

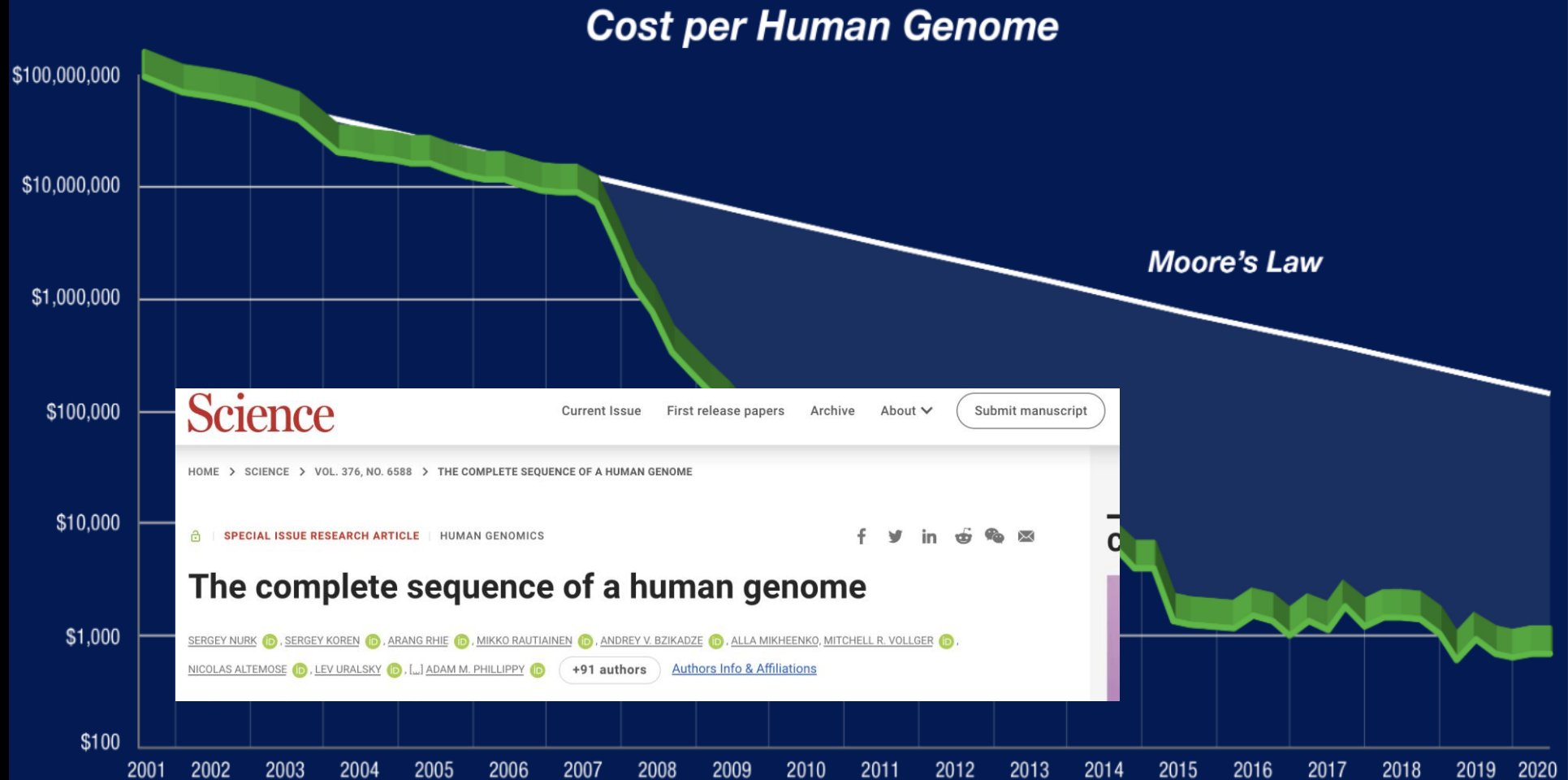
Benchmarking Methods for Metagenomics

ClinGen

April 19th 2022

Tierney

The fastest-ever technological pace

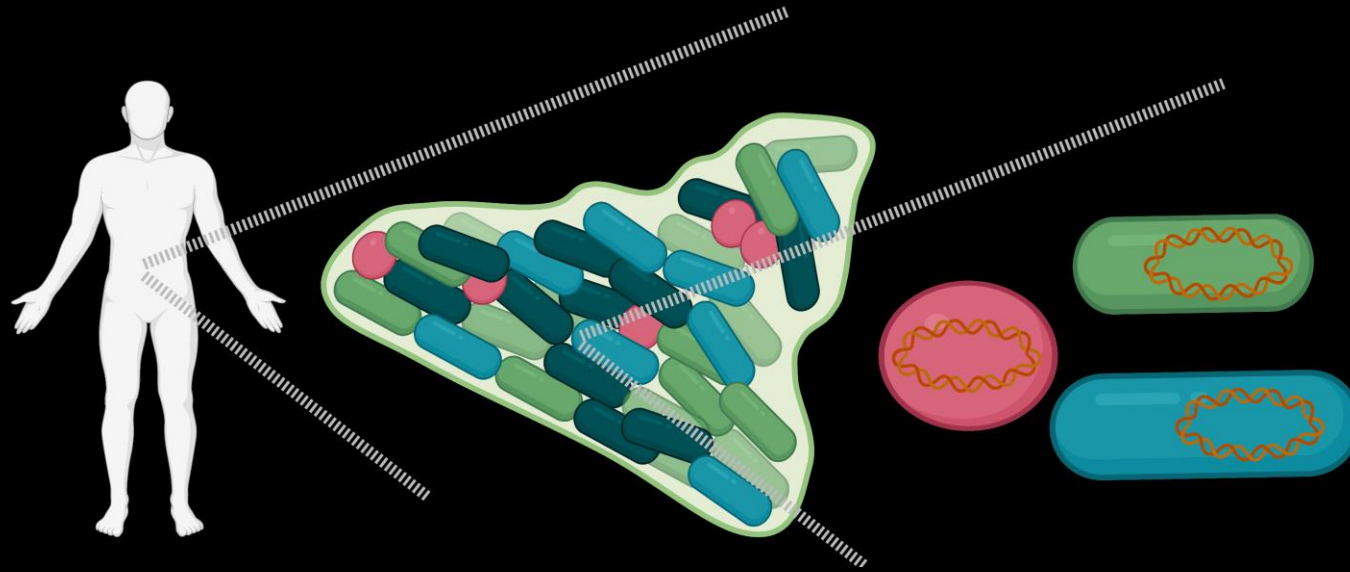


But!

**There is more
than one genome:**

The human metagenome

The complex of microbial communities living in and on your body



Studied body sites include the gut, mouth, skin, nasal cavity, urogenital tract

Trillions of microbial cells, thousands (?) of species, millions (?) of genes

Microbiome content is heavily influenced by environment (e.g. diet)





200 pounds

3-5 pounds bacteria

195-97 pounds human

But why does the microbiome matter?



Altmetric: 1254

24% of the drugs with human targets, including members of all therapeutic classes, inhibited the growth of at least one strain *in vitro*

Article

Extensive impact of non-antibiotic drugs on human gut bacteria

Lisa Maier, Mihaela Pruteanu, Michael Kuhn, Georg Zeller[✉], Anja Telzerow, Exene Erin Anderson, Ana Rita Brochado, Keith Conrad Fernandez, Hitomi Dose, Hirotada Mori, Kiran Raosaheb Patil[✉], Peer Bork[✉] & Athanasios Typas[✉]

Recent advances in basic science

The role of gut microbiota in cancer treatment: friend or foe? ⁸

Wing Yin Cheng¹, Chun-Ying Wu²,  Jun Yu¹

Correspondence to Dr Jun Yu, Institute of Digestive Disease and The Department of Medicine and Therapeutics, Chinese University of Hong Kong, New Territories, Hong Kong; junyu@cuhk.edu.hk



PDF



XML

[Home](#) > [Science Magazine](#) > [8 June 2012](#) > Hood, 336 (6086): 1209

Article Views

- [Summary](#)
- [Full Text](#)
- [Full Text \(PDF\)](#)

VERSION HISTORY

- [336/6086/1209](#)

Published Online June 6 2012

Science 8 June 2012:

Vol. 336 no. 6086 p. 1209

DOI: 10.1126/science.1225475

EDITORIAL

Tackling the Microbiome

Leroy Hood

36%

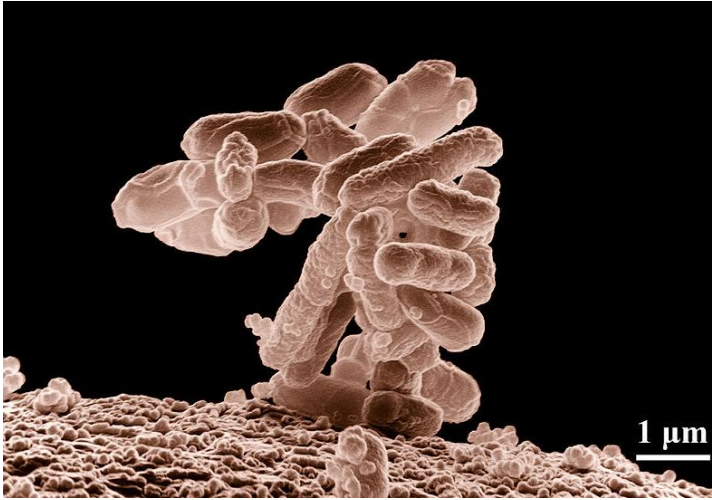
of the small molecules
in your blood
transit the microbiome

The Fecal Microbiome Transplant as the microbiome's greatest clinical success

GASTROENTEROLOGY & HEPATOLOGY
The Independent Peer-Reviewed Journal

Gastroenterol Hepatol (N Y). 2012 March; 8(3): 191–194. PMCID: PMC3365524

Fecal Transplantation for the Treatment of *Clostridium difficile* Infection
[Lawrence J. Brandt](#), MD, MACG, AGAF, FASGE*
[Author information](#) ▶ [Copyright and License information](#) ▶



SCIENTIFIC AMERICAN™
[Sign In / Register](#)

[Subscribe](#) [News & Features](#) [Topics](#) [Blogs](#) [Videos & Podcasts](#) [Education](#) [C](#)

Health :: News :: June 12, 2013 :: 10 Comments :: [Email](#) :: [Print](#)

