MICROBIOME

Population-level analysis of gut microbiome variation

Gwen Falony,[1,2]* Marie Joossens,[1,2,3]* Sara Vieira-Silva,[1,2]* Jun Wang,[1,2]* Youssef Darzi,[1,2,3] Karoline Faust,[1,2,3] Alexander Kurilshikov,[4,5] Marc Jan Bonder,[6] Mireia Valles-Colomer,[1,2] Doris Vandeputte,[1,2,3] Raul Y. Tito,[1,2,3] Samuel Chaffron,[1,2,3] Leen Rymenans,[1,2,3] Chloë Verspecht,[1,2] Lise De Sutter,[1,2,3] Gipsi Lima-Mendez,[1,2] Kevin D'hoe,[1,2,3] Karl Jonckheere,[2,3] Daniel Homola,[2,3]† Roberto Garcia,[2,3] Ettje F. Tigchelaar,[6,7] Linda Eeckhaudt,[2,3] Jingyuan Fu,[6,8] Liesbet Henckaerts,[1,9] Alexandra Zhernakova,[6,7] Cisca Wijmenga,[6] Jeroen Raes[1,2,3]‡

Fecal microbiome variation in the average, healthy population has remained under-investigated. Here, we analyzed two independent, extensively phenotyped cohorts: the Belgian Flemish Gut Flora Project (FGFP; discovery cohort; *N* = 1106) and the Dutch LifeLines-DEEP study (LLDeep; replication; *N* = 1135). Integration with global data sets (*N* combined = 3948) revealed a 14-genera core microbiota, but the 664 identified genera still underexplore total gut diversity. Sixty-nine clinical and questionnaire-based covariates were found associated to microbiota compositional variation with a 92% replication rate. Stool consistency showed the largest effect size, whereas medication explained largest total variance and interacted with other covariate-microbiota associations. Early-life events such as birth mode were not reflected in adult microbiota composition. Finally, we found that proposed disease marker genera associated to host covariates, urging inclusion of the latter in study design.

Sequencing-based assessment of microbial communities in human fecal material has linked alterations in gut microbiota composition to disease, as well as chronically suboptimal health and well-being (*1*–*3*). The discovery of these associations has stimulated the search for specific microbiome-based biomarkers for a wide range of pathologies (*4*–*9*). However, major challenges still hamper the once assumed imminent translation of microbiome monitoring into diagnostic and clinical practice. One such hurdle is the lack of knowledge about the impact of host and environmental factors on microbiota variation within an average, healthy population. Such information is essential for robust disease marker identification in clinical metagenomics (*10*). To identify and characterize major microbiome-associated variables, the Flemish Gut Flora Project (FGFP) initiated a large-scale cross-sectional fecal sampling effort in a confined geographic region (Flanders, Belgium). FGFP collection protocols combined rigorous sampling logistics, including frozen sample collection and cold chain monitoring, with exhaustive phenotyping through online questionnaires, standardized anamnesis and health assessment by general medical practitioners (GPs), and extended clinical blood profiling (fig. S1). Encompassing an equilibrated range of age, gender, health, and lifestyle, the FGFP cohort is expected to be representative for the average gut microbiota composition in a Western European population (table S1). From this cohort, fecal samples of 1106 individuals (98.5% of Western or Eastern European ethnicity; 96.8% born in Belgium) with time-matched blood and questionnaire data were analyzed. Microbiome phylogenetic profiling was performed using 16*S* ribosomal RNA (rRNA) gene amplicon sequencing. In addition, a Dutch cohort (*N* = 1135, LifeLines-DEEP, LLDeep;

[1]KU Leuven–University of Leuven, Department of Microbiology and Immunology, Leuven, Belgium. [2]VIB, Center for the Biology of Disease, Leuven, Belgium. [3]Vrije Universiteit Brussel, Faculty of Sciences and Bioengineering Sciences, Microbiology Unit, Brussels, Belgium. [4]Institute of Chemical Biology and Fundamental Medicine SB RAS, Novosibirsk, Russia. [5]Novosibirsk State University, Novosibirsk, Russia. [6]University of Groningen, University Medical Center Groningen, Department of Genetics, 9700 RB Groningen, Netherlands. [7]Top Institute Food and Nutrition, Wageningen, Netherlands. [8]University of Groningen, University Medical Center Groningen, Department of Pediatrics, 9700 RB Groningen, Netherlands. [9]KU Leuven–University Hospitals Leuven, Department of General Internal Medicine, Leuven, Belgium.
*These authors contributed equally to this work. †Present address: Computation and Systems Medicine Section, Department of Surgery and Cancer, Imperial College London, London, UK.
‡Corresponding author. Email: jeroen.raes@kuleuven.be

