

PRECISION MICROBIAL GENOMICS: MEASURING MICROBIOME TRANSMISSION IN THE IMMUNOCOMPROMISED HOST



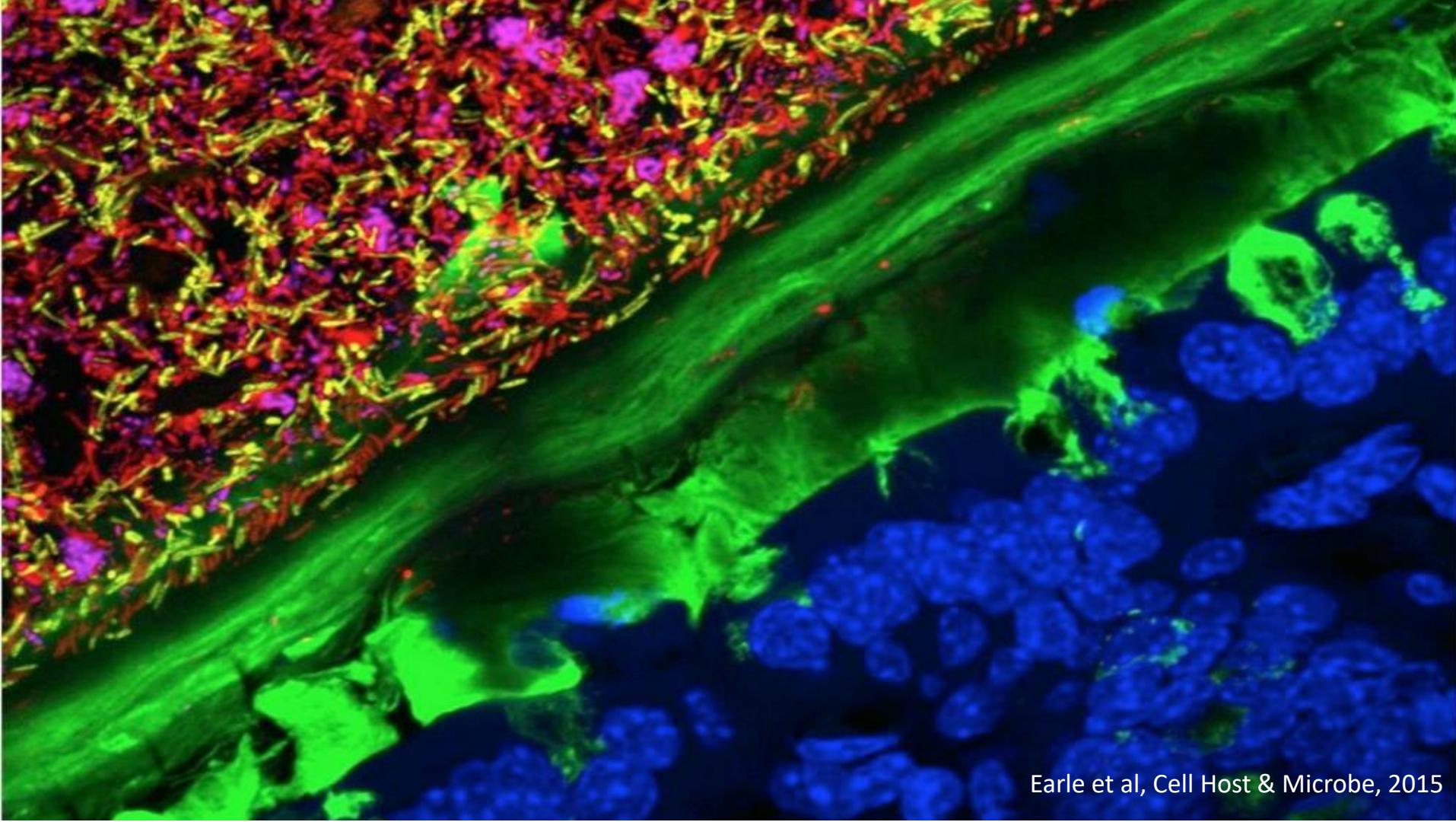
Ami S. Bhatt, MD, PhD
Stanford | bhattlab.com

The Barcelona Debates on the Human Microbiome 2019
June 21, 2019 – Barcelona, Spain

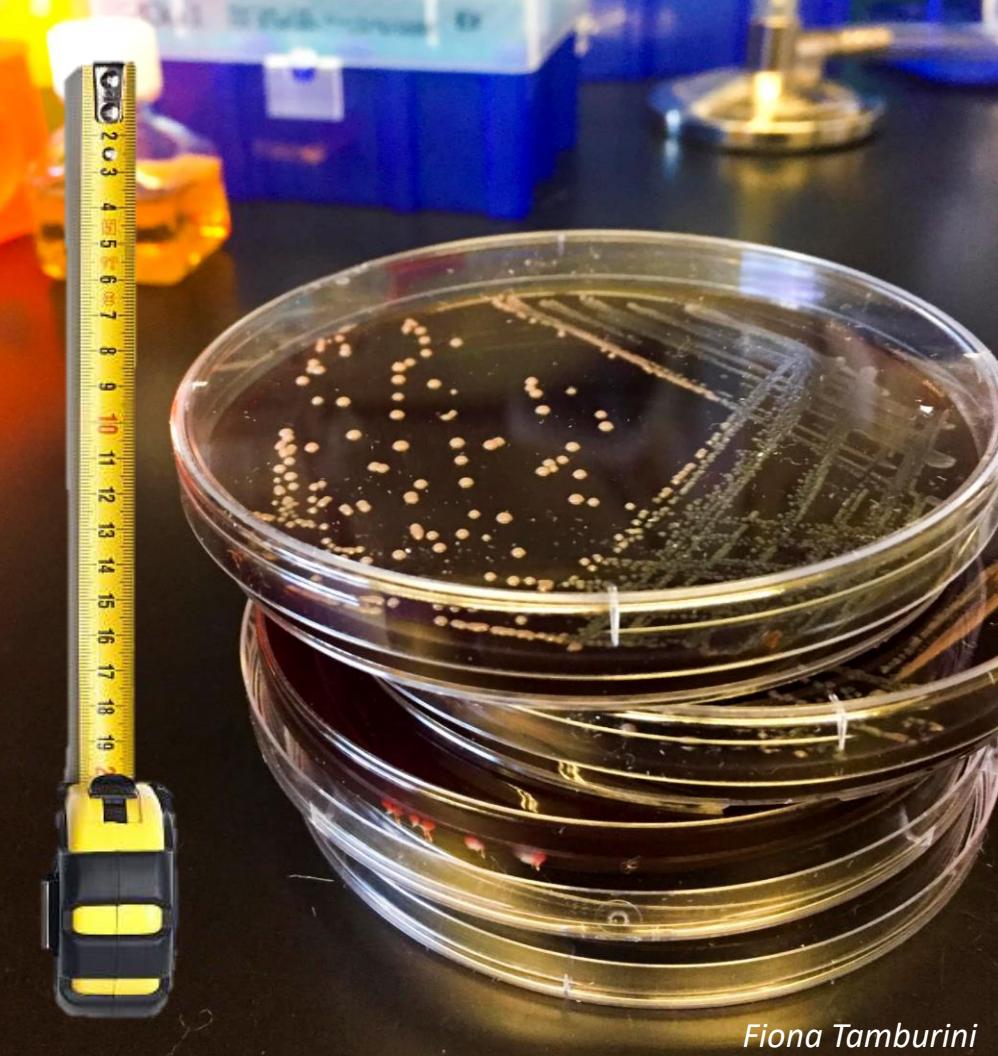
Disclosures

- Scientific Advisory Board: ArcBio, January.ai, Caribou Biosciences
- Paid Consultant: Kaleido Biosciences, Janssen Human Microbiome Institute
- Honoraria: Illumina
- Research collaboration without funding support: 10x Genomics, Illumina
- Research funding support: Agilent
- Nonprofit Boards: Global Oncology, Inc

