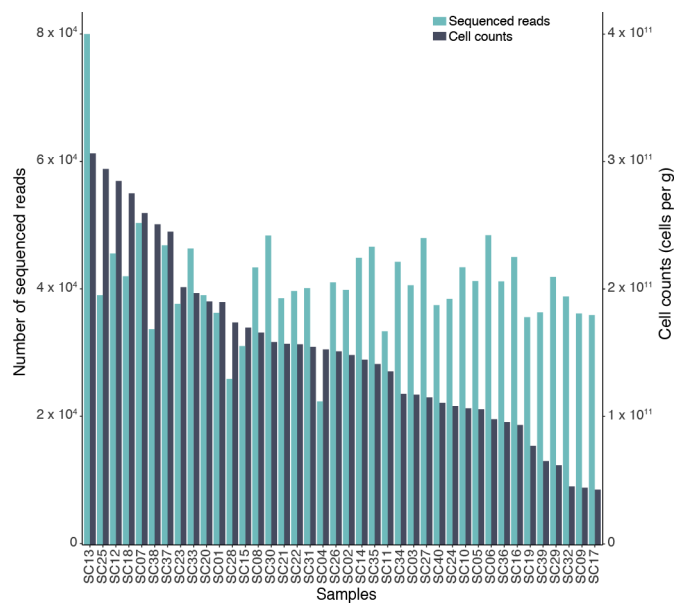


Extended Data Figure 3 | Faecal microbial loads correlate with sample moisture content and genus richness. **a**, Microbial cell counts and moisture content were negatively correlated in the longitudinal and the validation cohort, though not in the study cohort (study cohort, $n = 37$, Spearman's $\rho = -0.12$, two-sided FDR = 0.56; longitudinal cohort, $n = 132$,

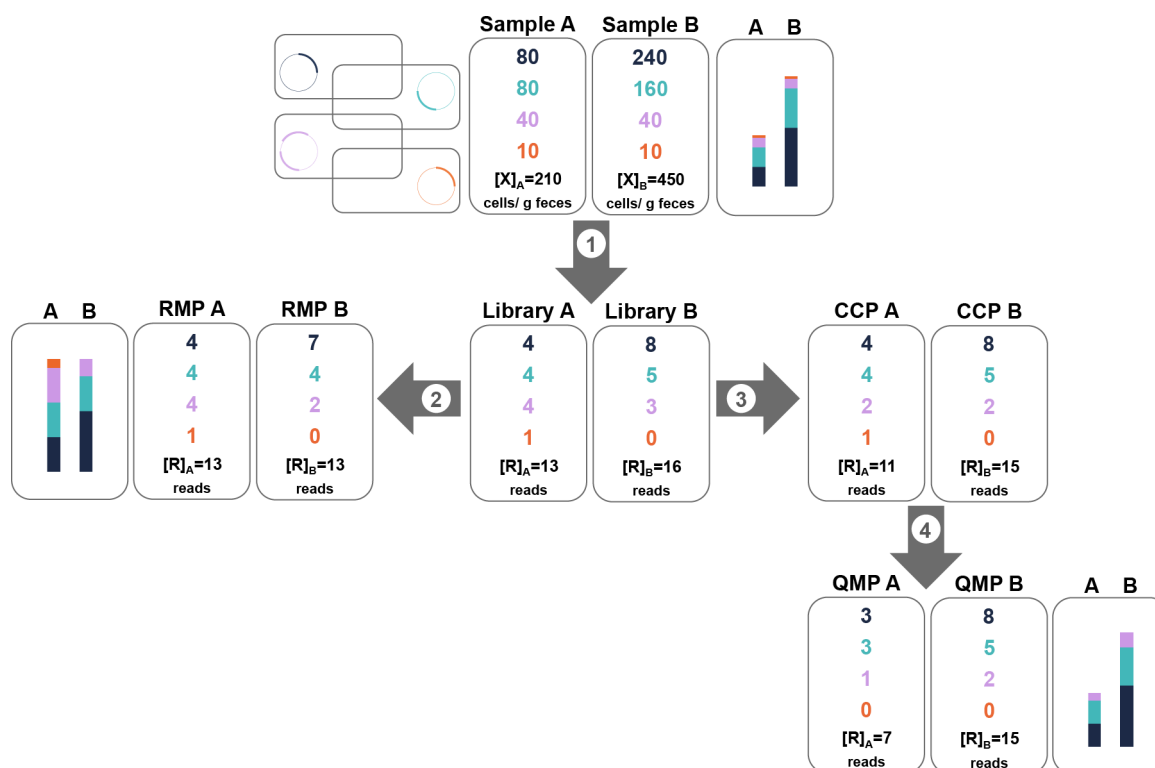
Spearman's $\rho = -0.52$, two-sided FDR = 1.6×10^{-9} ; validation cohort, $n = 54$, Spearman's $\rho = -0.56$, two-sided FDR = 9.1×10^{-5}). **b**, Microbial cell counts and observed richness correlated mildly (study cohort, $n = 40$; Spearman's $\rho = 0.36$, two-sided $P = 2.3 \times 10^{-2}$).



