Accurate estimation of intraspecific microbial gene content variation in metagenomic data with MIDAS v3 and StrainPGC

Byron J. Smith A (ORCID: 0000-0002-0182-404X)

Chunyu Zhao B (ORCID: 0000-0001-9589-2416)

Veronika Dubinkina A (ORCID: 0000-0002-4844-6795)  
Xiaofan Jin C (ORCID: 0000-0002-0802-7692)  
Liron Zahavi D,E (ORCID: 0000-0001-8867-1287)  
Saar Shoer D,E (ORCID: 0000-0003-0883-1434)

Jacqueline Moltzau-Anderson F (ORCID: 0000-0003-1398-5980)  
Eran Segal D,E

Katherine S. Pollard A,B,G,\* (ORCID: 0000-0002-9870-6196)

A The Gladstone Institute of Data Science and Biotechnology, San Francisco, CA, USA  
B Chan-Zuckerberg Biohub San Francisco, San Francisco, CA, USA  
C Department of Biomedical Engineering, University of Calgary, Calgary, AB, Canada  
D Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel  
E Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel  
F Department of Gastroenterology, University of California, San Francisco, CA, USA  
G Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, USA

\* Corresponding author:  
[katherine.pollard@gladstone.ucsf.edu](mailto:katherine.pollard@gladstone.ucsf.edu)

**Running Title:** Microbial gene content with MIDAS and StrainPGC