

community instability, this information is not used as a feature to improve classifier accuracy. Could microbiome instability become actionable by creating a new classifier that benefits from repeated measurements? If so, how many samples per individual are needed to assess instability?

We collected daily stool samples for up to 6 weeks from 19 CD subjects and 12 controls (see the analysis notebook for cohort description, methods and data, <https://github.com/knightlab-analyses/longitudinal-ibd>) over two separate periods of 2 or 4 weeks spread over 2 and 5 months, for a total of 960 samples. We believe that this is the most densely sampled longitudinal study of CD; previous studies collected samples every 1–3 months.^{1,3} Our cohort shows decreased alpha diversity and increased stability, as previously reported in CD and other subtypes of IBD.^{1,3–5} We also noted that subjects who underwent resection have lower alpha diversity than other CD-affected subjects (see analysis notebooks, <https://github.com/knightlab-analyses/longitudinal-ibd>).

A critical experimental design question for clinical studies is whether a finite budget should best be spent collecting samples from more patients or collecting more serial samples from each patient? Therefore, we created a Random Forests⁶ model based on per subject aggregation of longitudinal data for alpha diversity,⁷ beta diversity⁸ and abundances of two phylogenetic factors found to be associated with CD in ileal biopsies^{5,9} (figure 1). With one sample per subject, our model performs worse

Guiding longitudinal sampling in IBD cohorts

We read with interest the work by Pascal *et al* published recently in *Gut*.¹ Here, they report the volatile microbial signatures of patients with Crohn's disease (CD), a quality that greatly hinders our ability to classify healthy from affected subjects using 16S rRNA profiles from stool. Nonetheless, their work overcame these and other complications,² producing a decision tree that classifies subjects with CD, UC, irritable bowel syndrome and anorexia. Although the authors note that both subtypes of IBD, particularly CD, have increased microbial

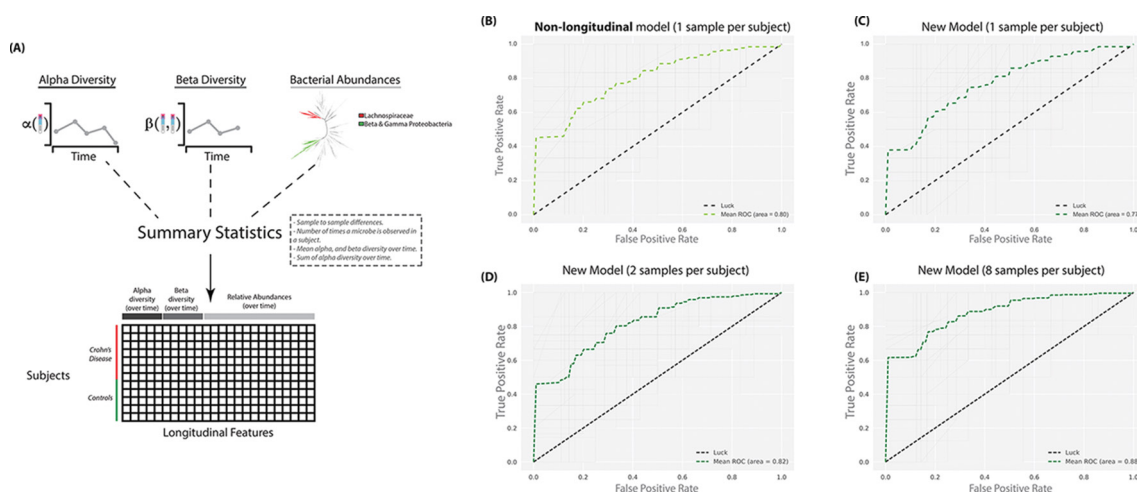


Figure 1 Diagram for the model creation and comparison of four receiver operating characteristic (ROC) curves. (A) Diagram describing the origin for the classifying features. (B) ROC curve for a model that relies on relative abundances and one sample per subject (as used in previous publications). (C–E) ROC curve for our new model at 1, 2 and 8 samples per subject. The grey lines represent the performance at each of the 100 iterations. The dotted black diagonal line represents the performance of a classifier that guesses the labels at random.

Table 1 Performance summary of the classifier at increased samples per subject for this cohort (daily samples) and a previously published cohort

AUC	Samples per subject	Controls	Crohn's disease	Sampling
0.80	1*	12	19	Daily samples
0.77	1	12	19	
0.82	2	12	19	
0.85	3	12	18	
0.86	4	12	18	
0.87	5	12	18	
0.87	6	12	18	
0.88	7	12	18	
0.88	8	12	18	
0.87	9	12	18	
0.87	10	12	17	
0.86	11	12	16	
0.80	1*	9	19	Monthly samples ³
0.80	1	9	19	
0.83	2	9	15	
0.86	3	8	14	
0.92	4	8	12	

The AUC summarises the performance; closer to 1 is better, of the model trained on the different sample sizes as described by the other columns. *Represents the performance of a classifier that relies on non-longitudinal relative abundances only. AUC, area under the curve.

than a classifier that uses microbial relative abundances at a single time point, but when more samples per subject are added, the classifier outperforms that approach and results previously only attained with biopsy samples.⁵ Furthermore, we replicate this observation with a different cohort (table 1).