Extended Data Figure 4 | Faecal microbial loads vary across enterotypes. Microbial load differences between the four enterotypes in the disease cohort ($n = 95$). Box plot representation of microbial load (cells per gram of faeces) distribution across the four enterotypes. The body of the box plot represents the first and third quartiles of the distribution and the median line. The whiskers extend from the quartiles to the last data point within $1.5 \times$ interquartile range, with outliers beyond. Two-sided Dunn's adjusted test, $**P < 0.01$, $***P < 0.001$.

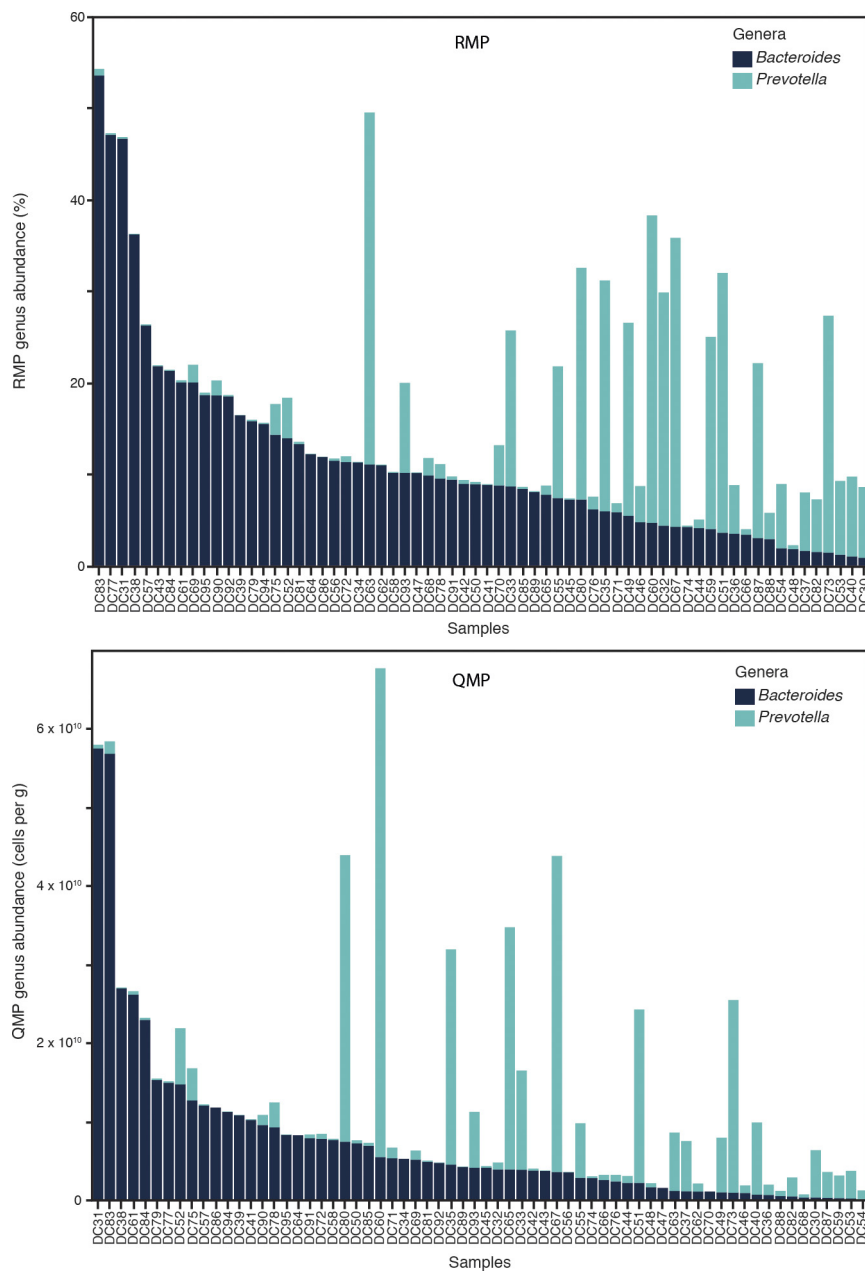


Extended Data Figure 5 | Microbial loads are not associated with sequencing depth. Sequencing depth did not reflect the total microbial load of a sample (Spearman's $\rho = 0.17$, two-sided $P = 0.28$). Samples are ranked according to decreasing cell counts.



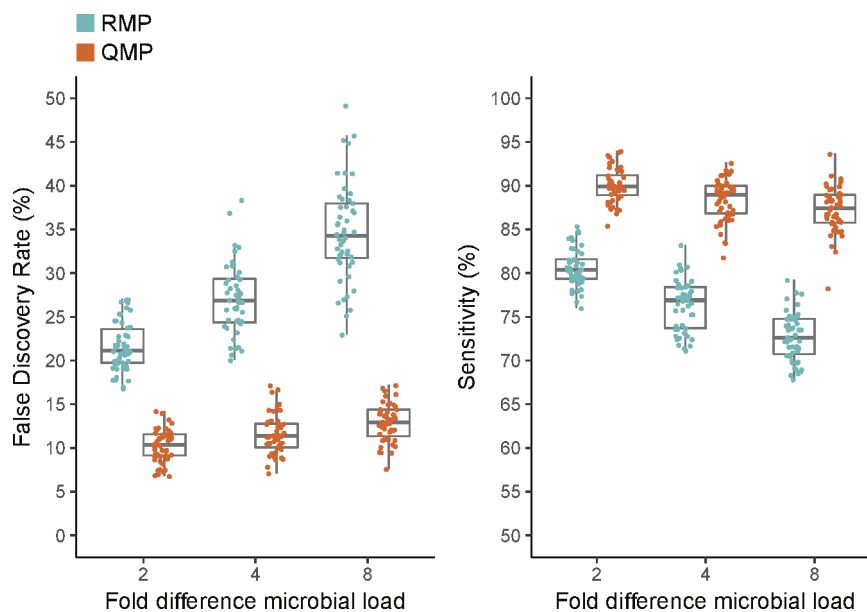
Extended Data Figure 6 | Illustration of differences between RMP and QMP methodology. Two samples, each containing four genera, are analysed (numbers are illustrative). Genus abundance distributions in sample A and B are markedly distinct, with the microbial load in sample B more than double that of the load in sample A. Genus 'purple' carries two copies of the 16S rRNA gene. (1) DNA extraction and library preparation. Neither RMP nor QMP correct for biases introduced by DNA extraction, primer specificity, PCR amplification, or other library preparation steps. The resulting sequencing depth is independent of microbial load. (2) By rarefying to an even number of reads per sample, RMP assumes similar genus abundance distributions in samples A and B: sample A is therefore sequenced far more intensively than sample B. The resulting profiles therefore poorly reflect the genus distribution in the original samples. Given the multiple copies of the 16S rRNA gene, the relative abundance of