FDA Comes to Grips with Fecal Transplants

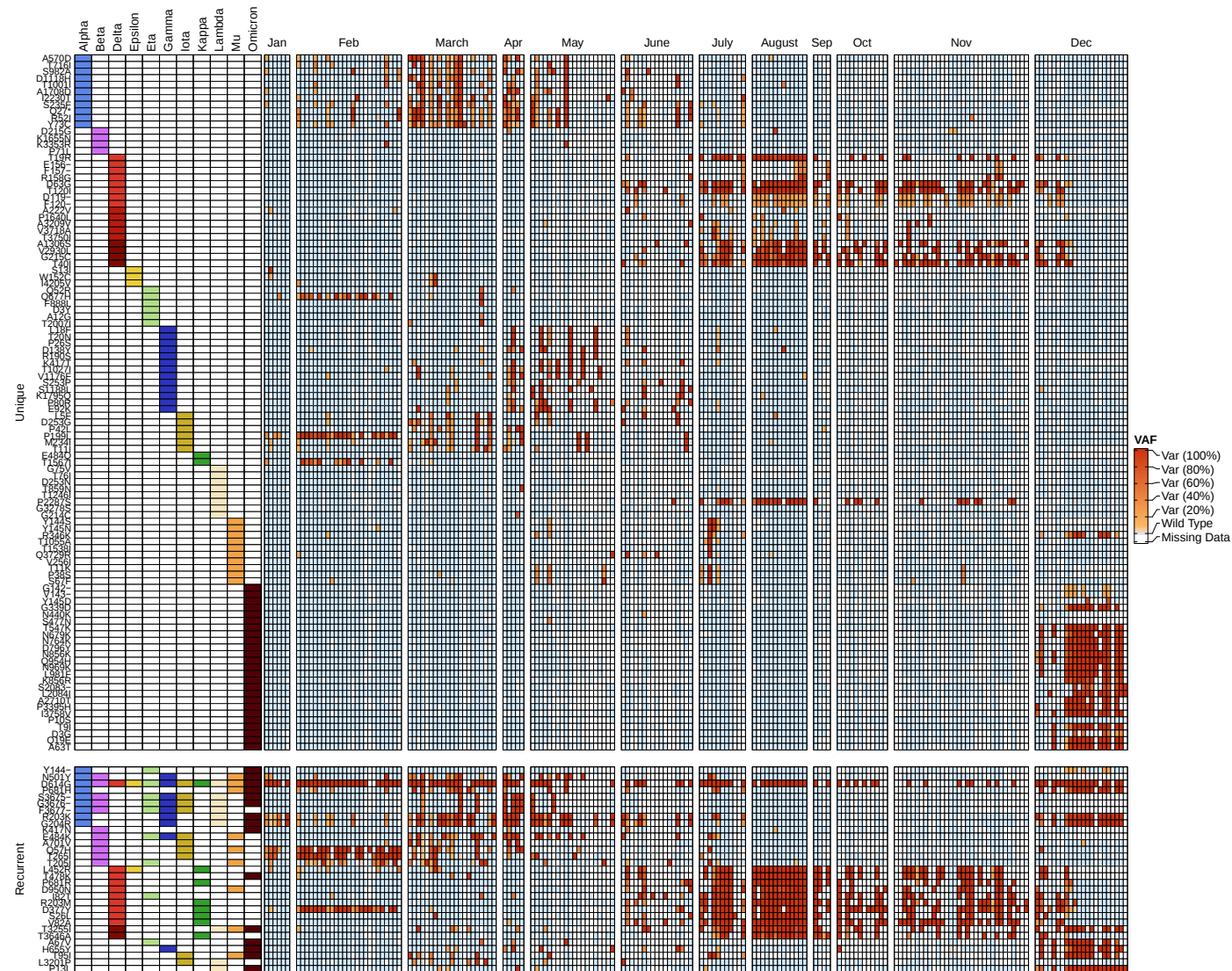
Efforts are under way to standardize fecal transplants, which helps re-establish healthy microbe populations in the guts of patients
By [Beth Mole](#) and [Nature magazine](#)

The POWER of POOP

promoting safe, accessible fecal microbiota transplant
for all who need it

<http://thepowerofpoop.com/>

Tracking signature mutations 2021 Miami wastewater samples



@ Jon Foox + U
Miami SF-RAD
consortium

Additional hypotheses about the human microbiome span the scientific lexicon

The Gut Microbiome as a Target for the Treatment of Type 2 Diabetes

Article | [Open Access](#) | Published: 01 November 2017

Human microbiome signatures of differential colorectal cancer drug metabolism

Leah Guthrie, Sanchit Gupta, Johanna Daily & Libusha Kelly 

npj Biofilms and Microbiomes **3**, Article number: 27 (2017) | [Cite this article](#)

4342 Accesses | 50 Citations | 132 Altmetric | [Metrics](#)



Clinical Gastroenterology and Hepatology

Volume 17, Issue 2, January 2019, Pages 307-321



Microbiome in specific diseases

Altered Microbiome in Patients With Cirrhosis and Complications

nature
International journal of science

Letter | [OPEN](#) | Published: 24 October 2018

The human gut microbiome in early-onset type 1 diabetes from the TEDDY study

Research | [Open Access](#)

Microbiota Transfer Therapy alters gut ecosystem and improves gastrointestinal and autism symptoms: an open-label study

Article | [Open Access](#) | Published: 19 October 2017

Gut microbiome alterations in Alzheimer's disease

Nicholas M. Vogt, Robert L. Kerby, Kimberly A. Dill-McFarland, Sandra J. Harding, Andrew P. Merluzzi, Sterling C. Johnson, Cynthia M. Carlsson, Sanjay Asthana, Henrik Zetterberg, Kaj Blennow, Barbara B. Bendlin  & Federico E. Rey 

Scientific Reports **7**, Article number: 13537 (2017) | [Cite this article](#)

How do we measure the microbiome
with sequencing data?

Taxonomy

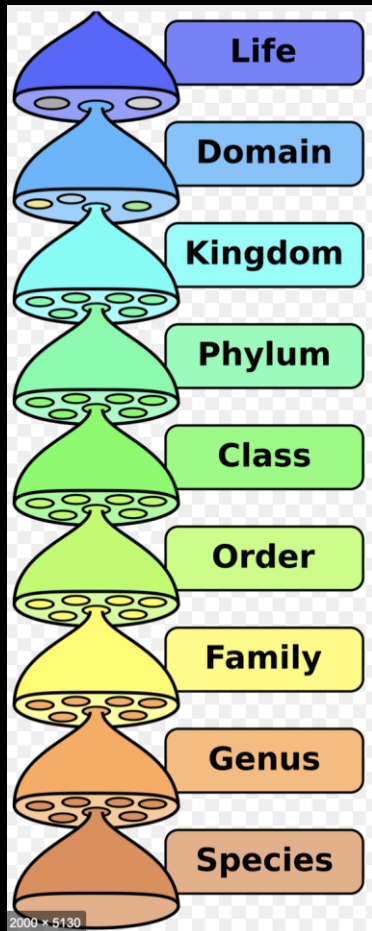
Pathways

Genes

SNPs

Genomes

The complexity of bacterial genetics and taxonomy



Bacteria

Firmicutes

Erysipelotrichia

Erysipelotrichales

Erysipelotrichaceae

Solobacterium

Solobacterium moorei

The 16S ribosomal RNA gene:

Essential for bacterial life

Minimal changes in sequence over time

Used to taxonomically classify microbes

Microbial genomes are circular and encode on average 4K genes

Microbial taxonomy is NOT necessarily correlated to genetics

Two strains of the same species can vary in up to 20% of genomic content

Strains!

Solobacterium moorei CC14Y
Solobacterium moorei CC57E
Solobacterium moorei DSM 22971
Solobacterium moorei F0204

Taxonomy

16S

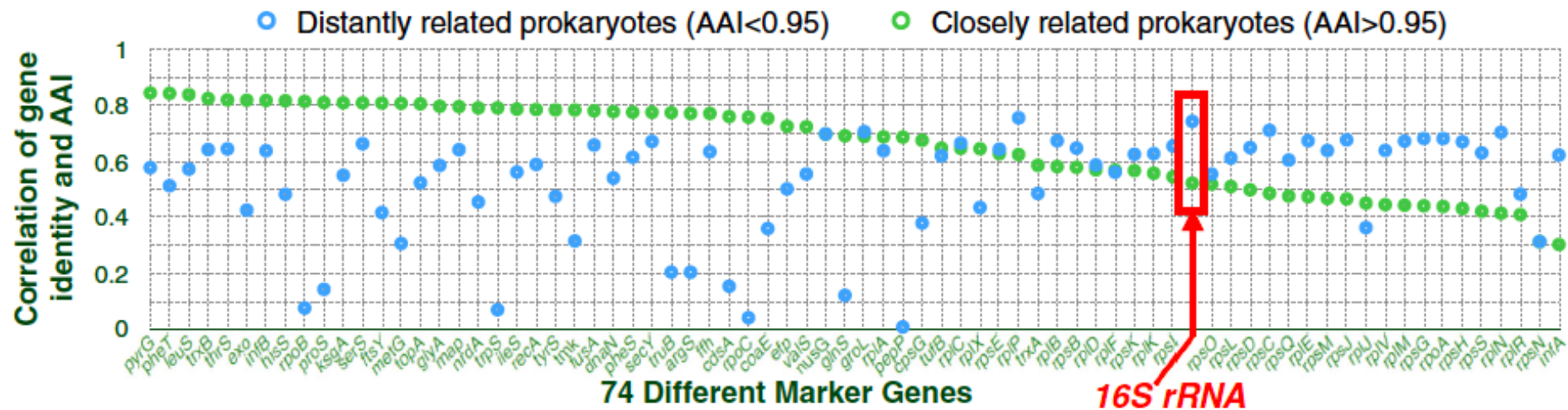
Metabolic function

Genetics

Zhu, A., Sunagawa, S., Mende, D. R., & Bork, P. (2015). Inter-individual differences in the gene content of human gut bacterial species. *Genome Biology*, 16(1), 82. <https://doi.org/10.1186/s13059-015-0646-9>

16S rRNA is limited

“16s rRNA predicts genome-wide levels of similarity very well for distantly related prokaryotes, but not for closely related ones.”