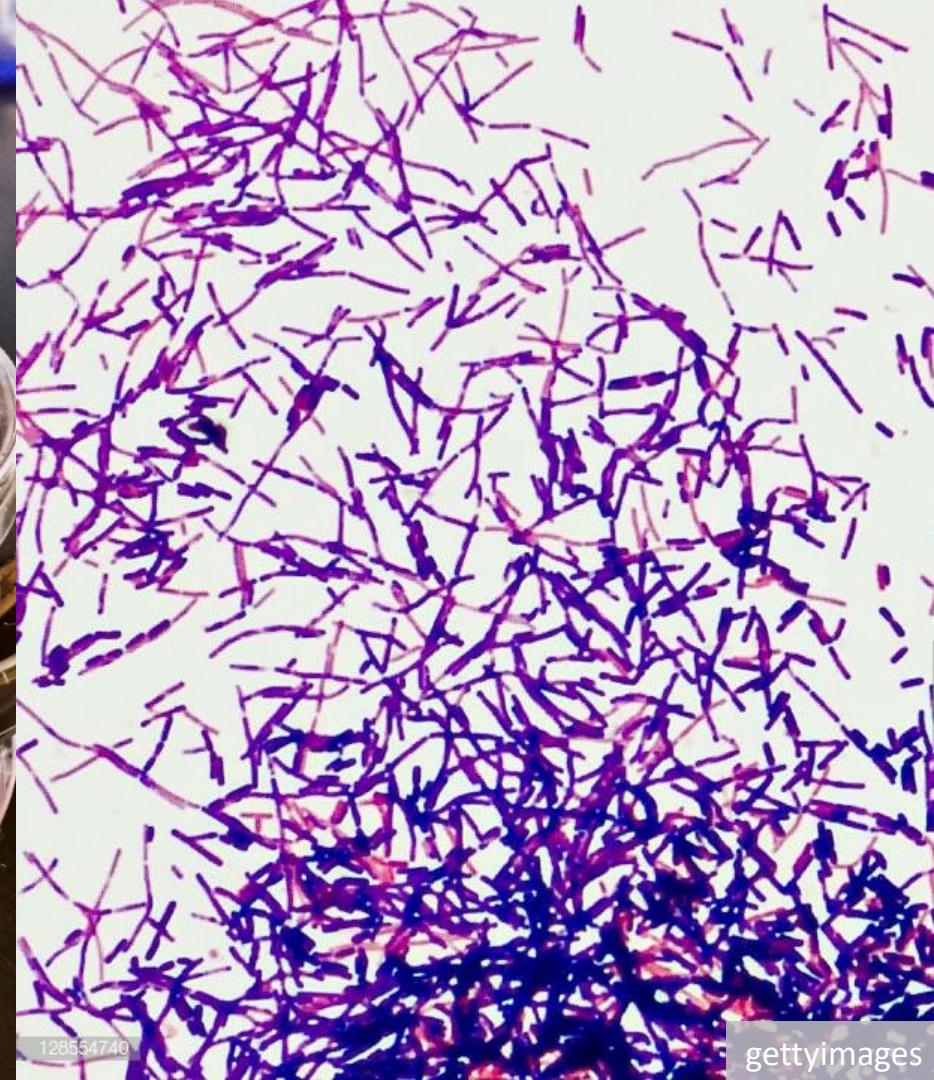




Earle et al, *Cell Host & Microbe*, 2015



Fiona Tamburini



128554740

gettyimages



GTCACCACTATATGGAGAGACCCTGATCTATTCTAGAGCTTTCC
GAGGTAGAGATTCCCCATATGCCCTATTGGGGCTCACACGCTCA
AGAGATTCCCCATATGCCCTATTGGGGCTCACACGCTCACGTCA



Most microbes have yet to be discovered

#1

**WE CAN SAVE CANCER PATIENT LIVES
BY MANIPULATING THE MICROBIOME**





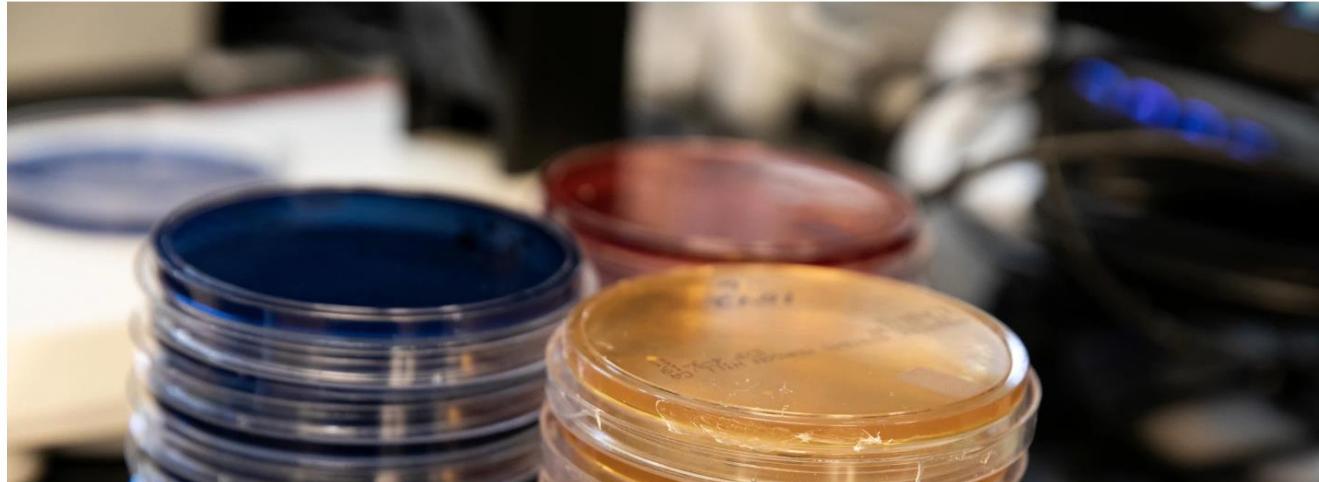


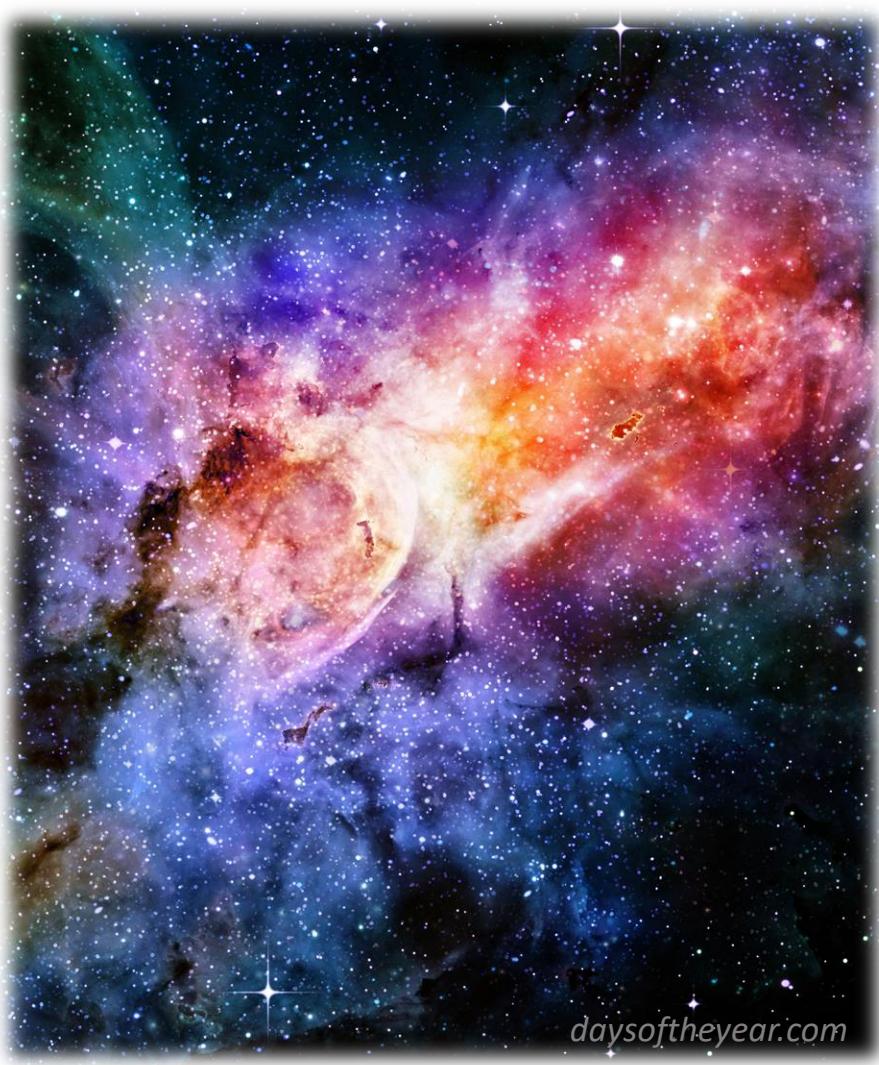
?