Downloaded from http://science.sciencemag.org/ on March 13, 2019

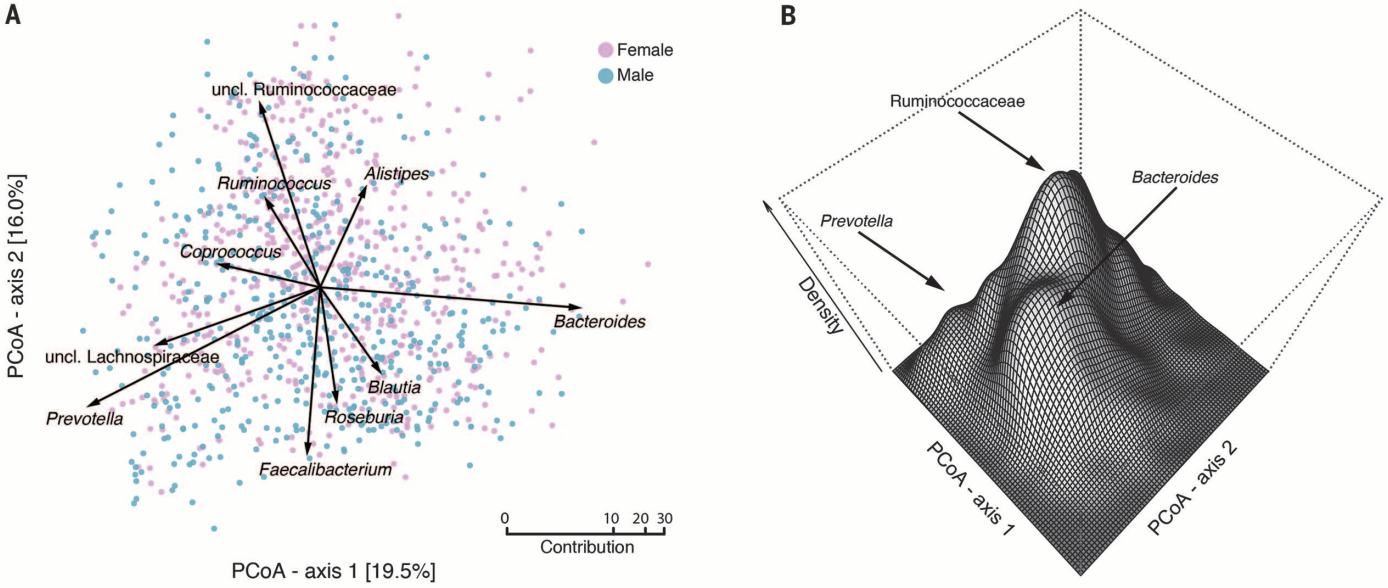

Fig. 1. Microbial community variation in the FGFP cohort, represented by principal coordinates analysis (PCoA, genus-level Bray-Curtis dissimilarity). (**A**) Top 10 contributors to community variation as determined by canonical correspondence analysis on unscaled genera abundances, plotted on the two first PCoA dimensions (arrows scaled to contribution). (**B**) FGFP sample density on the PCoA plot; arrows indicate density peaks enriched in the three previously proposed enterotype drivers: *Prevotella*, *Bacteroides*, and Ruminococcaceae genera.

the Netherlands) was profiled and analyzed (*11*) for validation purposes.

Characterizing the core microbiota

First, we identified a human core microbiota by combining the FGFP and LLDeep data with other U.K. and U.S. studies (*12–14*), yielding nearly 4000 well-profiled individuals. Combined, these data sets comprised a total richness of 664 genera (fig. S2A). Extrapolation estimated total western genus richness at 784 ± 40 (fig. S2B), suggesting that total western richness is still undersampled. Observing total richness would require sampling an estimated additional 40,739 individuals. The current data set yielded a core microbiota (i.e., the genera shared by 95% of samples) composed of 17 genera with a median core abundance (MA) of 72.20% (fig. S2, C and D, and table S2). Complementing this data set with 308 samples collected in Papua New Guinea (*15*), Peru (*16*), and Tanzania (*17*) reduced the size of the human core microbiota to 14 genera. Notably, *Alistipes*, *Clostridium IV*, *Parabacteroides*, and all Actinobacteria were excluded from the global core composition (fig. S3 and table S2). Within the FGFP data set specifically, 35 genera meet the core definition proposed (MA 90.40%), while a 99% cutoff reduced core composition to 20 genera (MA 80.67%; table S2). These 20 core genera also occurred among the top 33 most abundant taxa in the FGFP cohort (table S2). Independently of gender, genus richness correlated positively with age, whereas total core abundance decreased (fig. S4).

Based on unconstrained canonical correspondence analysis of genus-level community composition, we identified the main genera contributing to microbiome variation within the FGFP data set (table S3). Interindividual variation in microbiota composition mainly resulted from changes in relative abundance of core taxa (Fig. 1A). The taxa showing the largest variation in abundance were Ruminococcaceae, *Bacteroides*, and *Prevotella*; all previously proposed as enterotype identifiers (*18*). However, microbiota variation was not only defined by fluctuations in the core or dominant microbiota members, as less abundant genera, such as *Akkermansia* and *Methanobrevibacter*, were also discriminative (table S3). The density of individuals within the FGFP microbiome composition landscape resolved into three major peaks, coinciding with the three main contributors to variation identified above (Fig. 1B), as well as enterotypes [based on clustering (*18*) or Dirichlet multinomial mixtures (*19, 20*); fig. S5].

Identifying microbiome covariates

Building upon the extensive FGFP phenotyping, we tested 503 metadata variables (table S1) to identify microbiome covariates. To achieve a balance between number of phenotypes of interest and rates of false discovery, a stepwise approach was applied. After removing collinear variables (table S4), 69 factors were shown to correlate significantly [false discovery rate (FDR) <10%] with overall microbiome community variation (Bray-Curtis dissimilarity; Fig. 2 and table S5). Of those covariates, 26 had an analog in the LLDeep record (*11*). Despite differences in study population and sample analysis (e.g., DNA extraction methods), 24 matching covariates were found to be significantly associated with microbiome composition in the LLDeep cohort, leading to an overall replication success rate of 92% (Fig. 2). All 69 covariates identified correlated with alpha-diversity measures and individual taxa abundances (table S6). However, the predictive power of the linear covariate-based models was limited, as they only explained 1.50 to 14.74% of genus abundance variation (table S7), suggesting additional contribution from unknown factors, stochastic effects, and/or biotic interactions (*21*). Moreover, correlations were affected by interactions between specific covariates, notably medication (see below; table S8).

