Harvard T. H. Chan School of Public Health Microbiome Analysis Core

Jeremy E. Wilkinson[1], Lauren J. McIver[1], Anahit Mailyan[1], Kelsey N. Thompson[1], Curtis Huttenhower[1,2]

[1]Harvard T. H. Chan School of Public Health [2]Broad Institute of MIT and Harvard

The Microbiome Analysis Core at the Harvard T. H. Chan School of Public Health (HMAC) was established in response to the rapidly emerging field of microbiome research and its potential to affect studies across the biomedical sciences. HMAC's goal is to aid researchers with human microbiome study design and interpretation, reducing the gap between primary data and translatable biology. The mission of the HMAC is to provide formalized support, at the highest quality standards, for human microbiome-related studies and to foster collaborative initiatives in microbiome research. The HMAC provides a single point of contact for integrative microbiome informatics and analysis of data from amplicon (16S/18S rRNA gene and internal transcribed spacer [ITS] DNA), shotgun metagenomic, and metatranscriptomic sequencing studies, as well as metabolomics analysis of the human microbiota. The HMAC has extensive experience with the aforementioned data collections, including bioinformatics processing of such data from large cohorts, qualitative ecology, and subsequent microbial systems and human epidemiological analysis. By integrating microbial gene function and taxonomic profiles with host information, we enable researchers to interpret molecular activities of the microbiota and assess its impact on human health.

HMAC services

Consultation for microbiome project development.
This includes consultation on experimental design, sample handling and sequencing, grant proposal development, study power estimation, bioinformatics, and statistical data analysis. First five hours free.

Validated end-to-end meta'omic analysis of microbial community data. Using open-source analytical methods developed in the Huttenhower laboratory and by other leaders in the field, we provide cutting-edge microbiome informatics and analysis.

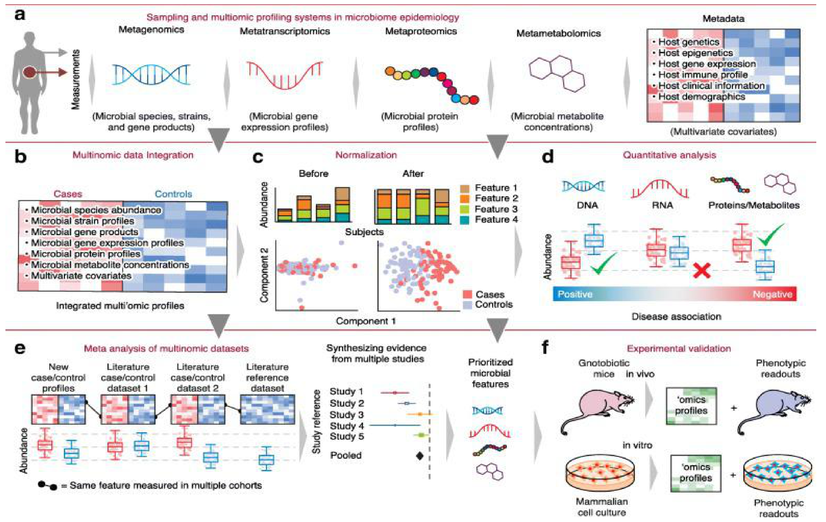
Support fully-collaborative grant-funded investigations.
Includes preliminary data development, hypothesis formulation, grant narrative development, data analysis and inference, custom software development, and co-authored dissemination of findings.

Study Design
- Consultation
- Grant assistance
- Power analysis
- Collection methods
- Wet lab
- Dry lab