Fecal Transplant Is Linked to a Patient's Death, the F.D.A. Warns

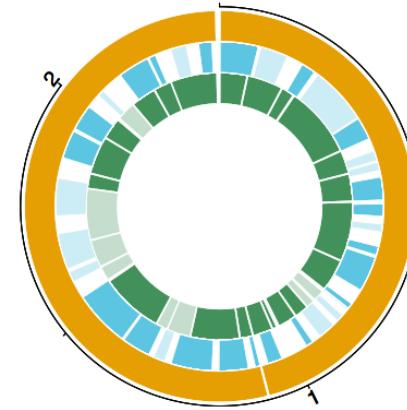
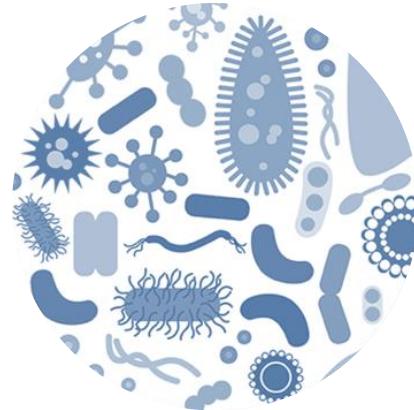
The agency said two patients received donated stool that had not been screened for drug-resistant germs, leading it to halt clinical trials until researchers prove proper testing procedures are in place.





daysoftheyear.com

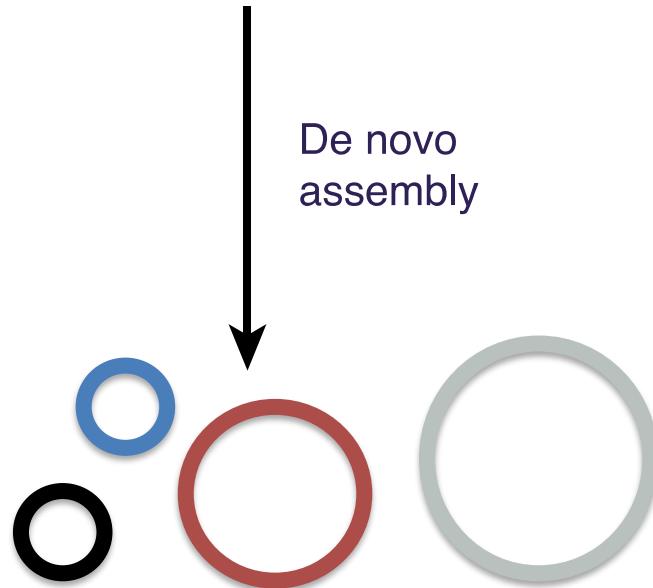
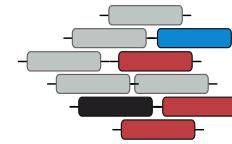




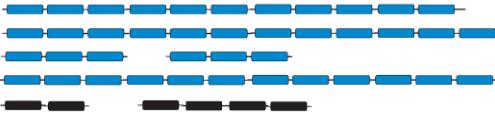
#1

GENOMES FROM METAGENOMES: ILLUMINATING THE DARK MATTER

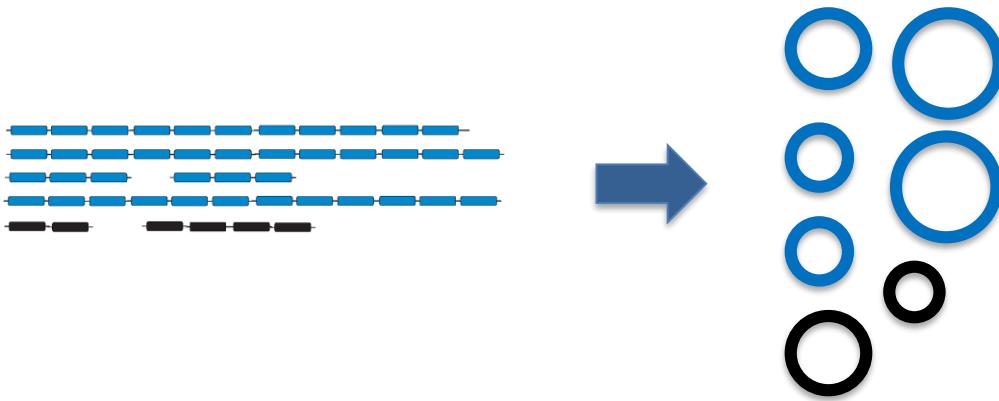
de novo ASSEMBLY:
GENOMES from
METAGENOMES



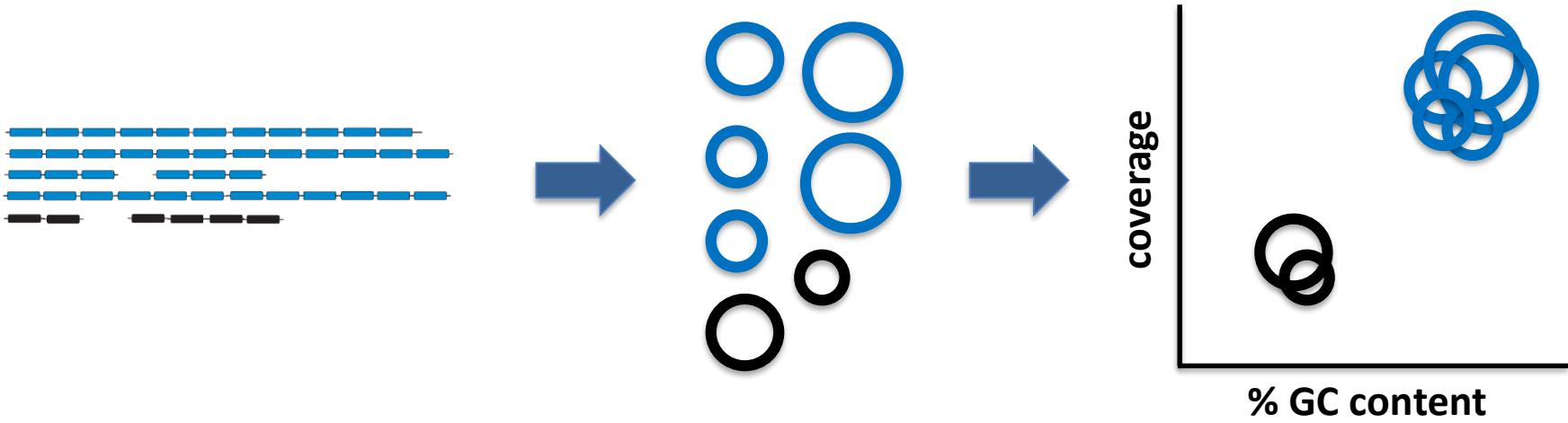
Generating *de novo* genomes (metagenome-assembled genomes)



Generating *de novo* genomes (metagenome-assembled genomes)

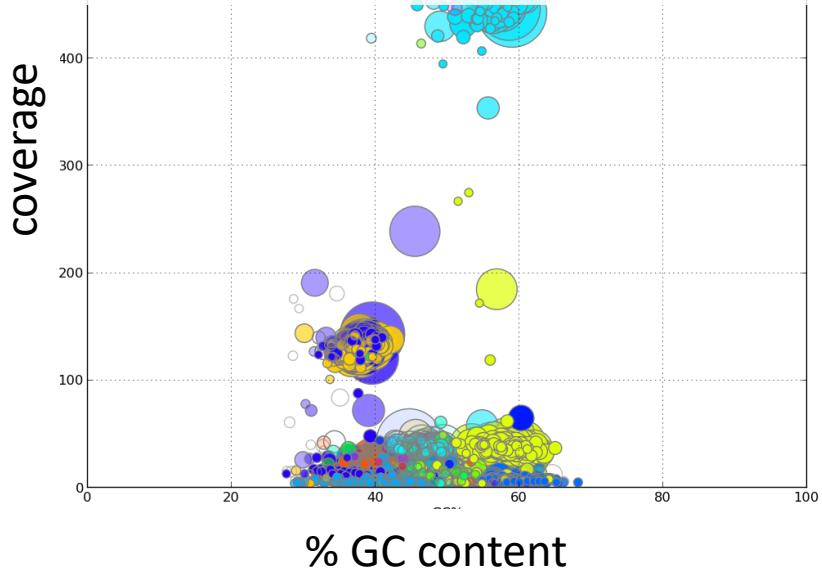


Generating *de novo* genomes (metagenome-assembled genomes)