'purple' is overestimated. (3) The first step of QMP corrects for 16S rRNA copy number variation. In the resulting copy-corrected profile (CCP), each read corresponds with a single bacterium sequenced. (4) By dividing the CCP reads total (R) by the microbial loads (X), sampling depth is estimated for each sample. For sample A and B, sampling depth is $[R]_A$ divided by $[X]_A$ and $[R]_B$ divided by $[X]_B$, respectively. The sampling depth for B is the lowest (3.33%) of the two; sample A is rarefied to the same level. This implies that 'orange' is no longer detected. As 'orange' was equally abundant in A and B, the fact that it is included in sample A RMP can be considered an artefact of uneven sampling intensity. The resulting rarefied genus abundances are proportional with sample microbial loads and can be extrapolated to generate QMPs expressed as number of cells per gram.



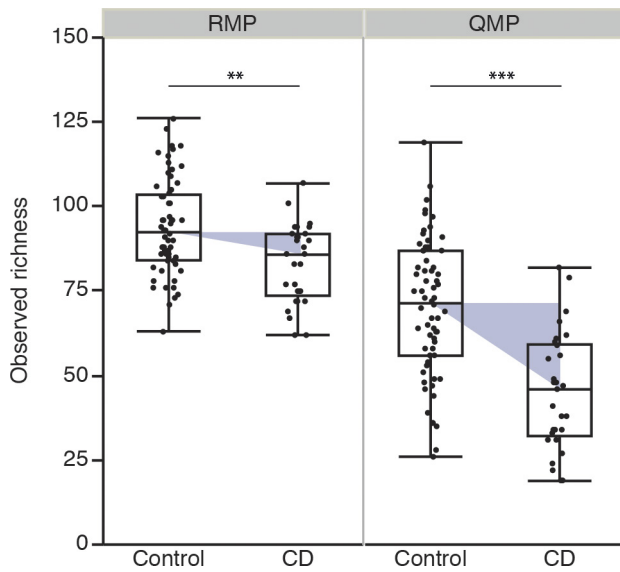
Extended Data Figure 7 | Distribution of RMP and QMP abundances of *Bacteroides* and *Prevotella* in healthy controls. Samples ($n = 66$) are ranked according to decreasing *Bacteroides* abundance in both the RMP and QMP panel (stacked bars). The trade-off between *Bacteroides* and

Prevotella (RMP; Spearman's $\rho = -0.59$, two-sided FDR = 2×10^{-4}) was no longer significant after correction for microbial load (QMP; Spearman's $\rho = -0.33$, two-sided FDR = 1).