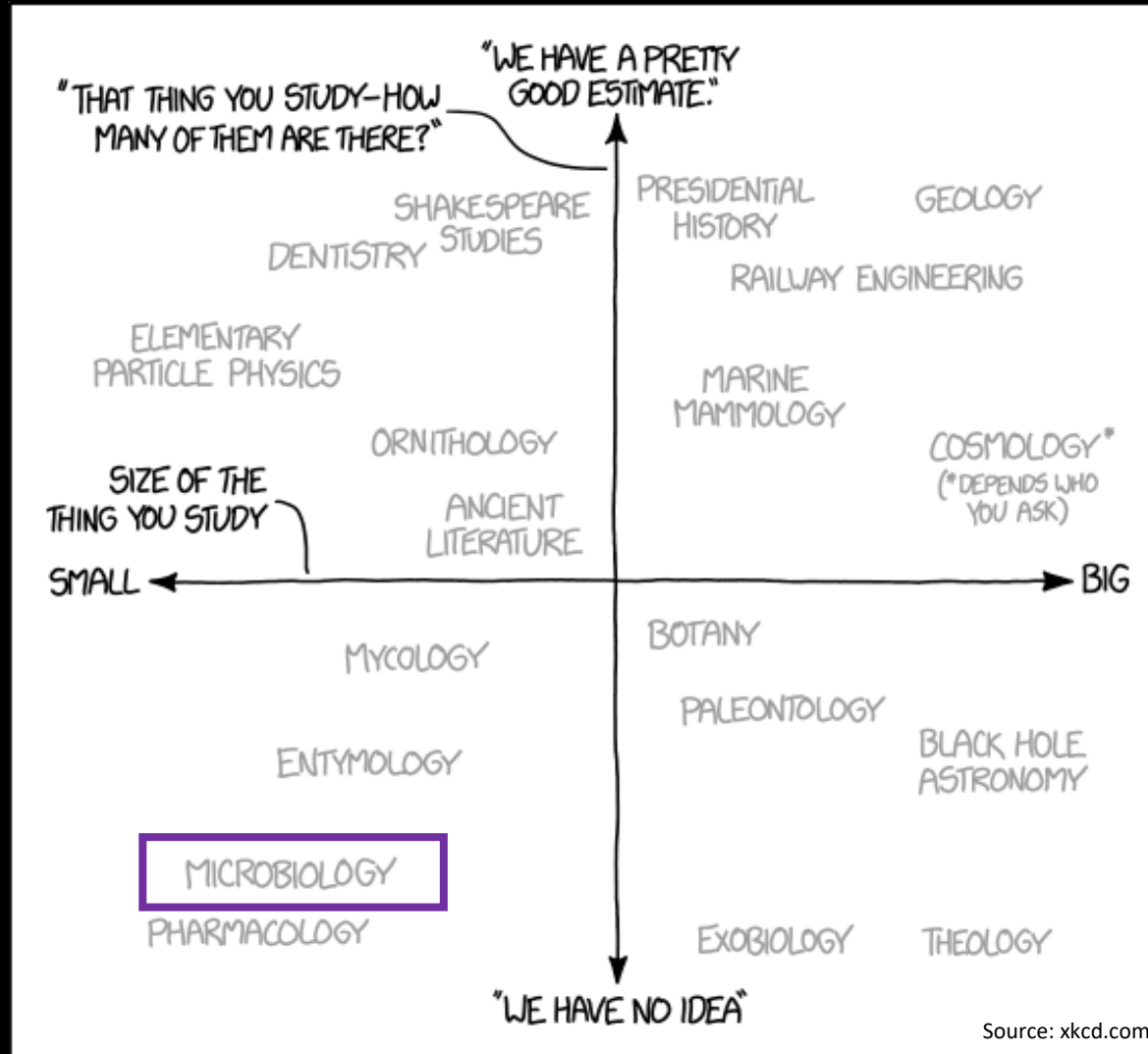
Lan Y, Rosen G, Hershberg R. "Marker genes that are less conserved in their sequences are useful for predicting genome-wide similarity levels between closely related prokaryotic strains." Microbiome. 2016.

Metagenomics can expand the microbiome to query across kingdoms AND data types!!

Data Type	16S	18S	ITS	Shotgun
Taxonomic Classification	Yes	Yes	Yes	Yes
Prokaryotes	Yes	No	No	Yes
Archaea	Yes	No	No	Yes
Eukaryotes	No	Yes	Yes	Yes
Parasites	No	Yes	No	Yes
Plasmids	No	No	No	Yes
Phages	No	No	No	Yes
Human Ancestry	No	No	No	Yes
Biosynthetic Gene Clusters	No	No	No	Yes
Antimicrobial Resistance (AMR) Markers	No	No	No	Yes
Kingdom Specificity	Yes	Yes	Yes	No
Approximate Raw Cost / Sample	\$100	\$100	\$125	\$300

From <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5359768/>

We are still today sorting out what is “in” the human microbiome



- Recent work has estimated there to be hundreds of thousands of strains in the gut microbiome alone (*Almeida et al, Nature 2020*)

Almeida, Alexandre et al. 2021. “A Unified Catalog of 204,938 Reference Genomes from the Human Gut Microbiome.” *Nature Biotechnology* 39 (1): 105–14.

- 2 strains of the same species can vary by 20% in gene content

Zhu, A., Sunagawa, et al. (2015). Inter-individual differences in the gene content of human gut bacterial species. *Genome Biology*, 16(1), 82. <https://doi.org/10.1186/s13059-015-0646-9>

- Variation in gene content drives variation in microbial phenotype

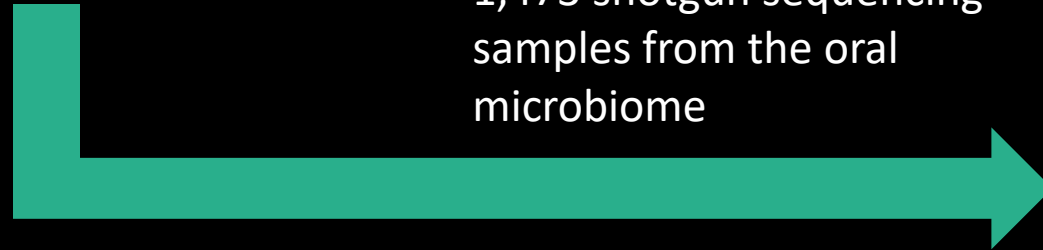
Can we even count the number of genes in the human microbiome?

Prior “complete” estimates of gut microbiome genetic diversity put the total number of genes at ~10 million from ~1000 samples

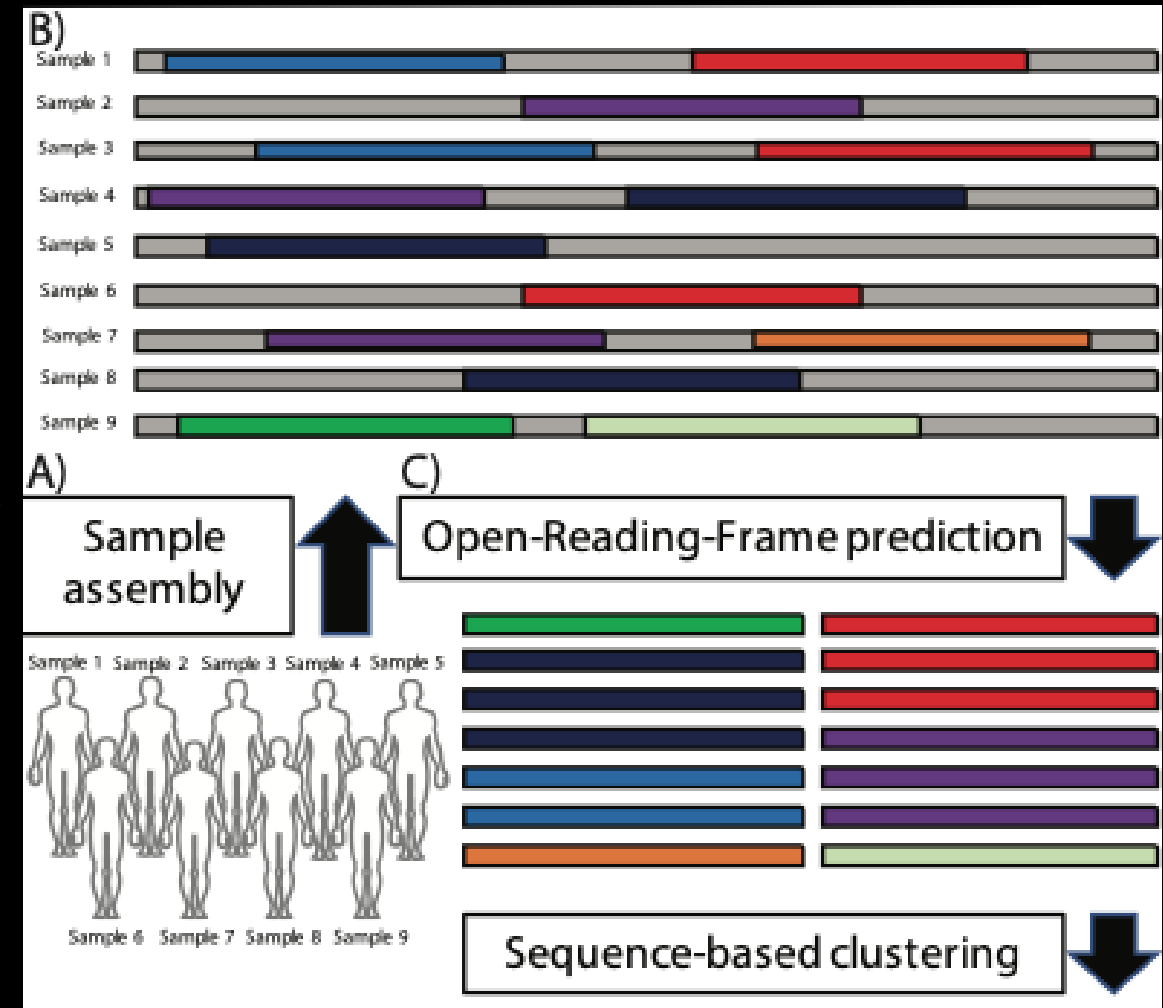
Li et al. 2014. “An Integrated Catalog of Reference Genes in the Human Gut Microbiome.” *Nature Biotechnology* 32 (8): 834–41.

1,473 shotgun sequencing samples from the oral microbiome

2,182 shotgun sequencing samples from the gut microbiome



Tierney, Braden T., Zhen Yang, Jacob M. Luber, Marc Beaudin, Marsha C. Wibowo, Christina Baek, Eleanor Mehlenbacher, Chirag J. Patel, and Aleksandar D. Kostic. 2019. “The Landscape of Genetic Content in the Gut and Oral Human Microbiome.” *Cell Host & Microbe* 26 (2): 283–95.e8.



Li, Weizhong, and Adam Godzik. 2006. “Cd-Hit: A Fast Program for Clustering and Comparing Large Sets of Protein or Nucleotide Sequences.” *Bioinformatics* 22 (13): 1658–59.

Sequence-based clustering: how do we define a "gene?"



ATGTACGTACCGATCGTGAT



ATGTACGTACCGATCGTGAT

Sequence-based clustering: how do we define a "gene?"