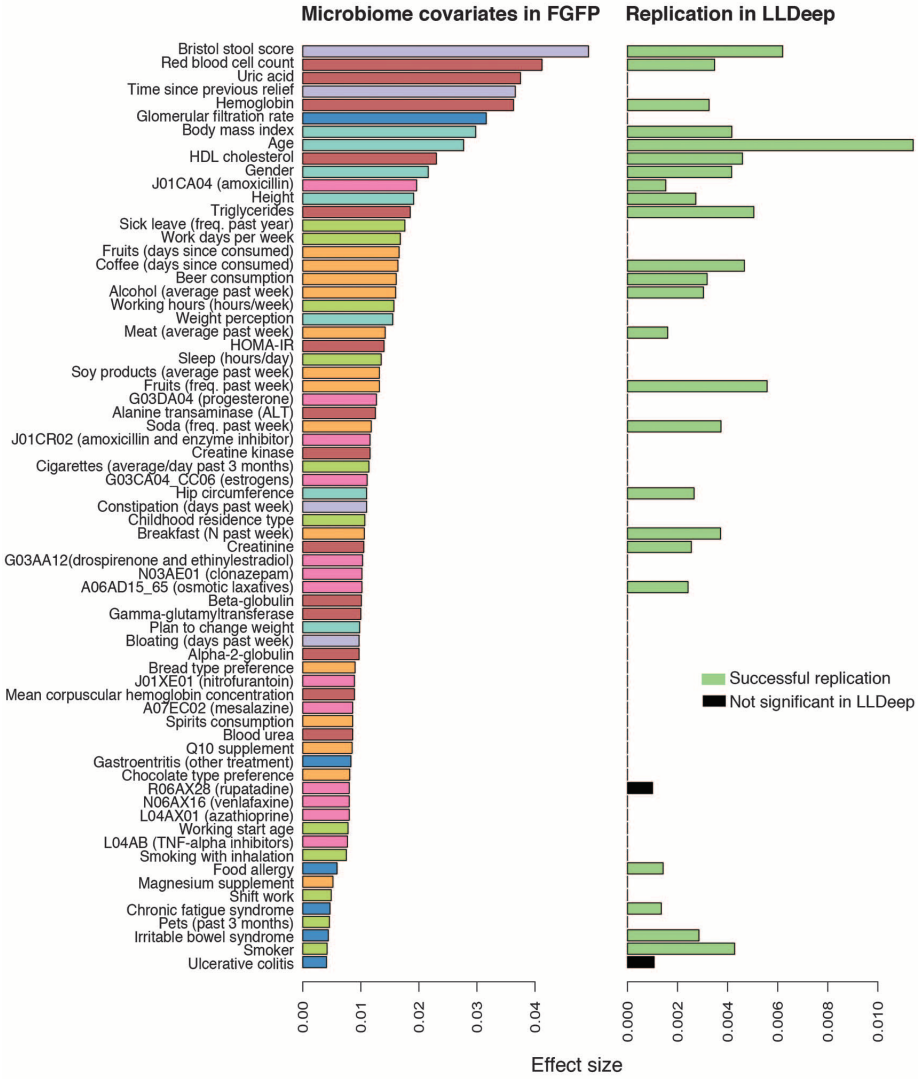

Fig. 2. Microbiome covariates identified in the FGFP cohort (left) and their replication in the LLDeep study (right). Factors are sorted according to their effect size in FGFP and colored based on metadata category (fig. S1), with "Medication" (pink) here split out of the "Health" category (dark blue). Covariates identified in the FGFP and successfully replicated in LLDeep ($P < 5\%$) are colored in green; nonreplicated covariates are in black.

Calculation of the covariates' combined effect size per phenotypical category revealed that medication had the largest explanatory power on microbiome composition, including 10.04% of community variation (Fig. 3A and table S9). Blood parameters, bowel habits, health status, anthropometric features, and lifestyle followed with decreasing combined correlation, raising the total additive effect size of all categories to 16.43%. To identify nonredundant covariates of microbiome variations from our shortlist of 69 correlating factors, we performed a forward stepwise redundancy analysis (RDA) that resulted in a set of 18 variables (Fig. 3B and table S10) with a cumulative (nonredundant) effect size on community variation of 7.63%. Here, we identified stool consistency as the top single, nonredundant microbiome covariate in the FGFP metadata (see below) (*22, 23*). Among the other

Downloaded from http://science.sciencemag.org/ on March 13, 2019

nonredundant covariates were age (*12*) and gender (*24*), but also the intake of specific drugs and dietary information (including fiber uptake, bread preference, and fruit consumption; Fig. 3B). Regarding the ongoing debate on the association between microbiome composition and body mass index (BMI) (*25*, *26*), our analyses revealed that effect size is small but significant (table S10). Notably, previously unidentified factors such as red blood cell (RBC) count and hemoglobin concentration indicated covariation of microbiome composition with blood oxygen uptake capacity (*27*). Previous work in mice has shown an effect of oxygen diffusion on the microbiota (*28*). Moreover, correlations between RBC counts and *Faecalibacterium* abundances are in line with the known oxygen requirements of this genus (*29*). Of the 18 covariates with nonredundant contributions to microbiome variation, 10 were found to be significant by generalized linear model analysis (table S11). This approach confirmed the top covariate status of stool consistency (*22*, *23*) and revealed associations between genus abundances and hip circumference, uric acid concentrations, amoxicillin intake, and chocolate-type preference (namely, an increased abundance of unclassified Lachnospiraceae in participants with a preference for dark chocolate).