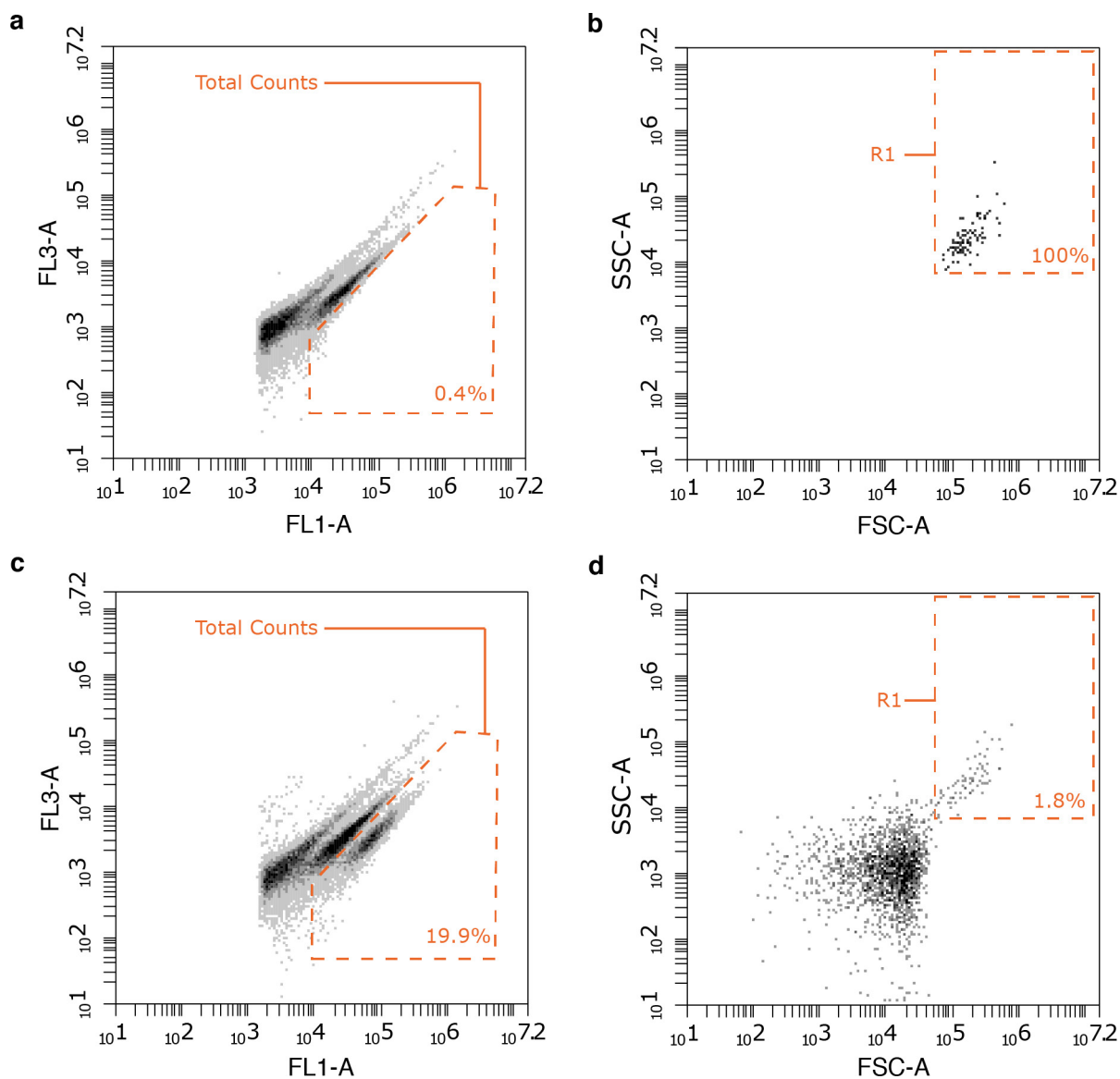


Extended Data Figure 8 | False discovery rate and sensitivity in network reconstruction based on QMP and RMP simulated data. QMP resulted in increased sensitivity and decreased FDR compared to RMP (two-sided t -test, $P < 10^{-15}$). For a two-, four- and eightfold maximum difference in microbial load, QMP FDRs were 11%, 15%, and 22% lower, respectively,

than for RMP; and QMP true positive discovery rate (sensitivity) was increased by 10%, 12%, and 15%. Data points depict repetitions ($n = 50$). The body of the box plot represents the first and third quartiles of the distribution and the median line. The whiskers extend from the quartiles to the last simulated data point within $1.5 \times$ interquartile range.



Extended Data Figure 9 | RMP underestimates the decrease in microbiome richness associated with Crohn's disease. Observed richness in healthy controls ($n = 66$) and patients with Crohn's disease ($n = 29$). The decrease in richness associated with Crohn's disease is more pronounced in QMP. The body of the box plot represents the first and third quartiles of the distribution and the median line. The whiskers extend from the quartiles to the last data point within $1.5 \times$ interquartile range. Two-sided Wilcoxon test, *** $P < 0.001$, ** $P < 0.01$.