ATGTACGTACCGATCGTGAC



ATGTACGTACCGATCGTGAT

Sequence-based clustering: how do we define a "gene?"



ATGTACGTACCGATCGTG**GC**



ATGTACGTACCGATCGTGAT

Sequence-based clustering: how do we define a "gene?"



ATGTACGTACAATCGAAAGC



ATGTACGTACCGATCGTGAT

Sequence-based clustering: how do we define a "gene?"



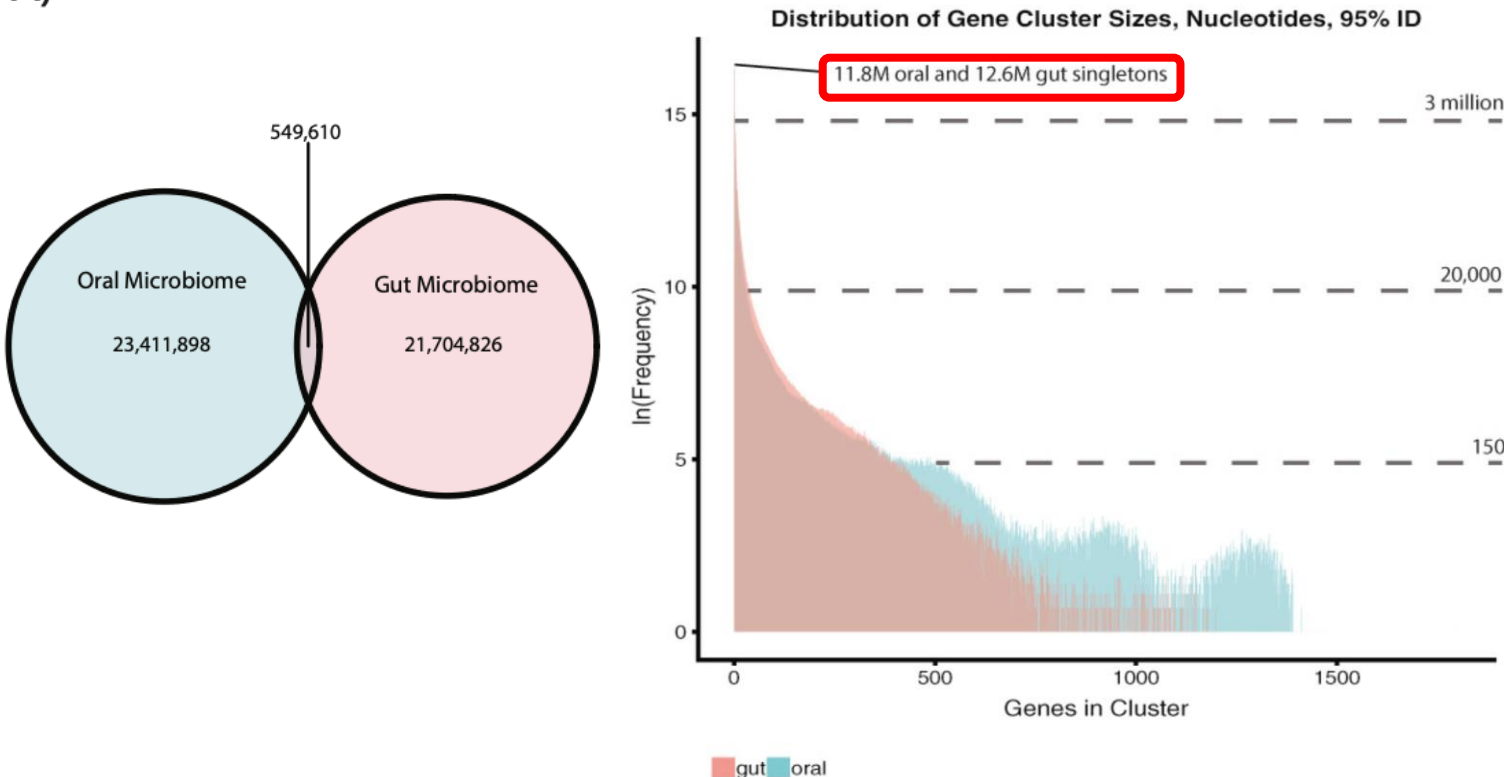
ATGTACGTACCGATCGTGAC



ATGTACGTACCGATCGTGAT

Most studies use 95% nucleotide or
90% amino acid identity to
discriminate between genes

A)



“Singletons” dominate the oral and gut microbiomes

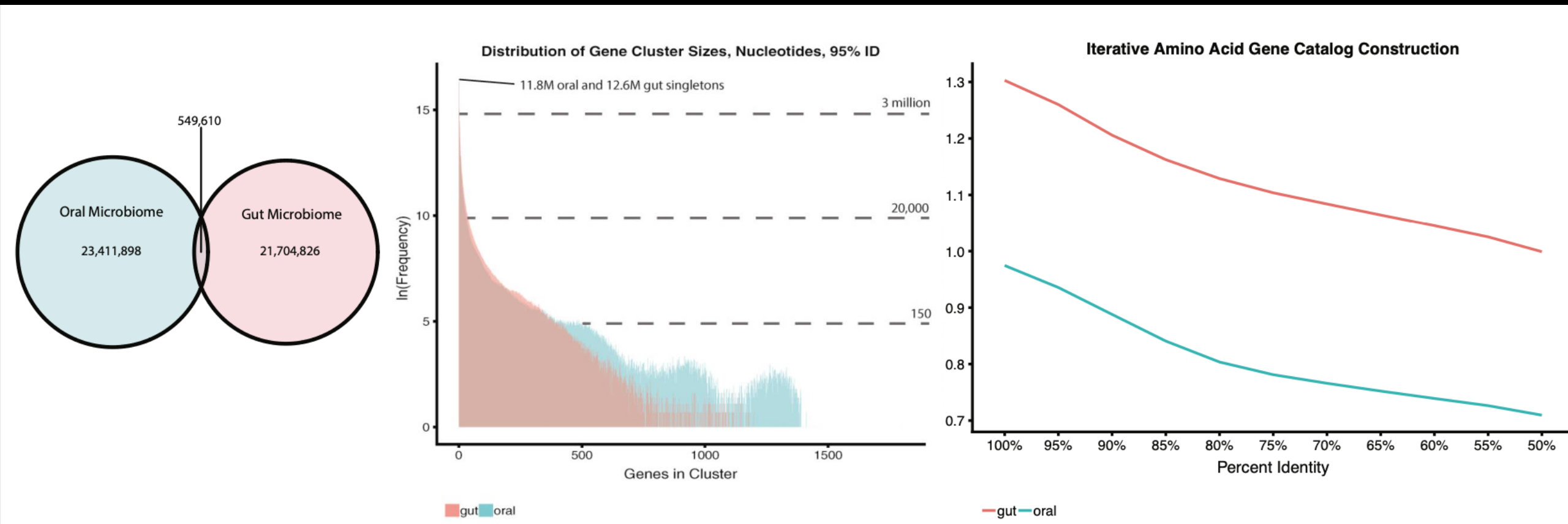


ATGTACGTAC**AATCGAAAGC**

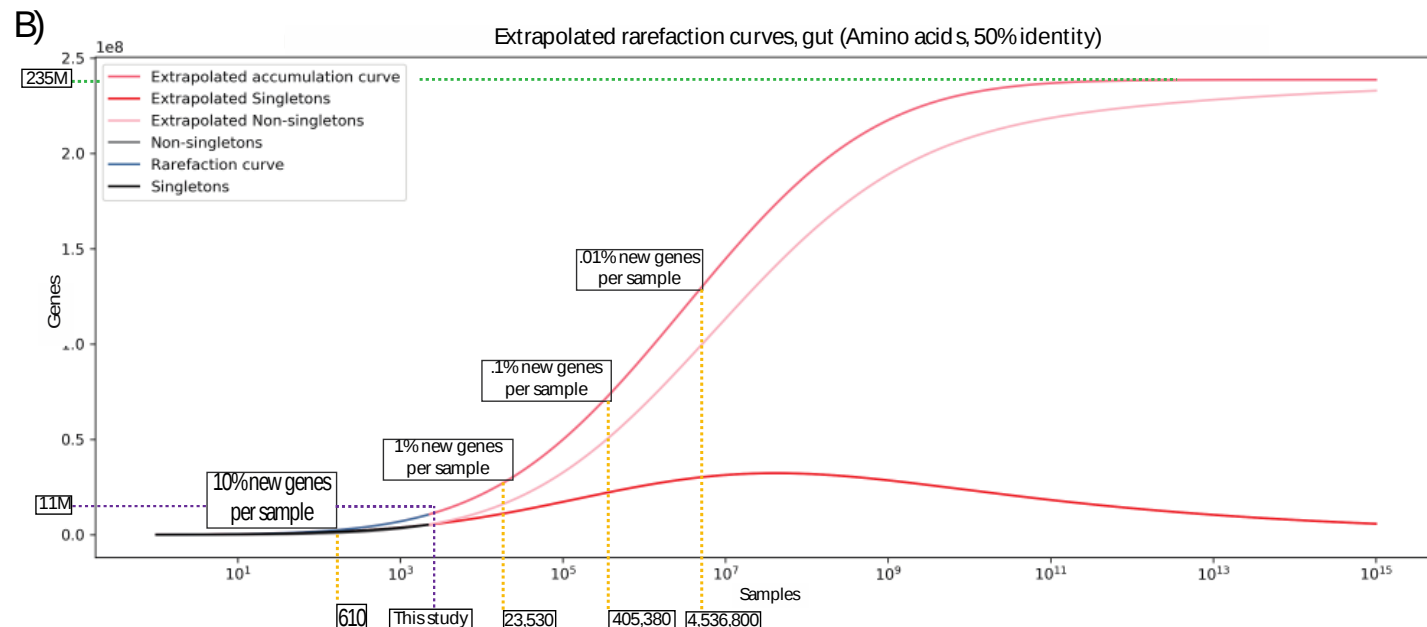
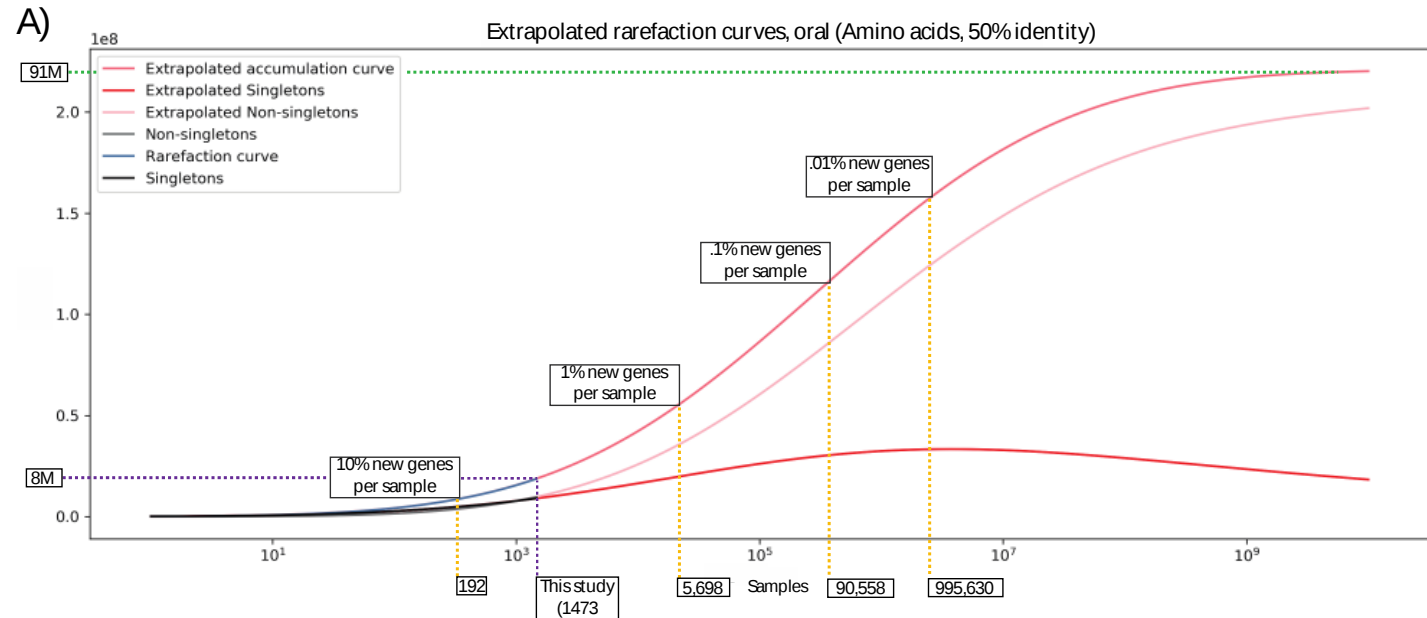