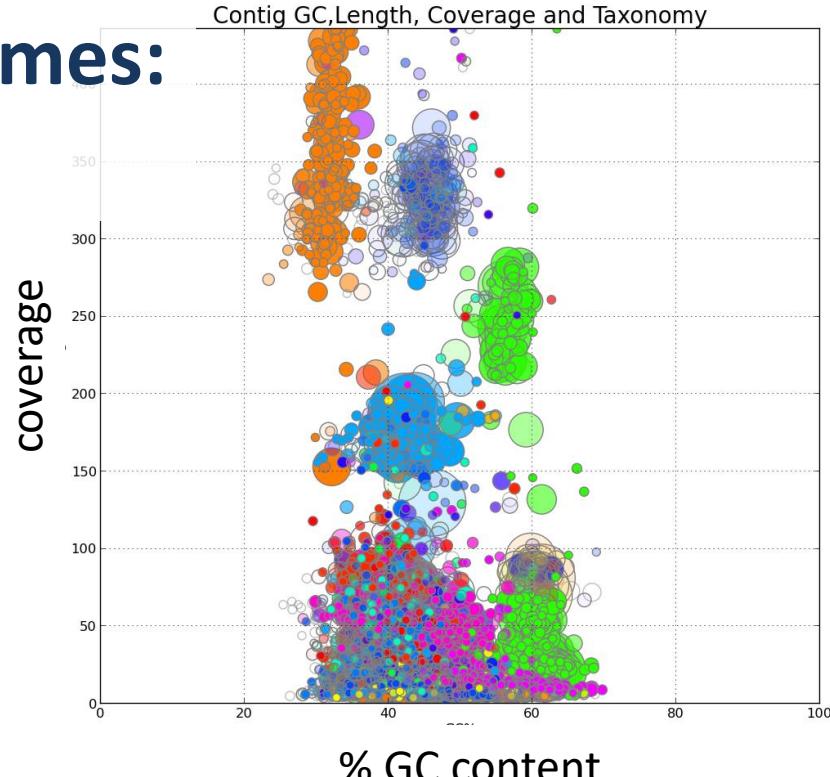
Contig GC,Length, Coverage and Taxonomy

Generating *de novo* genomes: a real life example



GVHD

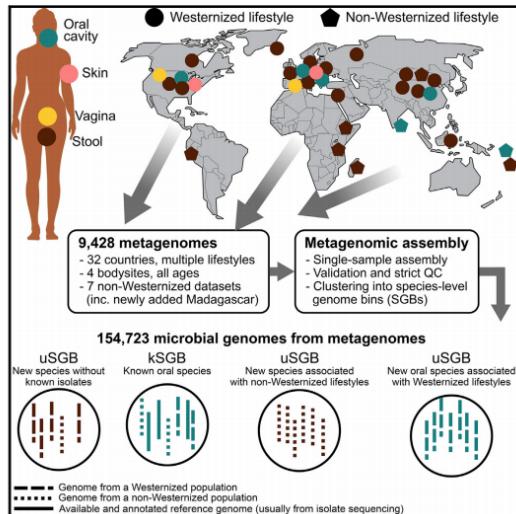
Contig GC,Length, Coverage and Taxonomy



No colitis

Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle

Graphical Abstract



Authors

Edoardo Pasolli, Francesco Asnicar,
Serena Manara, ..., Christopher Quince,
Curtis Huttenhower, Nicola Segata

Correspondence

nicola.segata@unitn.it

ARTICLE

OPEN

<https://doi.org/10.1038/s41586-019-1058-x>

New insights from uncultivated genomes of the global human gut microbiome

Stephen Nayfach^{1,2*}, Zhou Jason Shi^{3,4}, Rekha Seshadri^{1,2}, Katherine S. Pollard^{3,4,5,6,7,8} & Nikos C. Kyrpides^{1,2*}

ARTICLE

OPEN

<https://doi.org/10.1038/s41586-019-0965-1>

A new genomic blueprint of the human gut microbiota

Alexandre Almeida^{1,2*}, Alex L. Mitchell¹, Miguel Boland¹, Samuel C. Forster^{2,3,4}, Gregory B. Gloor⁵, Aleksandra Tarkowska¹, Trevor D. Lawley² & Robert D. Finn¹

METAGENOME-ASSEMBLED GENOMES

	Nayfach et al	Pasolli et al	Almeida et al
Datasets	15 studies of GMB from 15 countries; Westernized urban + non-Western rural/hunter-gatherer pops	46 studies from 31 countries; 5 body sites; + new GMB samples from Madagascar & Ethiopia	75 GMB studies from European Nucleotide Archive – most from Europe, N. Amer, Asia
Assembly	MegaHIT	metaSPAdes or MegaHIT. Reads mapped w/ Bowtie2	metaSPAdes. Reads mapped w/ BWA
Binning	DAS Tool (MaxBin + MetaBAT2 + CONCOCT) Contigs \geq 1kb	MetaBAT2. Contigs \geq 1kb	MetaBAT2. Contigs \geq 2kb
Quality	CheckM	CheckM, CMSeq (strain het)	CheckM
Grouping	Mash2.0 distance $<5\%$ \rightarrow species-level OTUs	Mash2.0 genetic distance + clustering \rightarrow “SGBs” ($\sim 5\%$ gen diversity)	Mash2.0 clustering \rightarrow individual metagenomic species (MGS)
Final	3,810 metagenomes \rightarrow 60,664 genomes (40% “HQ”) \rightarrow 2,058 new species	9,428 metagenomes \rightarrow 154,723 genomes (45% “HQ”) \rightarrow 4,930 species (77% unknown)	11,850 metagenomes \rightarrow 92,143 genomes (43% “near-complete”) \rightarrow 1,952 uncultured species

HUMAN GUT MICROBIOMES:

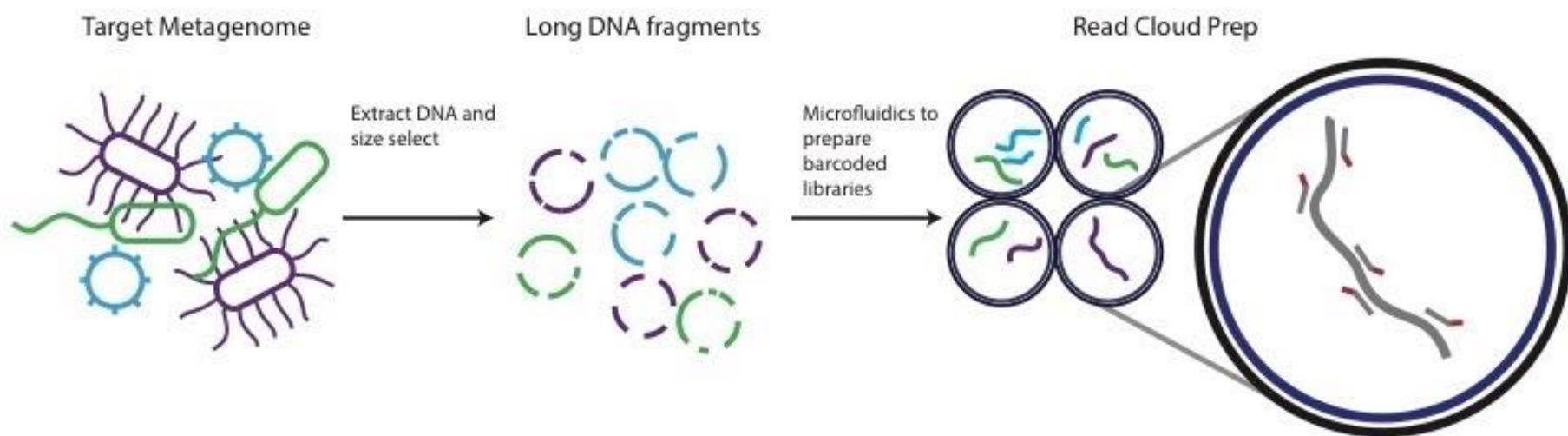
STANDARD APPROACHES GENERATE **HIGHLY FRAGMENTED GENOMES**



Phascolarctobacterium sp.

- Read Cloud
- SLR
- Short Read

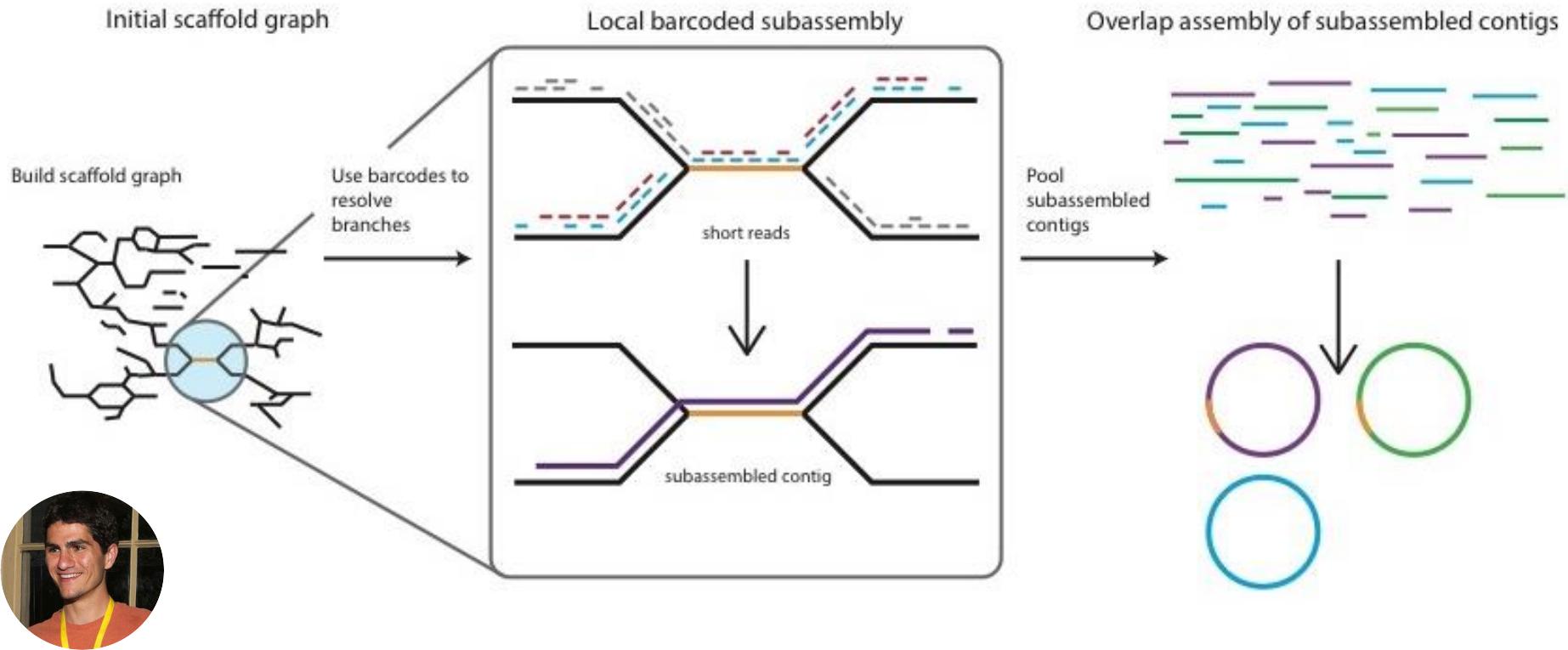
SOLUTION 1a: DNA partitioning (10x Chromium)