Extended Data Figure 10 | Flow cytometry gating strategy. A fixed gating and staining approach was applied²⁸. Both blank and sample solutions were stained with SYBR Green I. **a**, The FL1-A/FL3-A acquisition plot of a blank sample (0.85% w/v physiological solution) with gate boundaries indicated. A threshold value of 2,000 was applied on the FL1 channel. **b**, Secondary gating was performed on the FSC-A/SSC-A channels to further discriminate between debris or background and

microbial events. **c**, **d**, FL1-A/FL3-A count acquisition of a faecal sample (**c**) with secondary gating on FSC-A/SSC-A channels on the basis of blank analyses (**d**). Total counts were defined as events registered in the FL1-A/FL3-A gating area, excluding debris or background events observed in the FSC-A/SSC-A R1 gate. The flow rate was set at 14 μ l per minute and the acquisition rate did not exceed 10,000 events per second. Each panel reflects events registered over the course of a 30-s acquisition period.

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

► Experimental design

1. Sample size

Describe how sample size was determined.

No sample size calculation was performed. Cohort sizes were determined based on previous findings (Falony et al., Science 2016).

2. Data exclusions

Describe any data exclusions.

To assess microbiome feature and metadata associations, taxa unclassified at the genus level or present in less than 20% of samples were excluded from the statistical analyses. Taxa unclassified at the genus level or present in less than 50% of samples were excluded from co-abundance network reconstruction. The same 50% cut-off was handled in the simulated datasets.

3. Replication

Describe whether the experimental findings were reliably reproduced.

All findings were replicated. All replications were successful and this is reported in the manuscript.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

No experimental groups were defined. The study/longitudinal and validation cohorts were recruited among the staff of the KU Leuven research facilities located at university hospital (Leuven, Belgium) and staff and students of the Hogeschool PXL (Hasselt, Belgium), respectively. No inclusion or exclusion criteria were imposed. For a description of the disease cohort reanalyzed in the study, we refer to the original publication (doi: 10.1136/gutjnl-2015-311004).

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Investigators were not blinded during data collection and analyses.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- ☐ ☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ A statement indicating how many times each experiment was replicated
- ☐ ☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
- ☐ ☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- ☐ ☒ The test results (e.g. P values) given as exact values whenever possible and with confidence intervals noted
- ☐ ☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- ☐ ☒ Clearly defined error bars

See the web collection on [statistics for biologists](#) for further resources and guidance.

► Software

Policy information about [availability of computer code](#)

7. Software

Describe the software used to analyze the data in this study.

An open source QMP R-script is available on <http://www.raeslab.org/software/QMP> and the Github repository (<https://github.com/raeslab/QMP>) with public access.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* [guidance for providing algorithms and software for publication](#) provides further information on this topic.

► Materials and reagents

Policy information about [availability of materials](#)

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

n.a.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

n.a.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

n.a.

b. Describe the method of cell line authentication used.

n.a.

c. Report whether the cell lines were tested for mycoplasma contamination.

n.a.

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

n.a.

► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

n.a.

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

Study cohort: 20 male/20female. Median age: 29; BMI: 22.3; BSS: 4. Longitudinal cohort: 10 male/10 female. Median age: 27, BMI: 23, BSS: 4. Validation cohort: 10 male/44 female. Median age: 38, BMI: 23, BSS: 3.5. Disease cohort: 29 Crohn's disease patients/66 healthy controls (doi: 10.1136/gutjnl-2015-311004).

Flow Cytometry Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

► Data presentation

For all flow cytometry data, confirm that:

- ☒ 1. The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ 2. The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☒ 3. All plots are contour plots with outliers or pseudocolor plots.
- ☒ 4. A numerical value for number of cells or percentage (with statistics) is provided.

► Methodological details

- | | |
|----------------------------------------------------------------------------------------|----------------------------------------|
| 5. Describe the sample preparation. | <div>Reported in method section.</div> |
| 6. Identify the instrument used for data collection. | <div>Reported in method section.</div> |
| 7. Describe the software used to collect and analyze the flow cytometry data. | <div>Reported in method section.</div> |
| 8. Describe the abundance of the relevant cell populations within post-sort fractions. | <div>n. a.</div> |
| 9. Describe the gating strategy used. | <div>Reported in method section.</div> |

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information. ☒