Out of a total of 503 parameters, stool consistency, as measured by self-assessed Bristol stool scale (BSS) score, emerged as the top feature covarying with fecal microbiome composition. BSS score has been put forward as an indicative measure of transit time (*30*), but also reflects water availability and potential niche differentiation within the colon ecosystem (*23*). We confirmed previously reported associations of stool consistency with microbiota richness, prevalence of *Prevotella*-enterotyped samples, and *Akkermansia* and *Methanobrevibacter* abundances (*22*, *23*) (Fig. 4, fig. S6, and table S12). In addition, we showed that 12 out of 20 of the FGFP 99% core genera covary with BSS scores, with overall core abundance increasing in looser stools. We assessed the confounding effect of stool consistency on the remaining 68 microbiome covariates using RDA. Among the features losing most explanatory power were time since previous relief (also indicative of passage rates), blood uric acid and hemoglobin levels, BMI, gender, and frequency of beer consumption (table S12).

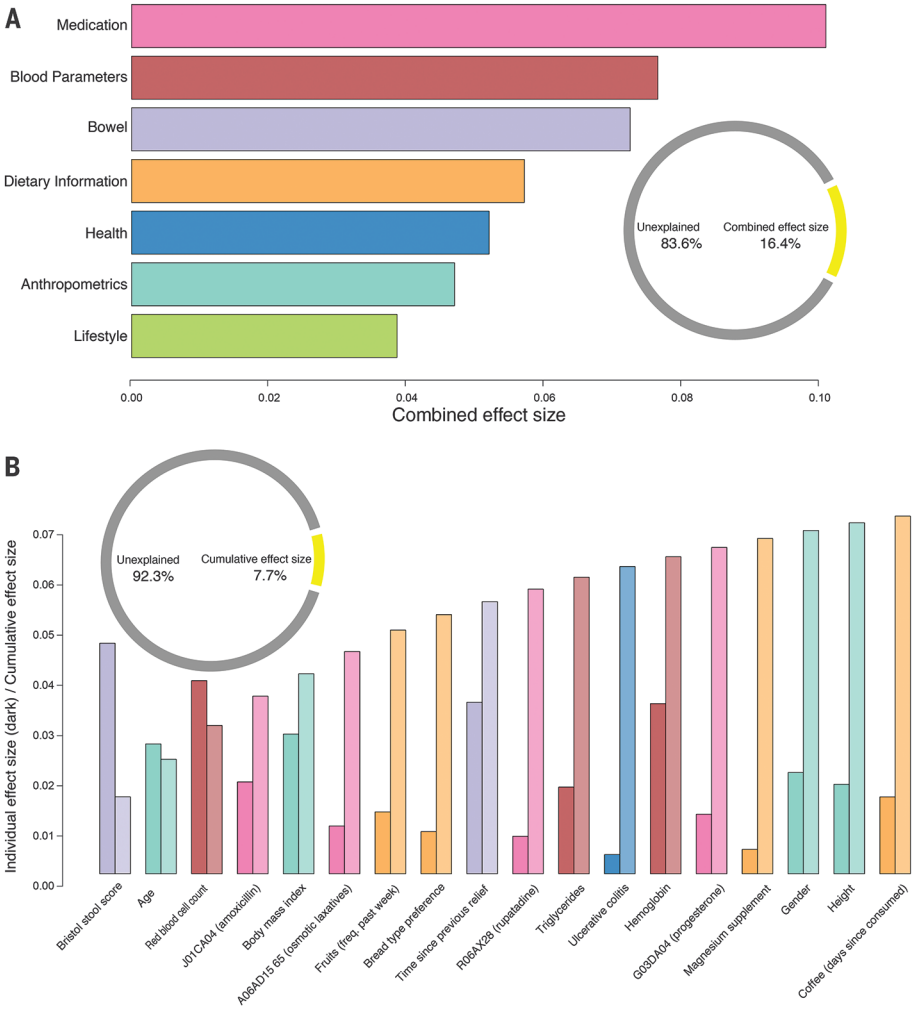

Fig. 3. Effect sizes of covariates of FGFP microbiome composition. (**A**) Combined effect size of FGFP covariates pooled in predefined categories (Fig. 2 color codes) with covariate distance-based selection. (**B**) Cumulative effect size of nonredundant covariates selected by stepwise RDA analysis (right bars) as compared to individual effect sizes assuming independence (left bars). Rings in each panel show the fraction of microbial variation explained with the approach.

Bacterial genera associated with disease

Years of disease-targeted microbiome research have generated an extensive inventory of bacterial genera with a reported association with one or more pathologies. We have assessed correlations between taxa that have been reported to be more abundant or depleted in individuals suffering from specific conditions (table S13) and the set of 18 nonredundant microbiome covariates identified. Our analyses confirmed previous work showing that *Akkermansia* abundance positively correlated with time since previous relief (*23*), but it was also negatively associated with insulin resistance risk factors such as BMI and blood triglyceride concentrations (*31*). *Faecalibacterium* numbers were, as discussed, dependent on RBC counts, but our analysis did find a decreased abundance in ulcerative colitis patients (*32*). The presence of *Fusobacterium* could not be linked to any of the nonredundant covariates identified in this study, which could indicate the specificity of its association with colorectal cancer (*8*). Given these associations, inclusion of the identified covariates in future clinical study design seems appropriate.