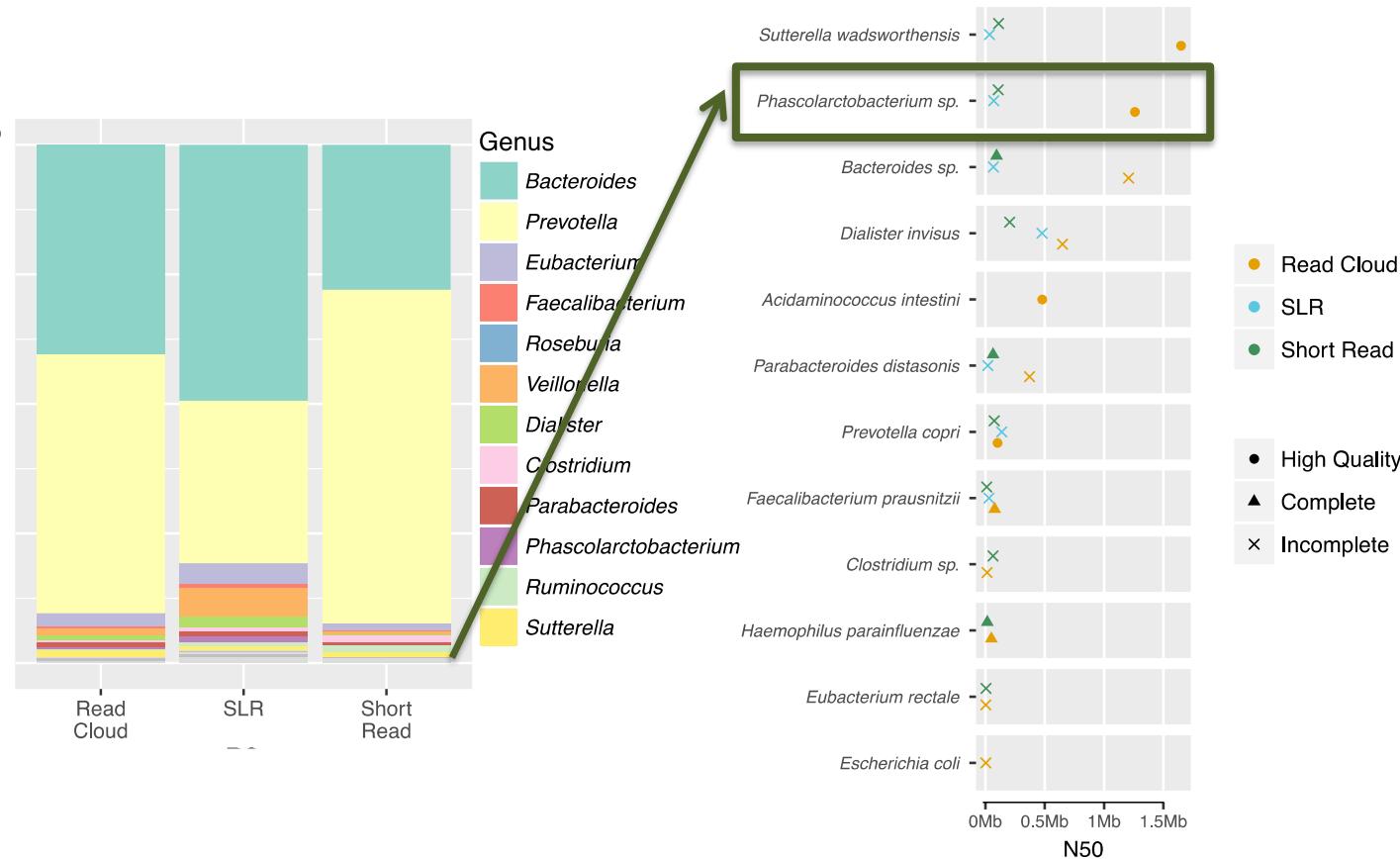
Commercially available technology

No software for metagenomic assembly application

SOLUTION 1b: Athena Assembly

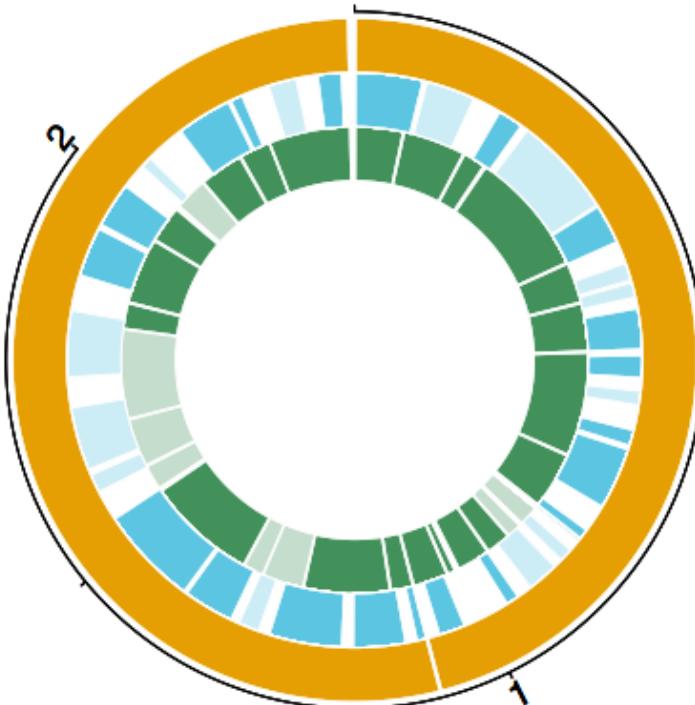


GOOD NEWS: Read clouds assemble some things well



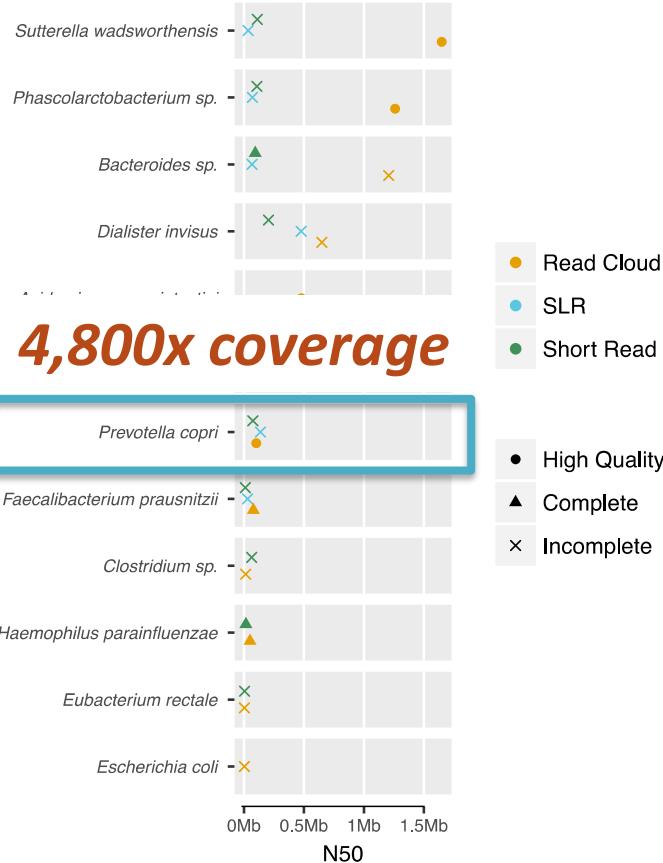
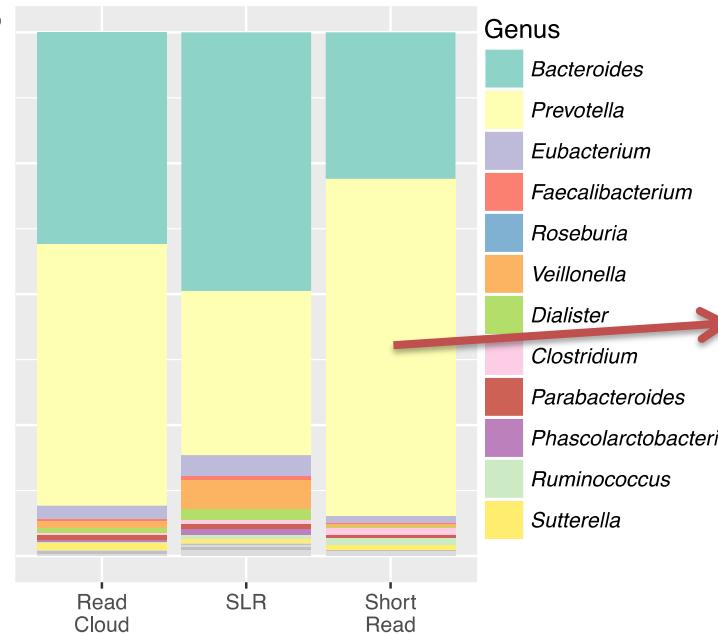
HUMAN GUT MICROBIOMES:

READ CLOUD SEQUENCING & ATHENA ASSEMBLY OUTPERFORMS THE COMPETITION

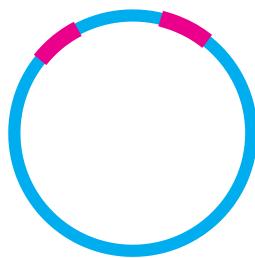


Phascolarctobacterium sp.