ATGTACGTACCGATCGTGAT

Most studies use 95% nucleotide or
90% amino acid identity to
discriminate between genes



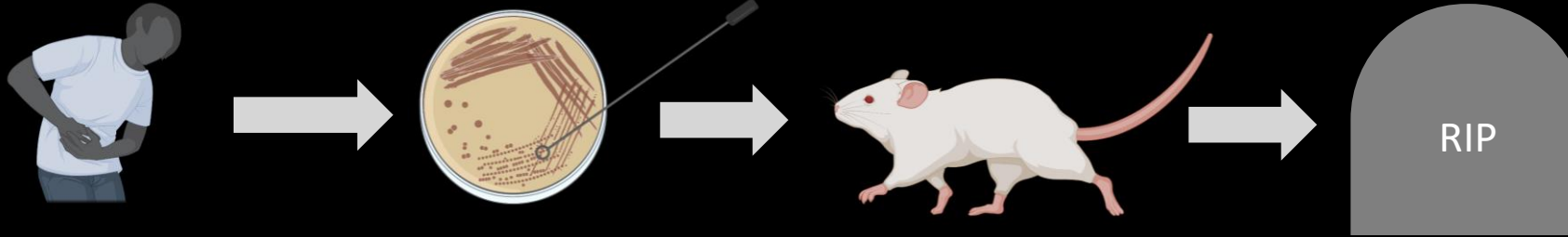
We need much more sequencing to accurately “measure” the microbiome



How do we study the human microbiome?

▶ The experimental scientific method (hypothesis testing)

1 Hypothesis → Experiment (e.g. Koch's postulates)



My thesis

▶ Data-driven “discovery” (hypothesis generation)

Limited causal/clinical
use by itself

“Big data” → Analysis → Many hypotheses



ATGTCGATG	AAAATTGTA	ATCGTGCTA
GTAGGATGA	GCTCGAGCC	ATCGAGTAA
ATGTGACGC	ACGTGATAT	ATCGTGCGT
CGCGCGCAA	ATCGAGTCT	ATCGTGCGT
ATCATTCTCT	ATCGATGCT	ATCGAAAAA
TATATATACC	ATCGATTTT	ATCGCGTGC

How can we
peruse gigantic,
next-generation
DNA sequencing
datasets to form
testable
hypotheses
regarding host-
microbe
interaction?

*** This is generic; specific workflows can vary on the order of steps here and how they are done.**

sequencing facility

@HISEQ2500:282:1:1101:1220:1944 1
ATCGGATCG...
+
<G.<G<AGGII...

demultiplex
(split samples by barodes)

fastqc/multiqc

Some tools:

- trimmomatic
- bbduk.sh (bbtools suite of tools)

>HISEQ2500:282:1:1101:1220:1944 1
ATCGGATCG...

no-assembly path

Analysis

Gene calling

Functional/taxonomic profiling

	Sample_A	Sample_B	...
obj_1	0	428	...
obj_2	306	323	...
obj_3	217	1	...
...	...	...	...

MetaQUAST is a great tool for comparing assemblies

assembly
path

(co)-assembly

Some assemblers and tools:

- Megahit (assembler)
- SPAdes (assembler)
- idba-ud (assembler)
- MetaAMOS (assembler and analysis pipeline)
- MetaCompass (reference-guided)
- MetagenomeScope (visualize assembly graphs)

consider testing assemblies with and w/o

digital normalization

- bbnorm
- diginorm

map individual sample reads to (co)-assembly

Generate coverage information (mapping)

Some tools:

- bowtie2
- bwa

— **A note on MAGs:**

MAGs (metagenome-assembled genomes) are not the same thing as isolate genomes. They are composite representative genomes of closely related genomic lineages.

Recovering genomes from metagenomes

Some tools:

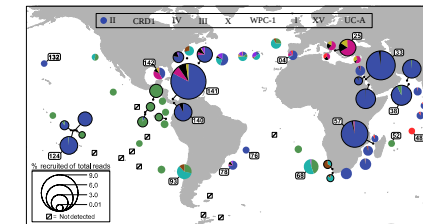
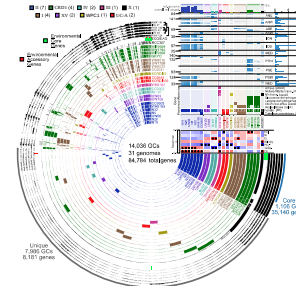
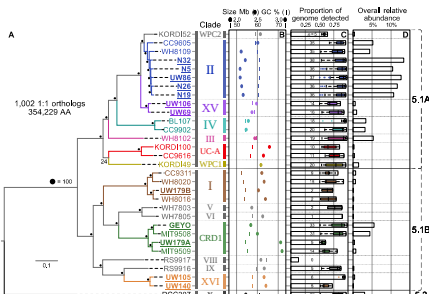
- anvi'o (interactive manual curation of bins; and much more)
- CONCOCT (kmer-based and coverage-based binning; also incorporated in anvi'o)
- COCACOLA (kmer-based, coverage-based, and incorporates paired-read linkage of contigs)
- MetaBAT2 (kmer-based and coverage-based binning tool)
- BinSanity (primarily coverage-based, optional second round kmer-based binning tool)
- checkm (genome-level taxonomy; and much more)
- DASTool (a tool for evaluating bins recovered by different methods)
- DESMAN (tool aimed at resolving strains)

Some common genomics stuff

Phylogenomics
Comparative
genomics
Pangenomics
Env. distributions

Some tools:

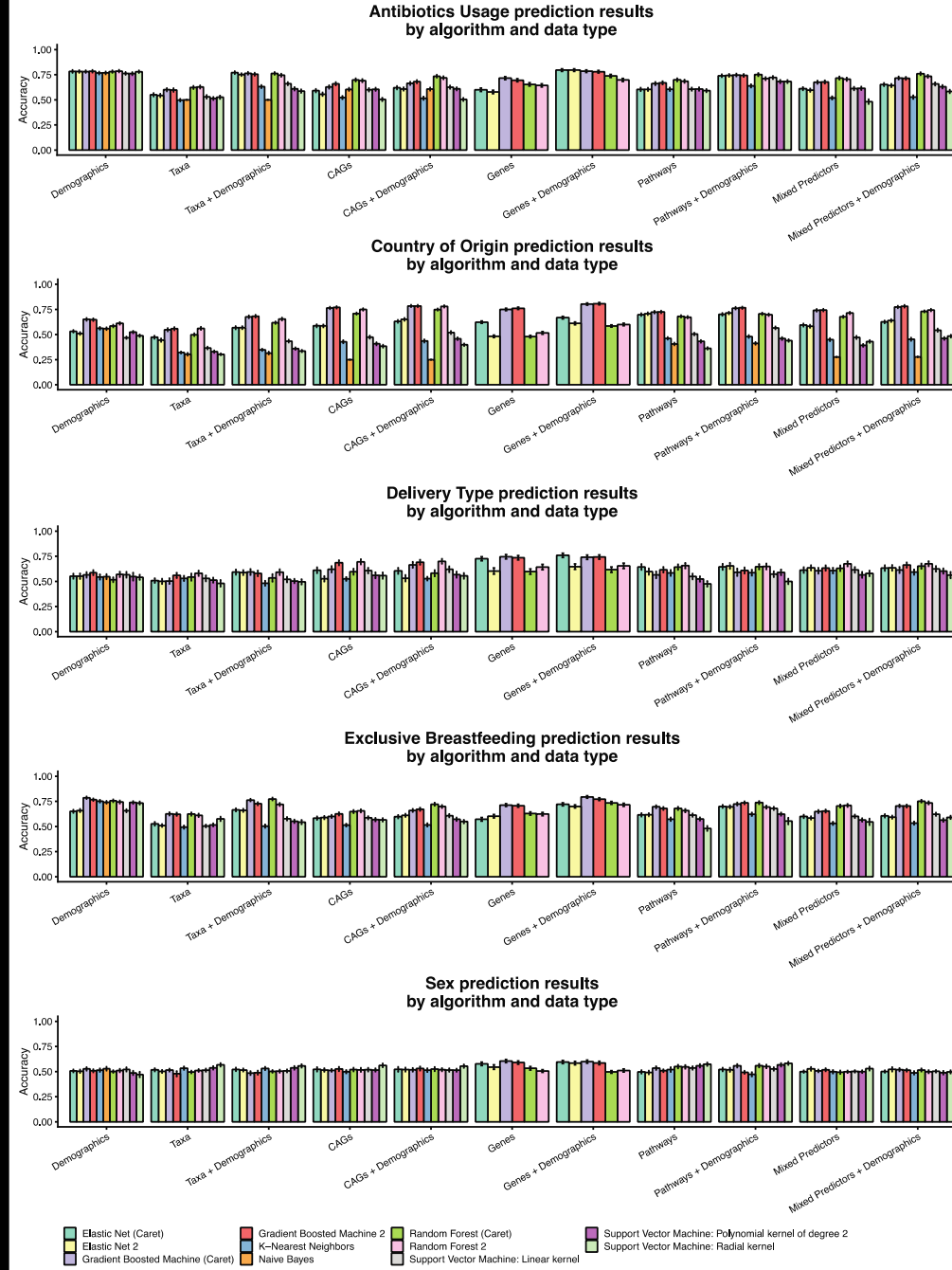
- anvi'o (integrated HMMs for common single-copy gene sets; integrated pangenomic workflow for identifying orthologs via OrthoMCL)
- PanOCT (identifies orthologs utilizing synteny information)
- StrainPhlAn/PanPhlAn (tools for strain-level analyses)
- MUSCLE (alignment software)
- FastTree (very fast, pseudo-maximum likelihood tree builder)
- RAxML (maximum likelihood tree builder)
- Mauve (whole-genome alignment)