Next, we identified sample subsets with specific taxonomic signatures using a biclustering approach (*33*). Two stable biclusters were detected, spanning 410 and 374 samples, respectively, with an intersection of 92 (table S14). The first bicluster comprised 15 genera, including several Clostridia, as well as hydrogenotrophic genera, such as *Methanobrevibacter* and *Desulfovibrio*. The cluster was predominantly composed of women, individuals with a lower weight, and participants with a longer transit time, as reflected both by stool consistency and time since previous relief. Both microbiota richness and evenness were elevated in this cluster. In contrast, the second bicluster, consisting of seven genera, including *Bacteroides* and *Parabacteroides*, comprised individuals with reduced microbiome diversity. Characterization of these individuals revealed a preference for white, low-fiber bread [bread being the major source of carbohydrates in an average Belgian diet (*34*)] and higher prevalence of recent amoxicillin treatment. Thus, this biclustering analysis hinted at microbiome configurations that at least partially overlap with previously described enterotypes. Indeed, while the *Ruminococcus* enterotype was overrepresented in the first bicluster, the second was enriched in *Bacteroides*-type individuals. This, together with the results from Fig. 1B, suggested that although not discrete, enterotypes do indeed represent "densely populated areas in a

Downloaded from http://science.sciencemag.org/ on March 13, 2019

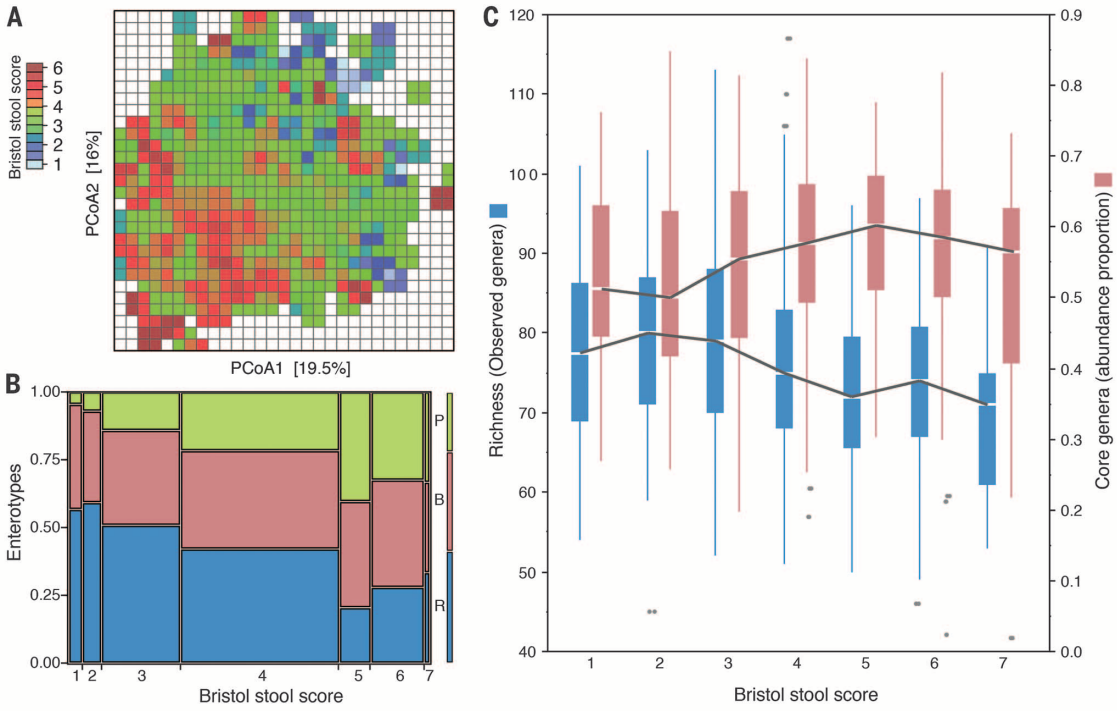

Fig. 4. BSS score association to microbiota variation. (**A**) BSS score variation across the FGFP cohort, as represented on the genus-level PCoA ordination (Bray-Curtis dissimilarity). Each cell is colored according to median BSS score of individual samples allocated to the cell coordinates. (**B**) Enterotype distribution over BSS scores [JSD enterotyping (*18*)] showing an increase in *Prevotella* individuals with looser stool consistency. (**C**) Median differences in abundance of the core microbiota (FGFP genus-level core at 99%) and in observed genus richness across BSS score.