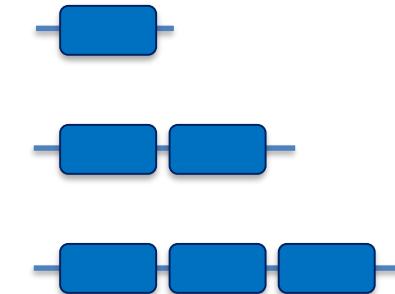
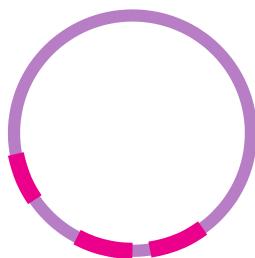
BAD NEWS: Read clouds assemble some things poorly



PREVOTELLA GENOME ASSEMBLY IS HARD BECAUSE OF PREVALENT REPEATED SEQUENCES



16S rRNA operon



Gene duplication

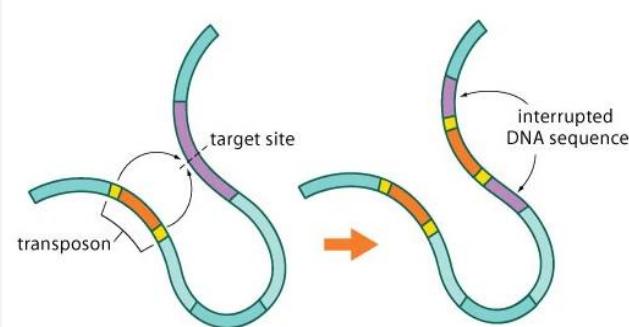
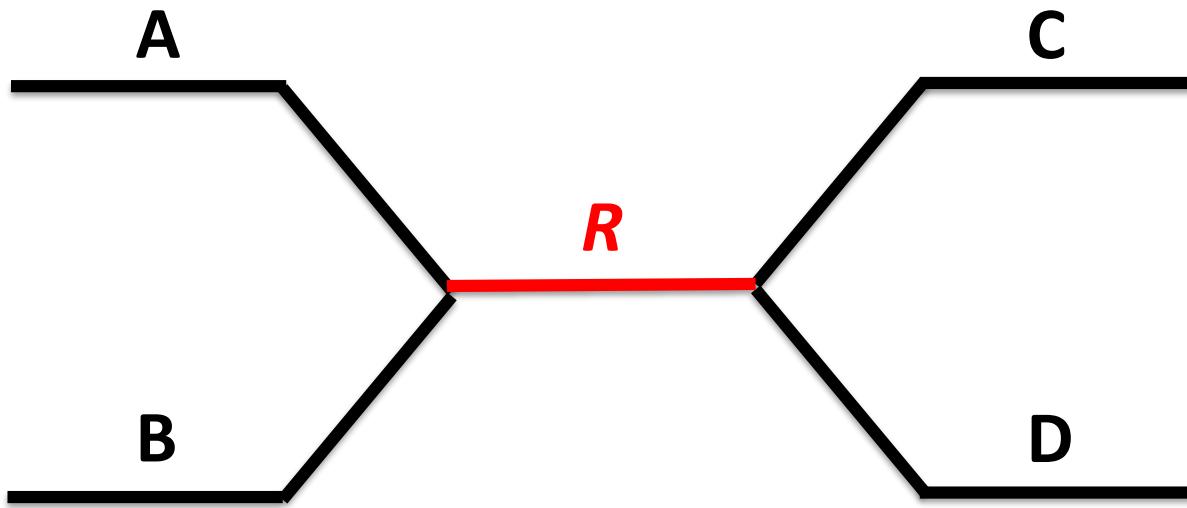
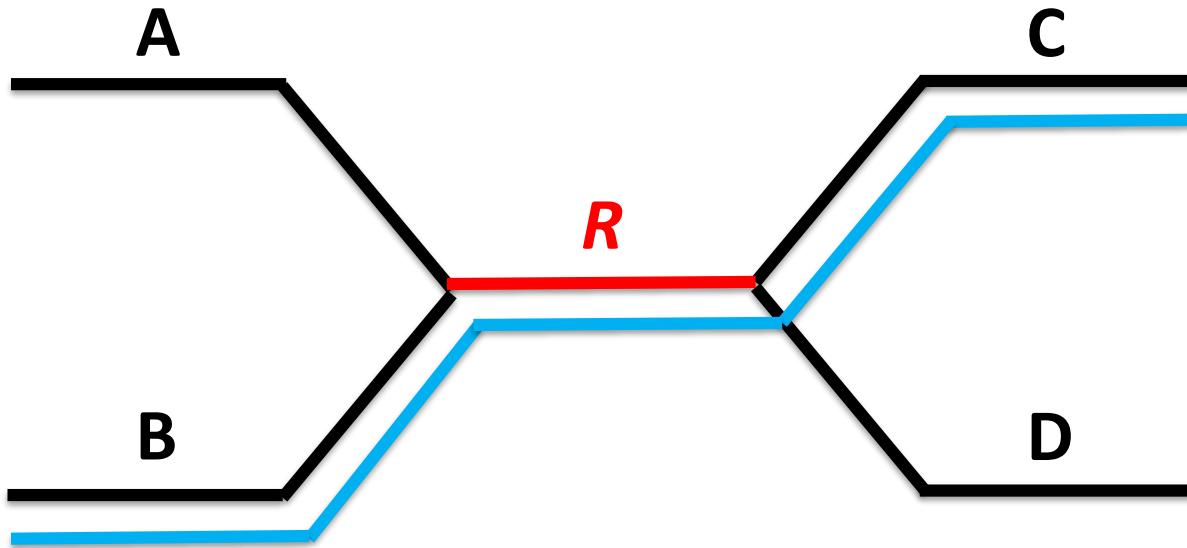