astrobiomike.github.io

Genomes

***Different data types
+ different algorithms
=
Different outputs***



***How you ask a
question determines
the answer you get***

Host environment can confound human-microbiome associations

Published: 02 December 2015

Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota

Kristoffer Forslund, Falk Hildebrand, [...] Oluf Pedersen 

Nature **528**, 262–266(2015) | [Cite this article](#)

6192 Accesses | **647** Citations | **343** Altmetric | [Metrics](#)

Article | Published: 06 May 2020

Statin therapy is associated with lower prevalence of gut microbiota dysbiosis

Sara Vieira-Silva, Gwen Falony, [...] Jeroen Raes 

Nature **581**, 310–315(2020) | [Cite this article](#)

12k Accesses | **2** Citations | **324** Altmetric | [Metrics](#)

The coffee problem: a fundamental challenge in microbiome research

CRC Study	Cohort information	Methods	Associated Microbes
Feng et al, 2015	~150 samples, Austrian	Stratified age, gender BMI Wilcoxon, Random Forest Bonferroni FDR	<i>Fusobacterium</i> <i>Parvimonas</i> , <i>Bacteroides</i> , <i>E. coli</i> , <i>Bilophila</i>
Yu et al, 2015	~70 samples, Chinese, Denmark validation set	Self collected samples No adjustments Wilcoxon, Random Forest BH FDR < 0.05	<i>Fusobacterium</i> , <i>Peptostreptococcus</i> , <i>Parvimonas</i> , <i>Solobacterium</i>
Zeller et al, 2014	~140 samples, multiple EU	Wilcoxon, LASSO Checked age, gender BMI BH FDR < 0.1	<i>Fusobacterium</i> , <i>Peptostreptococcus</i> <i>Porphyromonas</i>
Voigtmann et al, 2016	~100 samples, 1980s USA, 2000 French validation	Wilcoxon, Logistic Reg Age, gender BMI adjusted Bonferroni FDR < 0.05	<i>Fusobacterium</i> <i>Porphyromonas</i> , <i>Atopobium</i> , <i>Clostridia</i>

The coffee problem: a fundamental challenge in microbiome research

Cohort heterogeneity

CRC Study	Cohort information	Methods	Associated Microbes
Feng et al, 2015	~150 samples, Austrian	Stratified age, gender BMI Wilcoxon, Random Forest Bonferroni FDR	<i>Fusobacterium</i> <i>Parvimonas</i> , <i>Bacteroides</i> , <i>E. coli</i> , <i>Bilophila</i>
Yu et al, 2015	~70 samples, Chinese , Denmark validation set	Self collected samples No adjustments Wilcoxon, Random Forest BH FDR < 0.05	<i>Fusobacterium</i> , <i>Peptostreptococcus</i> , <i>Parvimonas</i> , <i>Solobacterium</i>
Zeller et al, 2014	~140 samples, multiple EU	Wilcoxon, LASSO Checked age, gender BMI BH FDR < 0.1	<i>Fusobacterium</i> , <i>Peptostreptococcus</i> <i>Porphyromonas</i>
Voigtmann et al, 2016	~100 samples, 1980s USA , 2000 French validation	Wilcoxon, Logistic Reg Age, gender BMI adjusted Bonferroni FDR < 0.05	<i>Fusobacterium</i> <i>Porphyromonas</i> , <i>Atopobium</i> , <i>Clostridia</i>

The coffee problem: a fundamental challenge in microbiome research

Small sample size

CRC Study	Cohort information	Methods	Associated Microbes
Feng et al, 2015	~150 samples, Austrian	Stratified age, gender BMI Wilcoxon, Random Forest Bonferroni FDR	<i>Fusobacterium</i> <i>Parvimonas</i> , <i>Bacteroides</i> , <i>E. coli</i> , <i>Bilophila</i>
Yu et al, 2015	~70 samples, Chinese, Denmark validation set	Self collected samples No adjustments Wilcoxon, Random Forest BH FDR < 0.05	<i>Fusobacterium</i> , <i>Peptostreptococcus</i> , <i>Parvimonas</i> , <i>Solobacterium</i>
Zeller et al, 2014	~140 samples, multiple EU	Wilcoxon, LASSO Checked age, gender BMI BH FDR < 0.1	<i>Fusobacterium</i> , <i>Peptostreptococcus</i> <i>Porphyromonas</i>
Voigtmann et al, 2016	~100 samples, 1980s USA, 2000 French validation	Wilcoxon, Logistic Reg Age, gender BMI adjusted Bonferroni FDR < 0.05	<i>Fusobacterium</i> <i>Porphyromonas</i> , <i>Atopobium</i> , <i>Clostridia</i>

The coffee problem: a fundamental challenge in microbiome research

Variable confounder adjustment

CRC Study	Cohort information	Methods	Associated Microbes
Feng et al, 2015	~150 samples, Austrian	Stratified age, gender, BMI Wilcoxon, Random Forest Bonferroni FDR	<i>Fusobacterium</i> <i>Parvimonas</i> , <i>Bacteroides</i> , <i>E. coli</i> , <i>Bilophila</i>
Yu et al, 2015	~70 samples, Chinese, Denmark validation set	Self collected samples No adjustments Wilcoxon, Random Forest BH FDR < 0.05	<i>Fusobacterium</i> , <i>Peptostreptococcus</i> , <i>Parvimonas</i> , <i>Solobacterium</i>
Zeller et al, 2014	~140 samples, multiple EU	Wilcoxon, LASSO Stratified age, gender, BMI BH FDR < 0.1	<i>Fusobacterium</i> , <i>Peptostreptococcus</i> <i>Porphyromonas</i>
Voigtmann et al, 2016	~100 samples, 1980s USA, 2000 French validation	Wilcoxon, Logistic Reg Stratified age, gender, BMI Bonferroni FDR < 0.05	<i>Fusobacterium</i> <i>Porphyromonas</i> , <i>Atopobium</i> , <i>Clostridia</i>

The coffee problem: a fundamental challenge in microbiome research

Methodological inconsistency

CRC Study	Cohort information	Methods	Associated Microbes
Feng et al, 2015	~150 samples, Austrian	Stratified age, gender BMI Wilcoxon, Random Forest Bonferroni FDR	<i>Fusobacterium</i> <i>Parvimonas</i> , <i>Bacteroides</i> , <i>E. coli</i> , <i>Bilophila</i>
Yu et al, 2015	~70 samples, Chinese, Denmark validation set	Self collected samples No adjustments Wilcoxon, Random Forest BH FDR < 0.05	<i>Fusobacterium</i> , <i>Peptostreptococcus</i> , <i>Parvimonas</i> , <i>Solobacterium</i>
Zeller et al, 2014	~140 samples, multiple EU	Wilcoxon, LASSO Checked age, gender BMI BH FDR < 0.1	<i>Fusobacterium</i> , <i>Peptostreptococcus</i> <i>Porphyromonas</i>
Voigtmann et al, 2016	~100 samples, 1980s USA, 2000 French validation	Wilcoxon, Logistic Reg Age, gender BMI adjusted Bonferroni FDR < 0.05	<i>Fusobacterium</i> <i>Porphyromonas</i> , <i>Atopobium</i> , <i>Clostridia</i>

