

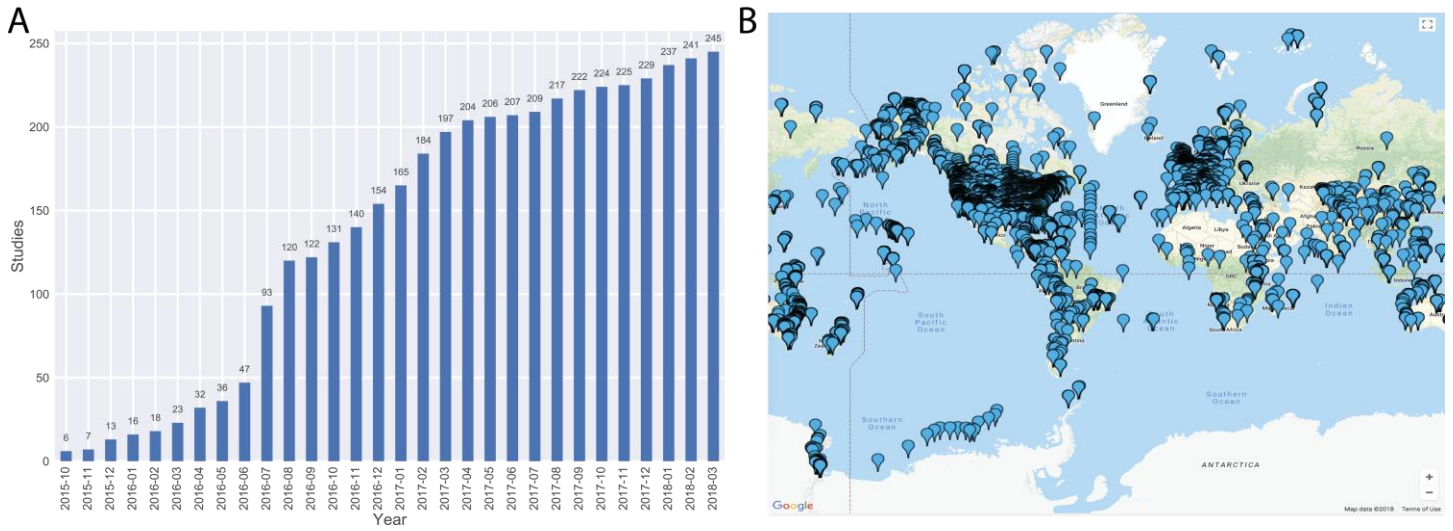
In the format provided by the authors and unedited.

Qiita: rapid, web-enabled microbiome meta-analysis

Antonio Gonzalez^{1,12}, Jose A. Navas-Molina ^{1,2,10,12}, Tomasz Kosciolk¹, Daniel McDonald¹, Yoshiki Vázquez-Baeza¹, Gail Ackermann¹, Jeff DeReus¹, Stefan Janssen¹, Austin D. Swafford³, Stephanie B. Orchanian³, Jon G. Sanders¹, Joshua Shorenstein^{1,11}, Hannes Holste ^{1,2}, Semar Petrus⁴, Adam Robbins-Pianka⁵, Colin J. Brislawn⁶, Mingxun Wang⁷, Jai Ram Rideout ⁸, Evan Bolyen⁸, Matthew Dillon⁸, J. Gregory Caporaso ^{8,9}, Pieter C. Dorrestein ^{1,3,7} and Rob Knight ^{1,2,3*}

¹Department of Pediatrics, School of Medicine, University of California, San Diego, La Jolla, CA, USA. ²Department of Computer Science and Engineering, University of California, San Diego, La Jolla, CA, USA. ³Center for Microbiome Innovation, University of California, San Diego, La Jolla, CA, USA.

⁴Department of Biology, University of California, San Diego, La Jolla, CA, USA. ⁵Department of Computer Science, University of Colorado, Boulder, Boulder, CO, USA. ⁶Earth & Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA, USA. ⁷Collaborative Mass Spectrometry Innovation Center, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA, USA. ⁸Pathogen and Microbiome Institute, Northern Arizona University, Flagstaff, AZ, USA. ⁹Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, USA. ¹⁰Present address: Google LLC, Mountain View, CA, USA. ¹¹Present address: Inscripta, Inc., Boulder, CO, USA. ¹²These authors contributed equally: Antonio Gonzalez and Jose A. Navas-Molina. *e-mail: robknight@ucsd.edu



Supplementary Figure 1

Data loaded in Qiita and uploaded to EBI.

A. Monthly studies and sample depositions to EBI-ENA via Qiita. B. Geographical distribution of the samples present in Qiita.

Supplementary Tables

Step	What's allowed?	What's the expectation?
Sample info file	At least 1 column: sample_name; any column name matching a MIMARKS column is Qiita admin verified to determine MIMARKS compliance	A full MIMARKS compliant file describing the samples and all important columns used for publication.
Preparation info file	At least 1 column: sample_name, must match the sample names in the sample_name; any column name matching a MIMARKS column is Qiita admin verified to determine MIMARKS compliance	A full MIMARKS compliant file describing the processing done on the samples and all important columns used for publication.
Processing - from raw data to BIOM tables	Processing appropriate for the data. For instance, metagenomic read mapping is not exposed for amplicon data.	The user selects a set of parameters that make sense given their data.
Analyses - combining BIOM tables	The user can specify any combination of BIOM tables in the system.	The system groups the selected tables by processing parameters and data types, and merges the tables within each group together such that differing parameter sets are not merged. The merge is done by the specific parameters in the commands that generated the files.

Supplementary Table 1. Main Qiita steps and what's allowed within the system and the expectations for each of those steps.

Study name	Qiita Study ID	Number of Samples
Dynamic changes in short- and long-term bacterial composition following fecal microbiota transplantation for recurrent <i>Clostridium difficile</i> infection	1924	152
The Treatment-Naive Microbiome in New-Onset Crohn's Disease	1939, 1998	1359 284
Dynamics of the gut microbiome in Inflammatory Bowel Disease	1629	683
The Inflammatory Bowel Disease Multi'omics Database	11484	2951
Structure, function and diversity of the healthy human microbiome (V35)	1928	6346
Structure, function and diversity of the healthy human microbiome (V13)	1927	3530

Supplementary Table 2. Studies used for meta-analysis.