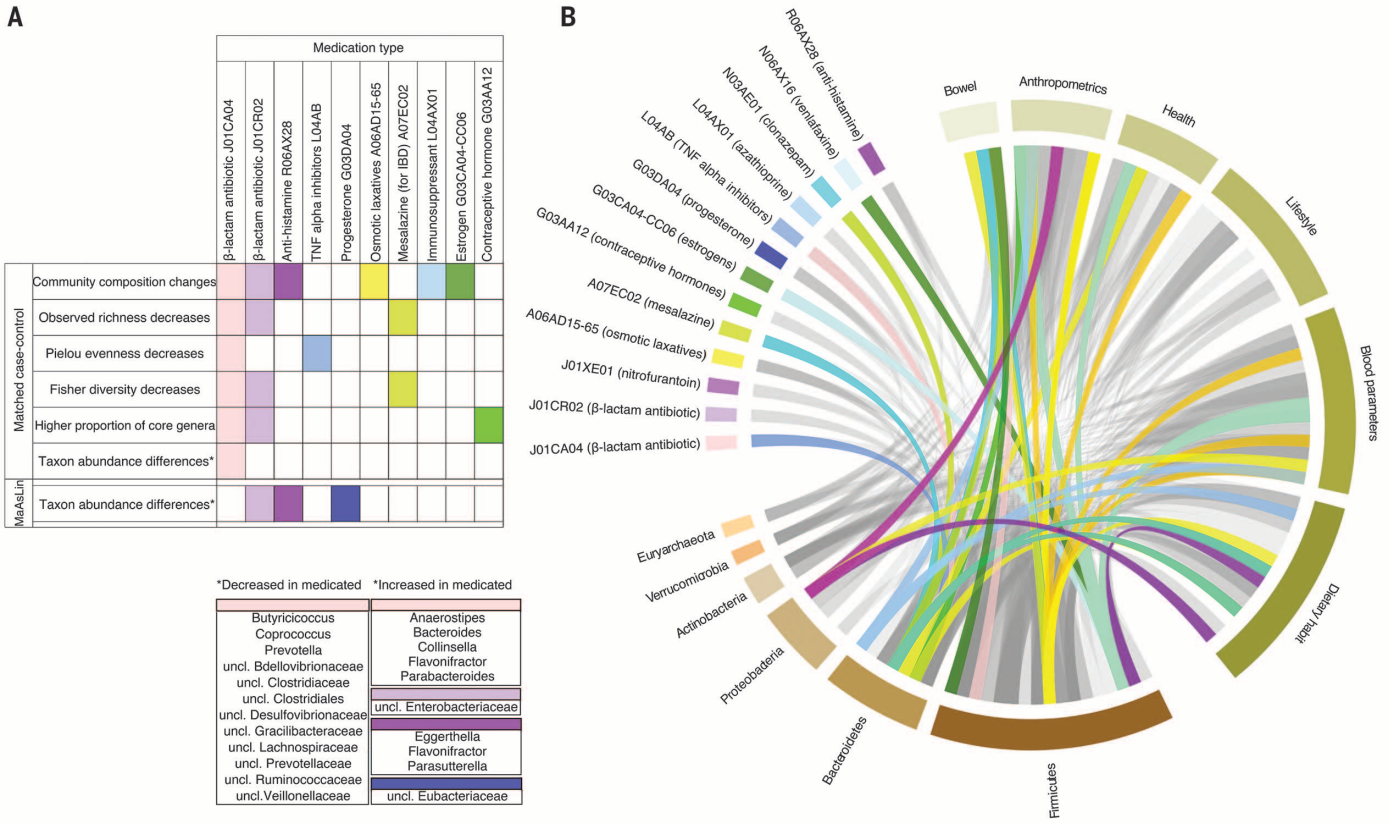

Fig. 5. Drug interactions in the FGFP. (**A**) Overview of the association between different types of medication and microbiome composition. Colored boxes (color coding according to medication) represent a significant result in the matched case-control (FDR<5%) or boosted additive general linear modeling (FDR<10%, table S11) analyses. The effect (decrease/increase) of medication on genera abundances is specified. (**B**) Circos plot showing correlations between covariates and genus abundances (FDR<10%) interacting with drugs. Genera are grouped at phylum level; ribbons represent genus-phenotype associations and are colored according to the confounding medication (gray indicates nonconfounded).

Downloaded from http://science.sciencemag.org/ on March 13, 2019

multidimensional space of community composition," as stated in the original publication (*18*).

The effect of medical interventions

When combining FGFP covariates in predefined categories (fig. S1 and table S15), the use of medication showed the largest explanatory value for microbiome variation in our study. The use of medication in the FGFP cohort was widespread [with 1950 records of over-the-counter plus prescription drug intake during the past 12 (antibiotics) or 6 months (all others) prior to sampling]. On the shortlist of 69 FGFP microbiome covariates figured 13 drugs, including antibiotics, osmotic laxatives, inflammatory bowel disease (IBD) medication, female hormones, benzodiazepines, antidepressants, and antihistamine. Independently of other covariates, intake of several of these substances was associated with community composition variation (Fig. 5A and table S15). The only drugs significantly associated with the abundance of specific genera in phenotype-matched case-control analyses were β-lactam antibiotics (FDR <5%). As medication was shown to affect the outcome of microbiome association studies (*35*), we performed an interaction analysis of covariate-microbiome correlations in the FGFP data set (table S8). Of the covariate interactions detected, 63% was driven by medication (Fig. 5B). This result highlights the versatility of drug-microbiome associations and stresses their importance as potentially confounding factors in clinical studies.

Some early-life events that are generally thought to affect adult microbiota composition were not associated with microbiota composition variation in our study, including mode of birth [cesarean section (N = 36) or vaginal delivery (N = 1036)], place of birth [home (N = 207) or hospital (N = 899); increased diversity in home-born individuals, FDR>5% when controlling for age], and infant nutrition [breastfed (N = 537) or not breastfed (N = 359)] (fig. S7). Residence type [ranging from countryside (N = 77) over rural village (N = 500), small town (N = 272), suburb (N = 137), to city (N = 102)] during early childhood (up to 5 years old), one of the 69 FGFP microbiome covariates, was linked to adult microbial community composition, with a positive correlation between evenness and residence in more industrialized areas, though not statistically significant (FDR >5%) when correcting for age, gender, and BMI. Although the lack of signal in the data was unexpected, these results by no means imply that early-life events do not affect microbiota assembly during infancy, nor do they question previous associations with disease or allergy (*36*, *37*); our analyses only indicated that such events were not significantly associated with microbiome composition at adult age in the FGFP cohort.