The coffee problem: a fundamental challenge in microbiome research

Variable results

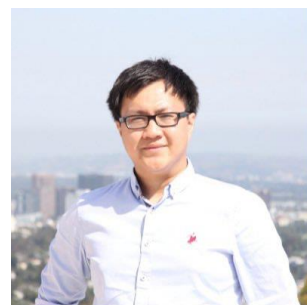
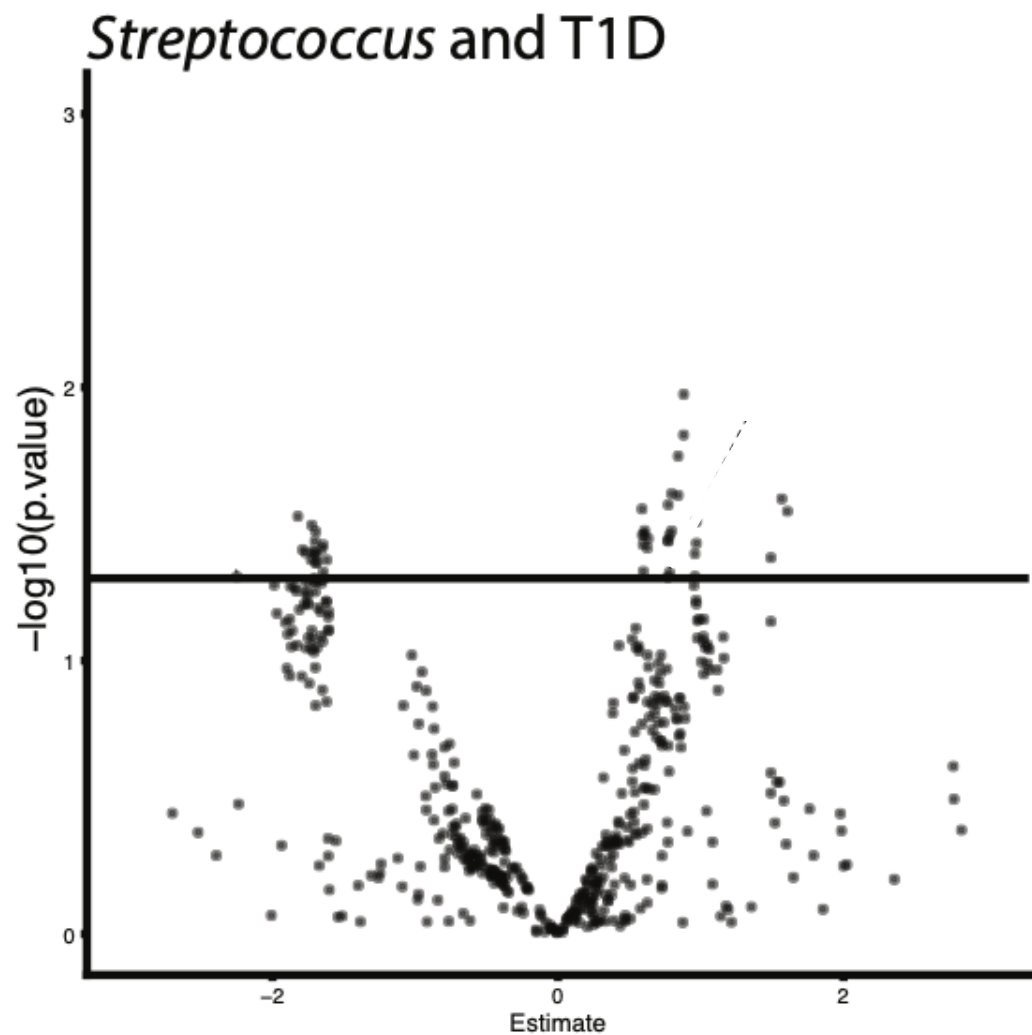
CRC Study	Cohort information	Methods	Associated Microbes
Feng et al, 2015	~150 samples, Austrian	Stratified age, gender BMI Wilcoxon, Random Forest Bonferroni FDR	<i>Fusobacterium</i> <i>Parvimonas</i> , <i>Bacteroides</i> , <i>E. coli</i> , <i>Bilophila</i>
Yu et al, 2015	~70 samples, Chinese, Denmark validation set	Self collected samples No adjustments Wilcoxon, Random Forest BH FDR < 0.05	<i>Fusobacterium</i> , <i>Peptostreptococcus</i> , <i>Parvimonas</i> , <i>Solobacterium</i>
Zeller et al, 2014	~140 samples, multiple EU	Wilcoxon, LASSO Checked age, gender BMI BH FDR < 0.1	<i>Fusobacterium</i> , <i>Peptostreptococcus</i> <i>Porphyromonas</i>
Voigtmann et al, 2016	~100 samples, 1980s USA, 2000 French validation	Wilcoxon, Logistic Reg Age, gender BMI adjusted Bonferroni FDR < 0.05	<i>Fusobacterium</i> <i>Porphyromonas</i> , <i>Atopobium</i> , <i>Clostridia</i>

The coffee problem: a fundamental challenge in microbiome research

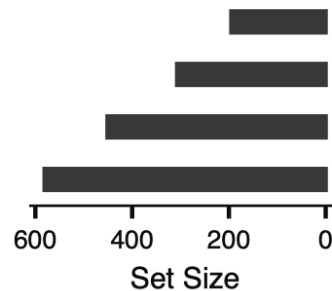
Variable results

CRC Study	Cohort information	Methods	Associated Microbes
Feng et al, 2015	~150 samples, Austrian	Stratified age, gender BMI Wilcoxon, Random Forest Bonferroni FDR	<i>Fusobacterium</i> <i>Parvimonas, Bacteroides, E. coli, Bilophila</i>
Yu et al, 2015	~70 samples, Chinese, Denmark validation set	Self collected samples No adjustments Wilcoxon, Random Forest BH FDR < 0.05	<i>Fusobacterium</i> , <i>Peptostreptococcus</i> , <i>Parvimonas, Solobacterium</i>
Zeller et al, 2014	~140 samples, multiple EU	Wilcoxon, LASSO Checked age, gender BMI BH FDR < 0.1	<i>Fusobacterium</i> , <i>Peptostreptococcus</i> <i>Porphyromonas</i>
Voigtmann et al, 2016	~100 samples, 1980s USA, 2000 French validation	Wilcoxon, Logistic Reg Age, gender BMI adjusted Bonferroni FDR < 0.05	<i>Fusobacterium</i> <i>Porphyromonas, Atopobium, Clostridia</i>

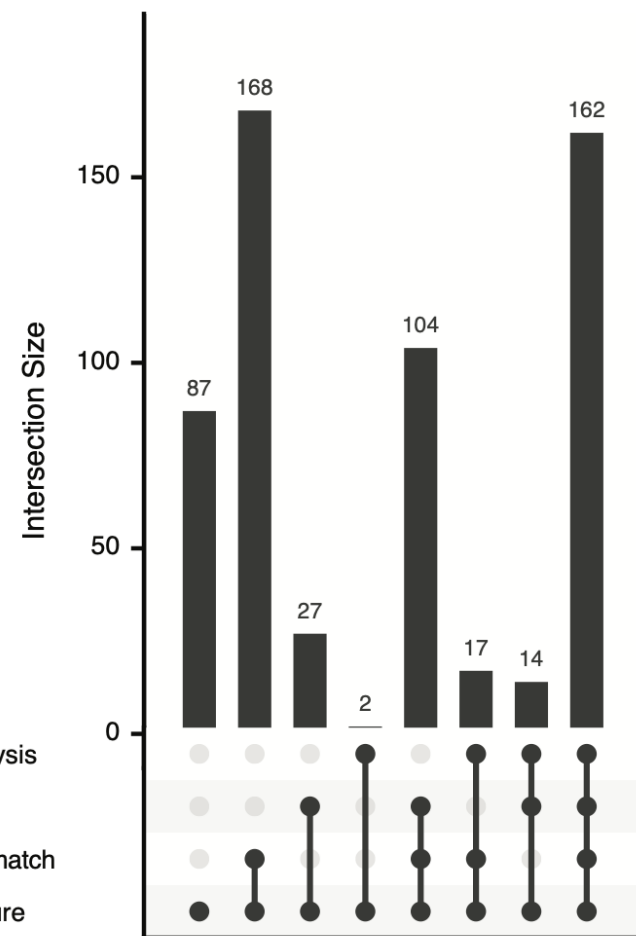
IS there an association between the genus *Streptococcus* and Type 1 diabetes?



~1/3 of reported associated were both robust and significant



Significant in our analysis
No janus effect
Association directions match
Reported in literature



Robust, gene-level,
metagenomic architectures
across 7 diseases yield
high-resolution disease
indicators

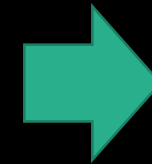


Jason Tan

1.1 million
taxon/pathway/
gene abundances

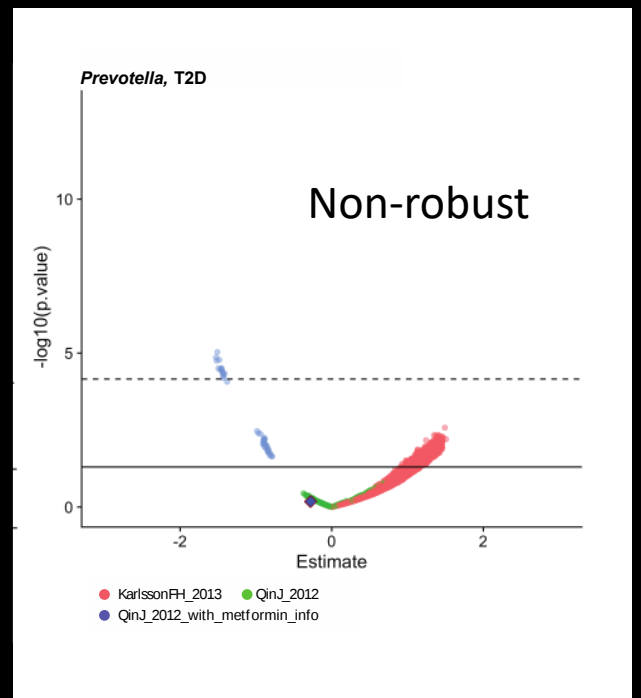
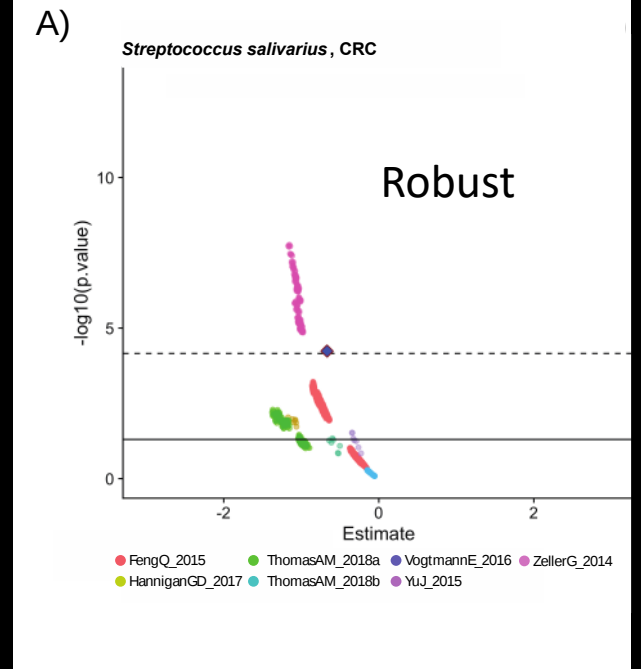