Image courtesy of Broad Institute

Transposons / IS



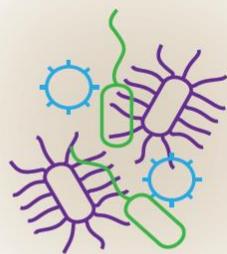
A-*R*-C
A-*R*-D
B-*R*-C
B-*R*-D



Long read sequencing

A-*R*-D
B-*R*-C

SOLUTION – STEP 1: Develop new method for high-yield, HMW DNA Extraction



Enzymatic cell wall
degradation



Phenol-chloroform
extraction



Proteinase K + RNase A
digestion



Gravity column
purification

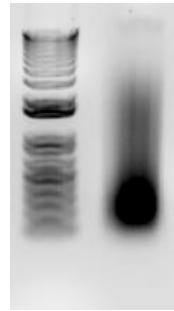


SPRI bead
size selection



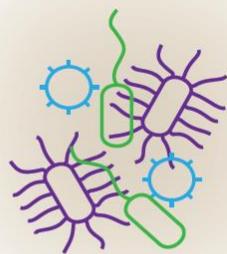


PROBLEM: Low yield, low MW DNA from stool





SOLUTION – STEP 1: Develop new method for high-yield, HMW DNA Extraction



Enzymatic cell wall
degradation



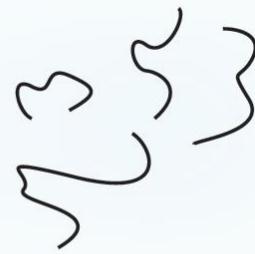
Phenol-chloroform
extraction



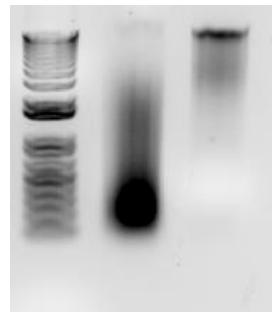
Proteinase K + RNase A
digestion

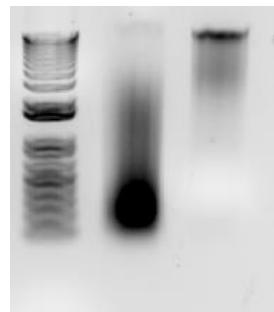


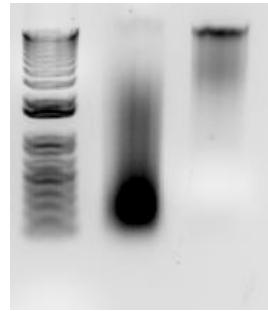
Gravity column
purification



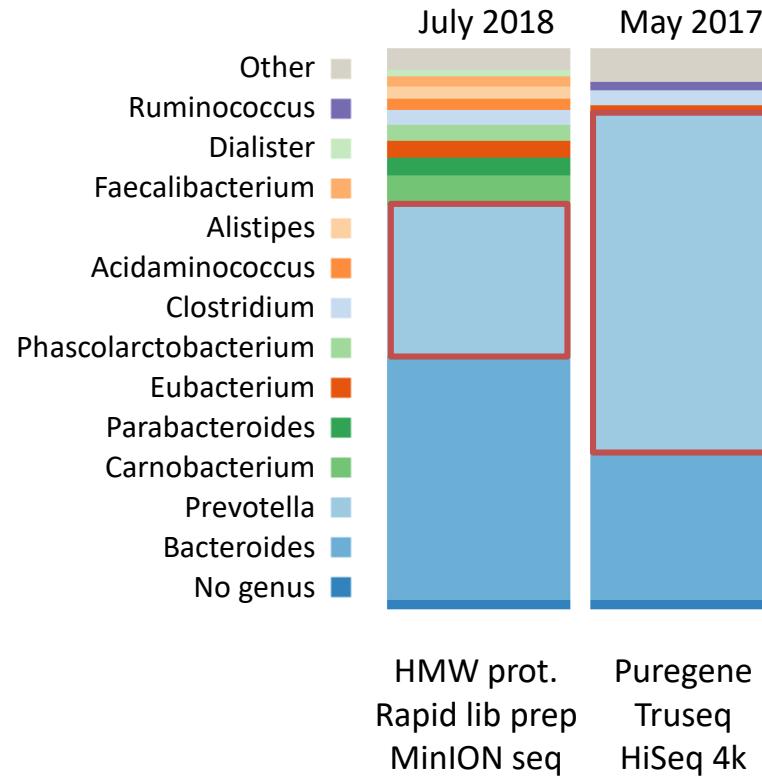
SPRI bead
size selection



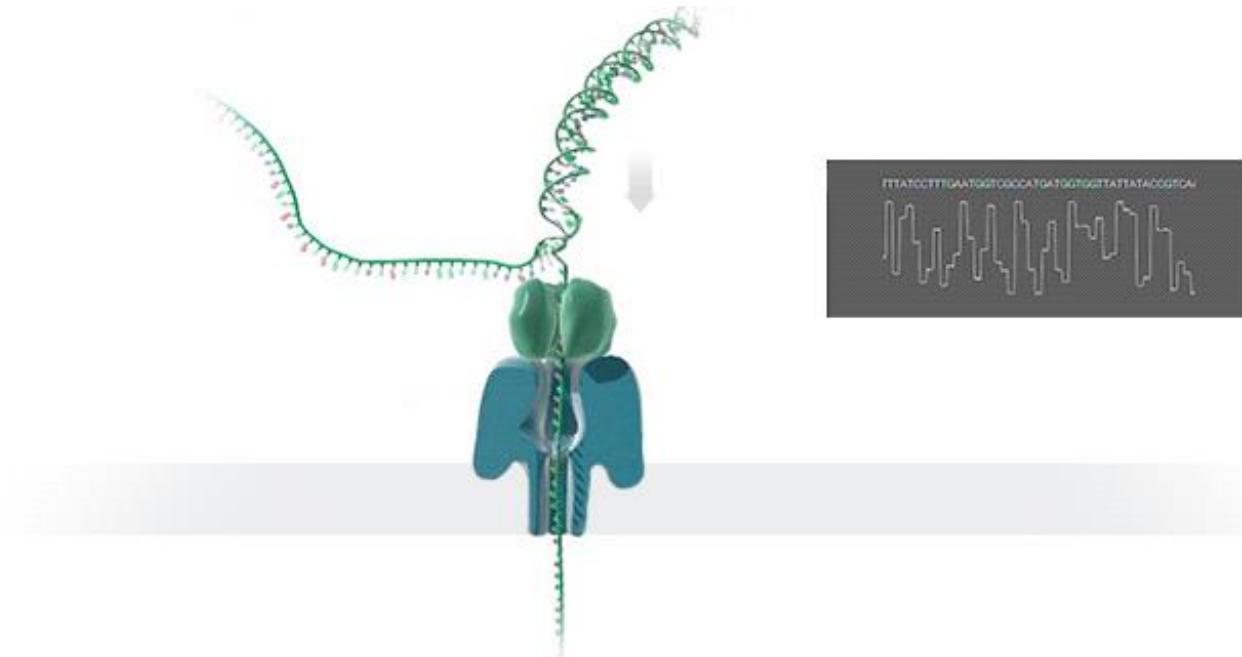




THERE ARE STILL *PREVOTELLA* IN THERE...

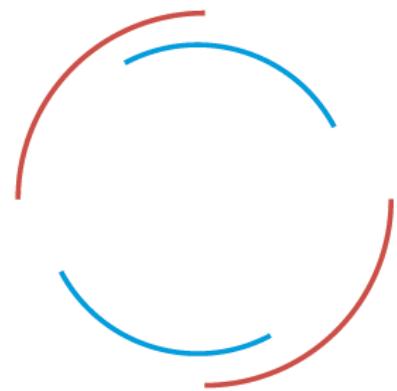


Nanopore sequencing – long reads

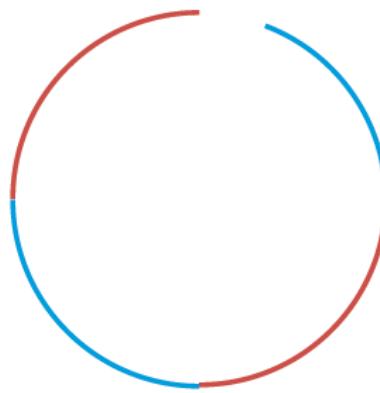


Oxford nanopore
technologies

STEP 2: COMPUTATIONAL WORKFLOW



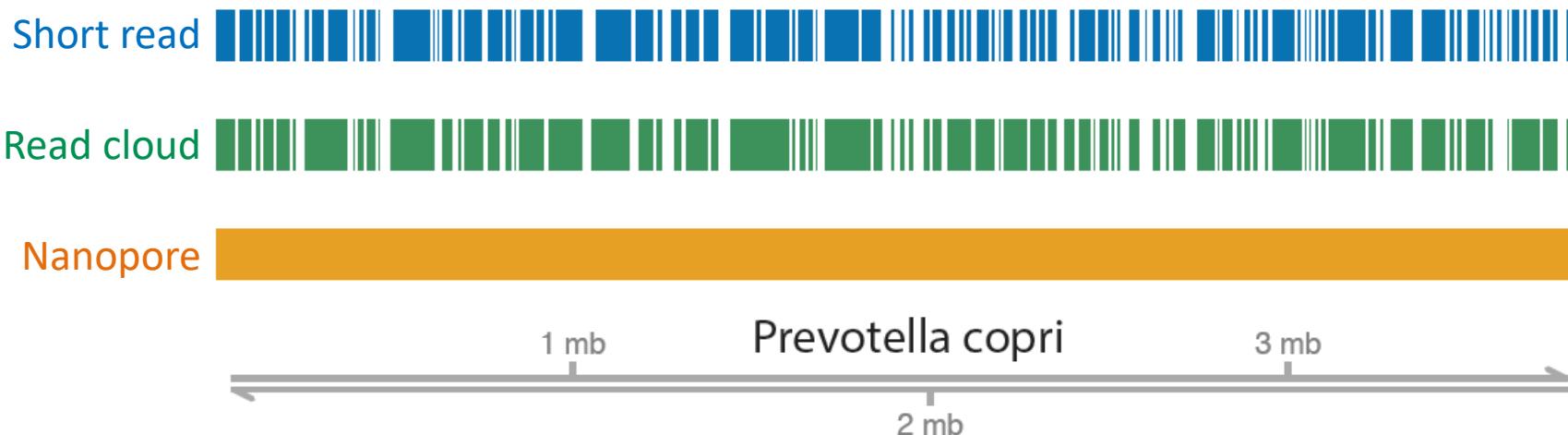
Twofold assembly



Merging

Assembler	Total sequence	Total Assembly Size	Assembly N50	Prevotella N50	Runtime	Prevotella Coverage
Unicycler Ir	0.5Gb	13Mb	0.9Mb	0.9Mb	0.4 hrs	16x
Unicycler Ir bold	0.5Gb	13Mb	0.9Mb	0.9Mb	0.2 hrs	16x
Flye	0.5Gb	20.1Mb	1Mb	3.8Mb, single contig	0.3 hrs	16x
Canu	0.5Gb	16Mb	3.8Mb	3.8Mb, single contig	3 hrs	16x
Spades Ir + sr	0.5Gb + 40Gb	148Mb	0.011Mb	0.12Mb	13 hrs	16x + 2,800x
Unicycler Ir+sr	0.5Gb + 40Gb	48Mb	0.1Mb	0.926Mb	40 hrs	16x + 2,800x
Athena	40Gb	142Mb	0.012Mb	0.1Mb	24+ hrs	2,800x