Power analysis and conclusions

Finally, the sample size and phenotypic breadth of the FGFP data set provided a unique opportunity to perform an informed power analysis for clinical microbiome studies. In a first approach, we calculated the number of samples needed to assess a difference in dominant microbiota members in a case-control setting where the type of microbiota shift is unknown (e.g., for a discovery project in an unstudied disease). We could detect a 9% difference between taxon proportions with 400 samples per group at a power above 95% and a 5% difference with 500 samples per group at a power of 80% (table S16). In a second approach, we estimated the sample size needed to identify a microbiome shift specific to a known association in a background of other factors (e.g., for intervention studies). Focusing on the prevalent concern of BMI increase and suboptimal health, we assessed the sample size needed to evaluate microbiota compositional changes associated to obesity. To do so, we calculated the independent effect sizes of obesity status, gender, age, and BSS on microbiota variation (table S16). This allowed us to estimate that 865 lean (BMI <25) and 865 obese (BMI ≥30) volunteers would be necessary to study microbiota compositional shifts with $P < 5\%$ significance level and a power of 80%. When taking into account gender, age, and BSS score as covariates, the estimated sample size was reduced to 535 (table S16).

Overall, this study identified a global human core microbiota, while also highlighting that total gut diversity is not yet covered, even combining microbiome data from almost 4000 individuals. Building upon rich metadata and a two-cohort design, we identified a set of microbiota covariates with a replication rate of over 92% and a cumulative, nonredundant effect size of 7.63%. This suggests the influence of additional, currently unknown covariates as well as intrinsic microbial ecological processes such as founder effects, species interactions, and dynamics. We showed that some of the medical conditions targeted by fecal microbiota research have much smaller microbiome effect sizes than commonly assumed. However, some of the covariates that we identified (such as BSS and medication) are currently largely ignored and should be taken into account in future clinical studies. Our power analyses showed that large-scale study design is indispensable for characterizing microbiome shifts, even in a controlled setting, confirming that scale indeed matters, but knowledge of confounders can help to ease power issues. The results from this study form a solid basis for the development of microbiome research as a clinical and diagnostic field.

REFERENCES AND NOTES

1. M. Rajilić-Stojanović *et al.*, *Am. J. Gastroenterol.* **110**, 278–287 (2015).
2. J. C. Clemente, L. K. Ursell, L. W. Parfrey, R. Knight, *Cell* **148**, 1258–1270 (2012).
3. J. F. Cryan, T. G. Dinan, *Nat. Rev. Neurosci.* **13**, 701–712 (2012).
4. E. Le Chatelier *et al.*, *Nature* **500**, 541–546 (2013).
5. J. Qin *et al.*, *Nature* **490**, 55–60 (2012).
6. F. H. Karlsson *et al.*, *Nature* **498**, 99–103 (2013).
7. J. Stock, *Atherosclerosis* **229**, 440–442 (2013).
8. M. R. Rubinstein *et al.*, *Cell Host Microbe* **14**, 195–206 (2013).
9. E. A. Eloe-Fadrosh, D. A. Rasko, *Annu. Rev. Med.* **64**, 145–163 (2013).
10. F. Bäckhed *et al.*, *Cell Host Microbe* **12**, 611–622 (2012).
11. A. Zhernakova *et al.*, *Science* **352**, 565–569 (2016).
12. T. Yatsunenko *et al.*, *Nature* **486**, 222–227 (2012).
13. J. K. Goodrich *et al.*, *Cell* **159**, 789–799 (2014).
14. P. J. Turnbaugh *et al.*, *Nature* **449**, 804–810 (2007).
15. I. Martínez *et al.*, *Cell Rep.* **11**, 527–538 (2015).
16. A. J. Obregon-Tito *et al.*, *Nat. Commun.* **6**, 6505 (2015).
17. S. L. Schnorr *et al.*, *Nat. Commun.* **5**, 3654 (2014).
18. M. Arumugam *et al.*, *Nature* **473**, 174–180 (2011).
19. M. J. Claesson *et al.*, *Nature* **488**, 178–184 (2012).
20. I. Holmes, K. Harris, C. Quince, *PLOS ONE* **7**, e30126 (2012).
21. K. Faust, J. Raes, *Nat. Rev. Microbiol.* **10**, 538–550 (2012).
22. E. F. Tigchelaar *et al.*, **65**, 540–542 (2016).
23. D. Vandeputte *et al.*, *Gut* **65**, 57–62 (2016).
24. C. Huttenhower *et al.*, *Nature* **486**, 207–214 (2012).
25. W. A. Walters, Z. Xu, R. Knight, *FEBS Lett.* **588**, 4223–4233 (2014).
26. M. M. Finucane, T. J. Sharpton, T. J. Laurent, K. S. Pollard, *PLOS ONE* **9**, e84689 (2014).
27. H. Mairbäurl, *Front. Physiol.* **4**, 332 (2013).
28. L. Albenberg *et al.*, *Gastroenterology* **147**, 1055–63.e8 (2014).
29. M. T. Khan *et al.*, *ISME J.* **6**, 1578–1585 (2012).
30. S. J. Lewis, K. W. Heaton, *Scand. J. Gastroenterol.* **32**, 920–924 (1997).
31. A. Tirosh *et al.*, *Diabetes Care* **31**, 2032–2037 (2008).
32. H. Sokol *et al.*, *Inflamm. Bowel Dis.* **15**, 1183–1189 (2009).
33. S. Hochreiter *et al.*, *Bioinformatics* **26**, 1520–1527 (2010).
34. S. Devriese, I. Huybrechts, M. Moreau, H. Van Oyen, The Belgian food consumption survey 1-2004 (IPH/EPI Reports N° 2006 - 016, Epidemiology Unit, Scientific Institute of Public Health, Brussels, 2006; www.wiv-isp.be/epidemio/epinl/foodnl/table04.htm.
35. K. Forslund *et al.*, *Nature* **528**, 262–266 (2015).
36. M. G. Dominguez-Bello *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11971–11975 (2010).
37. C. J. Lodge *et al.*, *Acta Paediatr.* **104** (suppl. S467), 38–53 (2015).