Compute overall
associations and
run vibration of effects
analysis

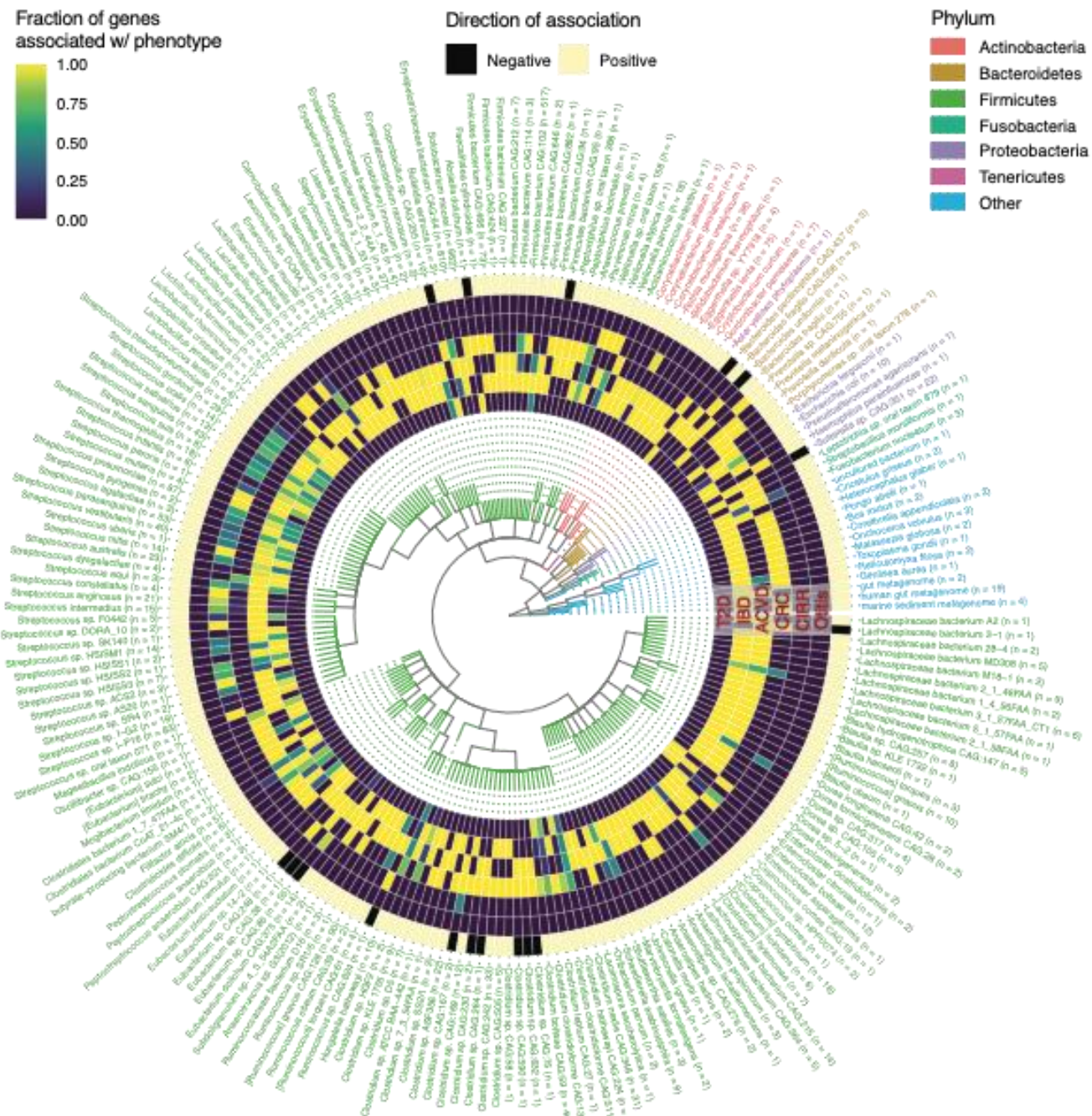


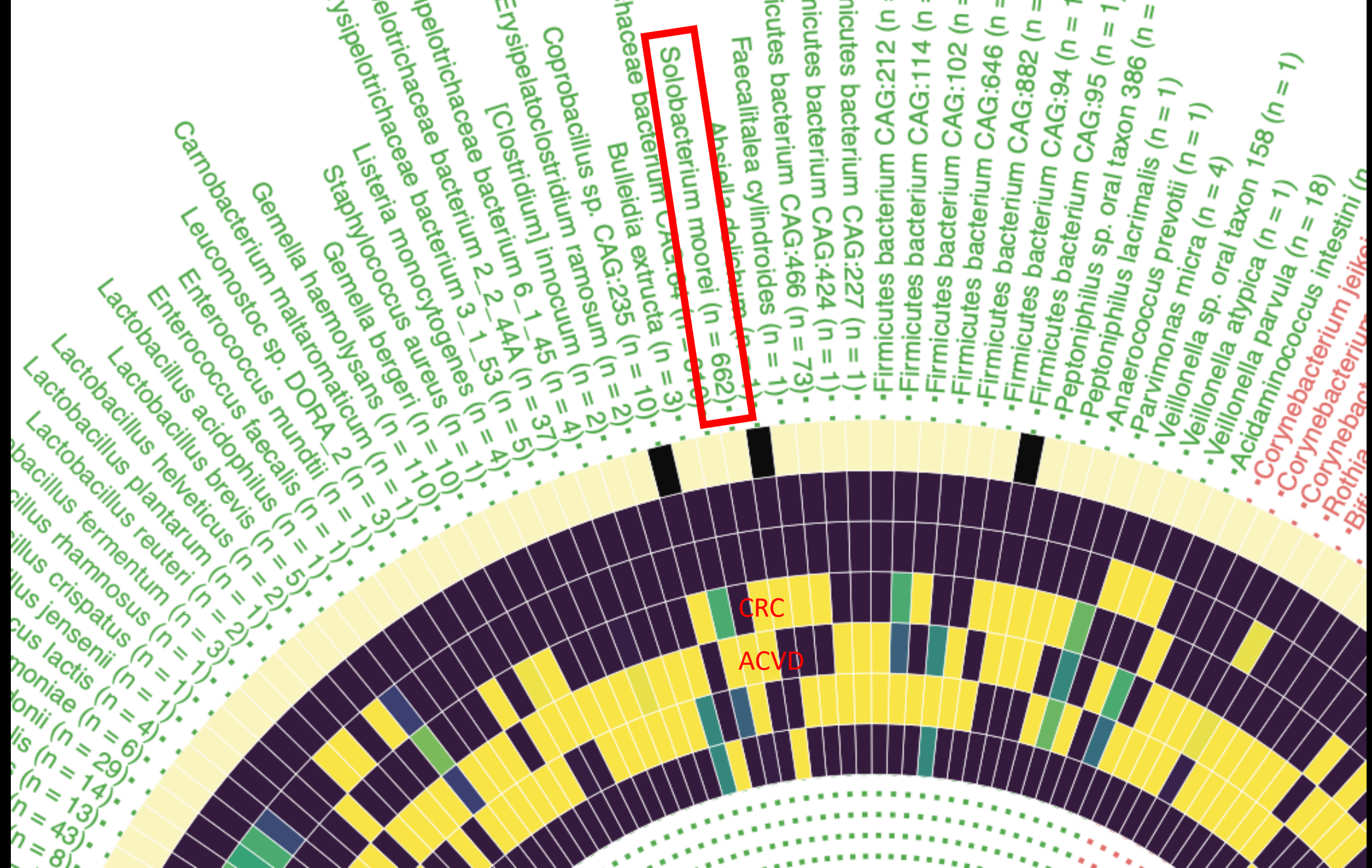
~70 million models fit

B.T. Tierney, Yingxuan Tan, Zhen Yang, A.D. Kostic, C.J. Patel. **Robust, gene-level, metagenomic architectures across 7 diseases yield high-resolution disease indicators.** (in press, *Nature Communications*)



Can we biologically
characterize the genes
associated with at least 2
phenotypes?





How can we peruse gigantic, next-generation DNA sequencing datasets to form testable hypotheses regarding host-microbe interaction?

Deep Gratitude to Many People:



@mason_lab

Mason Lab

Ebrahim Afshinnekoo
Yared Bayleyen
Chandrima Bhattacharya
Daniel Butler
Chris Chin
Rafael Colon
David Danko
Katie Dickieson
Ceyda Durmaz
Radwa Elshafey
Jonathan Foox
Kirill Grigorev
Sonia Iosim
Hannah Kelly
JJ Hastings
Cem Meyden
Chris Mozsary
Jiwoon Park
Jake Reed
Michael Retchin
Krista Ryon
Delia Tomoiaga
Craig Westover
Ben Young
WorldQuant Fellows

HudsonAlpha

Shawn Levy

Cornell/WCM

Alain Borczuk
David Erickson
Selina Chen-Kiang
Iwijn De Vlaminck
Olivier Elemento
Samie Jaffrey
Iman Hajirasouliha
Marcin Imielinski
Ari Melnick (Melnick Lab)
Margaret Ross
Rob Schwartz

Biotia

Niamh O'Hara (Biotia)
Rachid Ounit (Biotia)

NYU

Martin Blaser
Jef Boeke
Jane Carlton
Chris Park
Elizabeth Hénaff

Rockefeller

Jeanne Garbarino
Charles Rice

NYGC

Michael Zody
Dayna M. Oswald
Samantha Fennessey
Soren Gomer
Nicolas Robine
Tom Maniatis

Northwell Hospital

Lance Becker
Peter Gregersen

Miami / UM

George Grills
Helena Solo-Gabriele
Maria E Figueroa

MIT Media Lab

Kevin Slavin
Devora Najjar

FDA/SEQC/Fudan.

Leming Shi

ABRF/Vermont

Scott Tighe
Don Baldwin

Icahn/MSSM

Eric Schadt
Joel Dudley
Bobby Sebra

Illumina

Gary Schroth
Marc Van Oene

Duke

Stacy Horner
Nandan Gokhale

Jackson Labs

Sheng Li

Baylor

Fritz Sedlazeck
Jeff Rogers

MSKCC

Alex Kentsis
Christina Leslie
Ross Levine

AMNH

George Amato

Nanostring

Tyler Hether
Joe Beechem
Sarah Warren

New England Biolabs

Eileen Dimalanta
Nathan Tanner
Ted Davis
Fiona Stewart

UVA

Francine Garrett-Bakelman

NASA/JPL/Ames

Afshin Beheshti
Kate Rubins
Craig Kundrot
David Smith
Kasthuri Venkateswaran

Univ. Chicago

Yoav Gilad

UCSF

Charles Chiu

Acknowledgements



Dissertation Advisory and Thesis Committees: Matthew Meyerson, Mike Gilmore, Shamil Sunyaev, Ami Bhatt, Georg Gerber, Jon Clardy

Patel Lab	Kostic Lab	Ben Kent
Chirag Patel	Aleksandar Kostic	Michaela Walker
Chirag Lakhani	Sam Zimmerman	Bing Shui
Danielle Rasooly	Tao Xu	James Henriksen
Jason Tan	Marsha Wibowo	Erika Szymanski
Nam Pho	Ted Chavkin	Mike Gilmore
Elizabeth Anderson	Marc Beaudin	
Sivateja Tangirala	Melissa Tran	
Jake Chung	Zahir Islam	
Alan Le Gallec	Zhen Yang	
Andrew Deonarine	Other	
Christina Baek	George Church	
Kajal Claypool	Eran Segal	
Shreyas Bhav	Arjun Manrai	



Funding:

- NIH/ BIRT T15LM007092
- Microsoft Research
- NIH/ NIEHS R00 ES023504 & R21 ES025052
- NSF Big Data Spoke Grant
- NIAID R01R01 AI127250 02
- GCP/AWS/Oracle