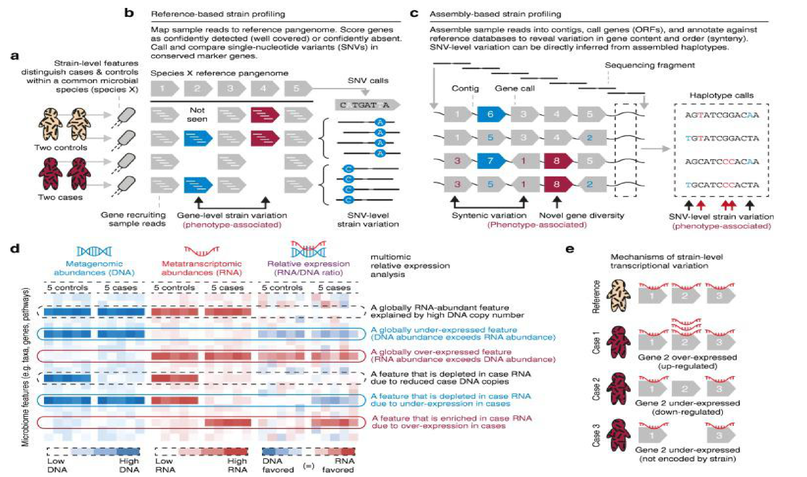
Analysis
- Bioinformatics (raw data processing, taxonomic and functional profiling)
- Downstream analysis and statistics

Interpretation
- Results
- Discussion
- Manuscript writing/editing
- Response to reviewers

Multiomics for microbiome epidemiology



HMAC supports microbiome epidemiology analysis for a variety of molecular data types in human populations or in model systems. Typical analysis workflow steps include **a)** molecular data generation of a variety of types, including but not limited to sequencing, which are **b)** bioinformatically processed into biologically interpretable features and **c)** quality controlled per dataset. This permits **d)** microbiome-tailored statistical methods to associate molecular features with covariates and outcomes, and optionally **e)** meta-analysis of multiple data types per project or across multiple projects. Finally, **f)** HMAC can assist with study design for downstream evaluation of statistical associations in in vivo or in vitro model systems.



Shotgun metagenomic and metatranscriptomic sequence data are particularly amenable to detailed computational analysis, including multiple complementary methods for **a)** strain tracking or differential microbial expression. **b)** Reference-based methods can identify strains using either single nucleotide or structural (genomic) variants, and **c)** can be used in tandem with assembly-based methods for novel microbial discovery. **d)** Whole-community microbial differential expression can additionally be detected either in tandem with or in addition to metagenomic copy number changes, and **e)** analyzed per gene, pathway, microbe, or human individual.

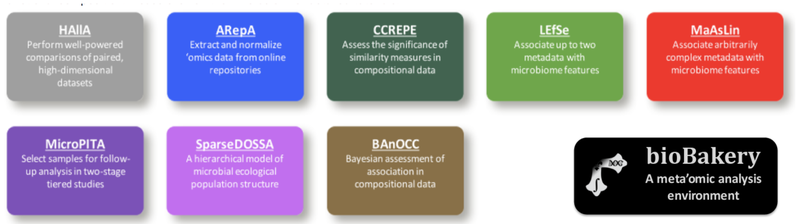
Mallick, H. et al. Experimental design and quantitative analysis of microbial community multiomics. Genome Biology. 18:228 (2017).

Microbial community profiling

The first step in microbiome molecular data analysis is quality control and profiling to transform raw data into biologically interpretable features. This includes identifying microbial species (MetaPhlAn2) and strains (PanPhlAn/StrainPhlAn), characterizing their functional potential or activity (HUMAnN2, ShortBRED), and integrating metagenomics with other data types (PICRUSt, MelonnPan), among others.

Downstream analysis and statistics

Once profiled, microbial communities are amenable to downstream statistics and visualization much like other molecular epidemiology such as human genetic or transcriptional profiles. Like these other data types, microbial communities often require tailored statistics for environmental, exposure, or phenotype association (LEfSe, MaAsLin) or for ecological interaction discovery (BAnOCC). HMAC also provides a variety of tools for bioinformaticians working in the microbiome space.



McIver, L. J. et al. bioBakery: a meta'omic analysis environment. Bioinformatics, 34:7, 1235-1237 (2018).

HARVARD T.H. CHAN SCHOOL OF PUBLIC HEALTH
Microbiome Analysis Core

HSPH Microbiome Analysis Core
Department of Biostatistics
Harvard T. H. Chan School of Public Health
655 Huntington Avenue
Boston, MA 02115
jewilkinson@hsph.harvard.edu
(617) 432-5065

http://hsph.harvard.edu/hmac
http://huttenhower.sph.harvard.edu