ACKNOWLEDGMENTS

FGFP procedures were approved by the medical ethics committee of the University of Brussels–Brussels University Hospital (approval 143201215505, 5/12/2012). A declaration concerning the FGFP privacy policy was submitted to the Belgian Commission for the Protection of Privacy. Written informed consent was obtained from all participants. The FGFP was funded with support of the Flemish government (IWT130359), the Research Fund–Flanders (FWO) Odysseus program (G.0924.09), the King Baudouin Foundation (2012-J80000-004), FP7 METACARDIS HEALTH-F4-2012-305312, VIB, the Rega institute for Medical Research, and KU Leuven. D.V. is funded by the Agency for Innovation by Science and Technology (IWT). S.V.S. is supported by Marie Curie Actions FP7 People COFUND–Proposal 267139 (acronym OMICS@VIB). M.J., S.V.S., G.L.M, S.C., K.F., J.W., and M.V., K.D. are supported by postdoctoral (six) and predoctoral (two) fellowships, respectively, from the FWO. The LifeLines-DEEP study was funded by the Top Institute Food and Nutrition (TiFN GH001), CardioVasculair Onderzoek Nederland (CVON 2012-03), the Netherlands Organization for Scientific Research (NWO-VIDI 864.13.013), and an FP7 ERC Advanced Grant to C.W. (Agreement no. 2012-322698). We thank P. Cornelis and all members of the Raes lab for lively discussions and feedback. We acknowledge the contribution of Flemish GPs and pharmacists to data and sample collection. Finally, we thank all FGFP volunteers for participating in the project. Data are available at the European Genotyping Agency (https://www.ebi.ac.uk/ega/)–study no. EGAS00001001689.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6285/560/suppl/DC1
Materials and Methods
Figs. S1 to S9
Additional Data tables S1 to S17
References (*38*–*57*)

1 September 2015; accepted 11 March 2016
10.1126/science.aad3503

Downloaded from http://science.sciencemag.org/ on March 13, 2019

Science

Population-level analysis of gut microbiome variation

Gwen Falony, Marie Joossens, Sara Vieira-Silva, Jun Wang, Youssef Darzi, Karoline Faust, Alexander Kurilshikov, Marc Jan Bonder, Mireia Valles-Colomer, Doris Vandeputte, Raul Y. Tito, Samuel Chaffron, Leen Rymenans, Chloë Verspecht, Lise De Sutter, Gipsi Lima-Mendez, Kevin D'hoe, Karl Jonckheere, Daniel Homola, Roberto Garcia, Ettje F. Tigchelaar, Linda Eeckhaudt, Jingyuan Fu, Liesbet Henckaerts, Alexandra Zhernakova, Cisca Wijmenga and Jeroen Raes

Science **352** (6285), 560-564.
DOI: 10.1126/science.aad3503

"Normal" for the gut microbiota

For the benefit of future clinical studies, it is critical to establish what constitutes a "normal" gut microbiome, if it exists at all. Through fecal samples and questionnaires, Falony *et al.* and Zhernakova *et al.* targeted general populations in Belgium and the Netherlands, respectively. Gut microbiota composition correlated with a range of factors including diet, use of medication, red blood cell counts, fecal chromogranin A, and stool consistency. The data give some hints for possible biomarkers of normal gut communities.

Science, this issue pp. 560 and 565

ARTICLE TOOLS http://science.sciencemag.org/content/352/6285/560

SUPPLEMENTARY MATERIALS http://science.sciencemag.org/content/suppl/2016/04/27/352.6285.560.DC1

RELATED CONTENT
http://science.sciencemag.org/content/sci/352/6285/530.full
http://science.sciencemag.org/content/sci/352/6285/532.full
http://science.sciencemag.org/content/sci/352/6285/535.full
http://science.sciencemag.org/content/sci/352/6285/539.full
http://science.sciencemag.org/content/sci/352/6285/544.full
http://science.sciencemag.org/content/sci/352/6285/565.full
http://stm.sciencemag.org/content/scitransmed/7/307/307fs39.full
http://stm.sciencemag.org/content/scitransmed/7/307/307ra152.full
http://stm.sciencemag.org/content/scitransmed/7/295/295ed8.full
http://stm.sciencemag.org/content/scitransmed/7/293/293ra103.full

REFERENCES This article cites 54 articles, 9 of which you can access for free
http://science.sciencemag.org/content/352/6285/560#BIBL

PERMISSIONS http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. The title *Science* is a registered trademark of AAAS.

Downloaded from http://science.sciencemag.org/ on March 13, 2019