Novel analyses aggregating features over time and combining both alpha and beta diversity over time using our intensive daily sampling demonstrate that the main benefits are already obtained by collecting between three and five faecal specimens, and no additional benefits are obtained beyond seven serial samples. Similar results are found for monthly sampling. These results highlight the importance of treating CD as a volatile, time-varying condition, even during clinical remission, but provide hope to clinicians in that a relatively small number of samples yield large additional benefits, facilitating patient compliance. This information can be used to design collection of faecal samples for a large prospective cohort of patients with CD for longitudinal studies of host-microbial interactions over time.

The methods demonstrated here have not previously been used for microbiome analyses but have been used for other engineering applications, for example, in production lines to predict product specification outcomes in a steel manufacturer's facility.¹⁰ We expect the results to generalise in other systems, including other GI and hepatic disorders, where

dynamic features of the microbiome, host gene expression or other accessible descriptors can act as indicators of underlying dysbiotic states.

Yoshiki Vázquez-Baeza,¹ Antonio Gonzalez,² Zhenjiang Zech Xu,² Alex Washburne,³ Hans H Herfarth,^{4,5} R Balfour Sartor,^{4,5,6} Rob Knight^{1,2}

¹Department of Computer Science and Engineering, University of California, San Diego, California, USA

²Department of Pediatrics, University of California, San Diego, California, USA

³Department of Microbiology and Immunology, Montana State University System, Bozeman, Montana, USA

⁴Center for Gastrointestinal Biology and Disease, University of North Carolina, Chapel Hill, North Carolina, USA

⁵Division of Gastroenterology and Hepatology, Department of Medicine, University of North Carolina, Chapel Hill, North Carolina, USA

⁶Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, North Carolina, USA

Correspondence to Professor Rob Knight, Department of Computer Science and Engineering, University of California, La Jolla, CA 92093-0763, USA; robknight@ucsd.edu

Acknowledgements We thank the research coordinators of the UNC Multidisciplinary IBD Center at UNC for patient recruitment. Additionally, we thank Gail Ackermann, Jamie Morton and Justin Silverman for their invaluable feedback and discussion provided in the preparation of this manuscript.

Contributors YVB wrote the manuscript, managed the data depositions and interpreted and analysed the data. AG contributed to the manuscript and interpreted and analysed the data. ZZX interpreted and analysed the data. AW contributed to the manuscript and interpreted and analysed the data. HHH contributed to the manuscript, designed the experiment and

interpreted and analysed the data. RBS wrote the manuscript, designed the experiment and interpreted and analysed the data. RK wrote the manuscript and interpreted, sequenced and analysed the data. All authors finalised and approved the current version of the manuscript.

Funding We acknowledge grant support by NIH (P01-DK094779) and the Crohn's and Colitis Foundation.

Competing interests None declared.

Ethics approval University of North Carolina.

Provenance and peer review Not commissioned; internally peer reviewed.

Data sharing statement The sequences will be deposited in EBI. In addition, the processed sequences and sample information can be found in the Qiita (<https://qiita.ucsd.edu/study/description/2538>) database under the study identifier 2538. Jupyter notebooks, and source code describing all the analyses in this paper can be found online (<https://github.com/knightlab-analyses/longitudinal-ibd>).



OPEN ACCESS

Open access This is an open access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <http://creativecommons.org/licenses/by/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

HHH, RBS and RK contributed equally.



To cite Vázquez-Baeza Y, Gonzalez A, Xu ZZ, *et al.* Gut 2018;**67**:1743–1745.

Received 27 September 2017

Revised 2 October 2017

Accepted 3 October 2017

Published Online First 21 October 2017

Gut 2018;**67**:1743–1745.

doi:10.1136/gutjnl-2017-315352

REFERENCES

- Pascal V, Pozuelo M, Borrueal N, *et al.* A microbial signature for Crohn's disease. *Gut* 2017;**66**:813–22.
- Sartor RB, Wu GD. Roles for intestinal bacteria, viruses, and fungi in pathogenesis of inflammatory bowel diseases and therapeutic approaches. *Gastroenterology* 2017;**152**:327–39.
- Halfvarson J, Brislawn CJ, Lamendella R, *et al.* Dynamics of the human gut microbiome in inflammatory bowel disease. *Nat Microbiol* 2017;**2**:17004.
- Frank DN, St Amand AL, Feldman RA, *et al.* Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci U S A* 2007;**104**:13780–5.
- Gevers D, Kugathasan S, Denson LA, *et al.* The treatment-naïve microbiome in new-onset Crohn's disease. *Cell Host Microbe* 2014;**15**:382–92.
- Breiman L. Random Forests. *Machine Learning* 2001;**45**:5–32.

- 7 Faith DP, Baker AM. Phylogenetic diversity (PD) and biodiversity conservation: some bioinformatics challenges. *Evol Bioinform Online* 2007;2:121–8.
- 8 Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* 2005;71:8228–35.
- 9 Washburne AD, Silverman JD, Leff JW, *et al.* Phylogenetic factorization of compositional data yields lineage-level associations in microbiome datasets. *PeerJ* 2017;5:e2969.
- 10 Christ M, Kempa-Liehr AW, Feindt M. Distributed and parallel time series feature extraction for industrial big data applications. 2016 <https://arxiv.org/abs/1610.07717>