NANOPORE SEQUENCING & CANU ASSEMBLY GENERATES THE FIRST FULL LENGTH *PREVOTELLA COPRI* GENOME



PROBLEM: POOR GENOME QUALITY



- Measures “Completeness” and “Contamination” of a genomes
- Completeness: % core genes present
- Contamination: % single copy core genes present in copy # > 1
- **Low levels of “completeness”**

SOLUTION: SHORT READ POLISHING WITH PILON

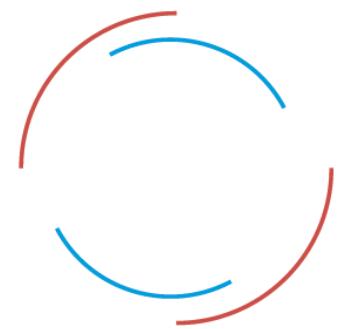


- Short read “polishing” substantially improves genome accuracy
- Improved prediction of core genes

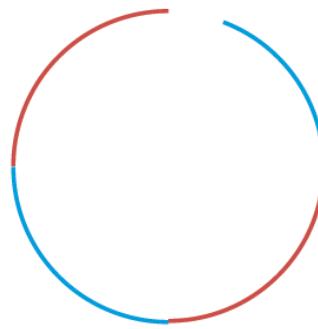


- **High levels of “completeness”**

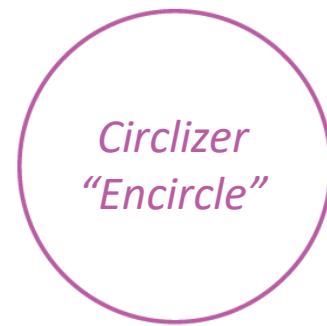
STEP 3: Refined Computational Workflow



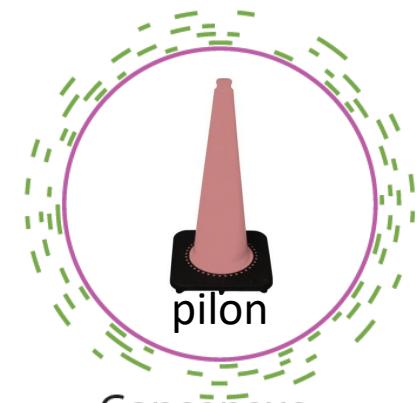
Twofold assembly



Merging



Circularization



Consensus
refinement

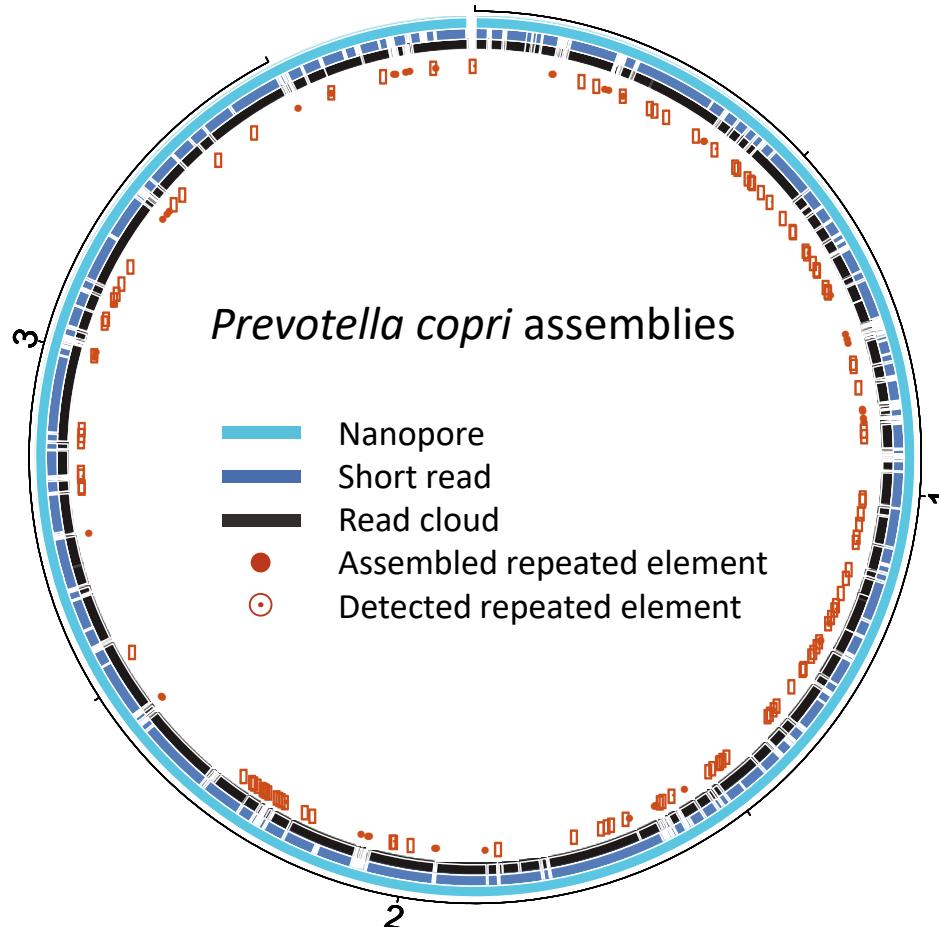
INSERTION SEQUENCES BREAK SHORT READ ASSEMBLY AND BINNING

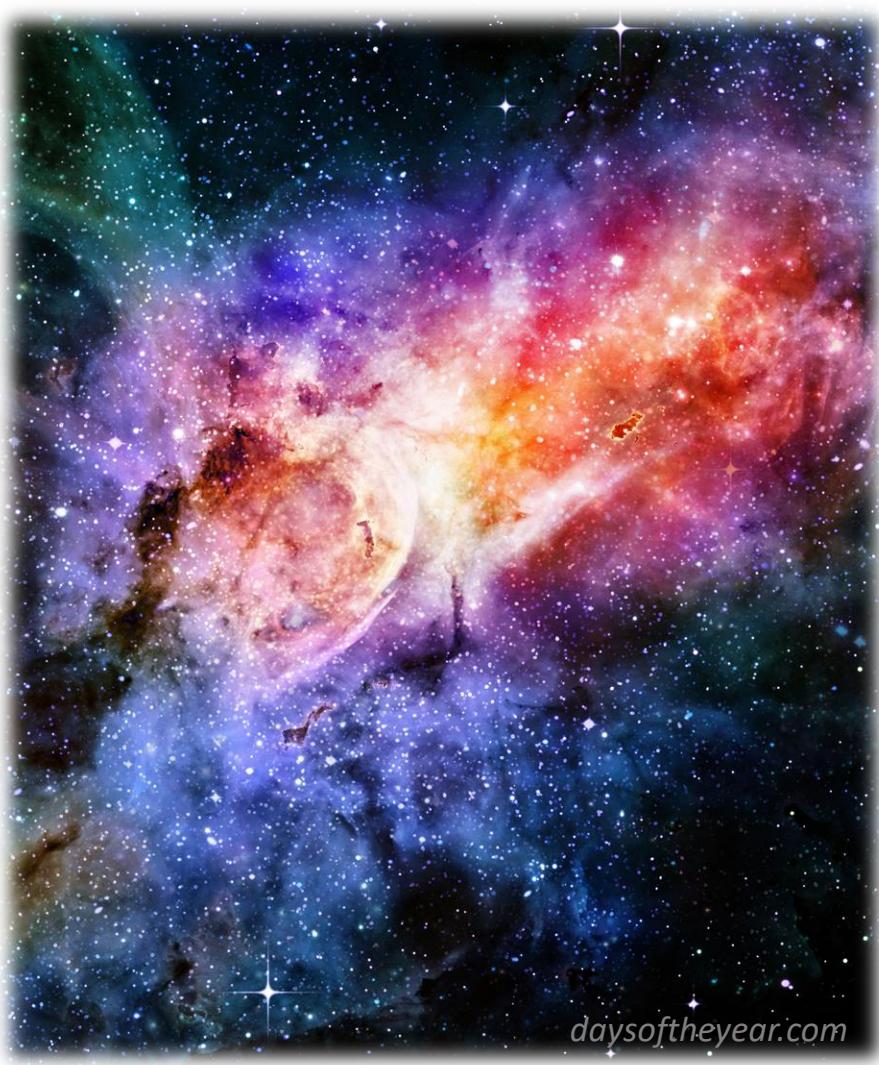
Short reads:

- small contigs split across 5 bins
- 2196 genes in total
- no 16S

Nanopore:

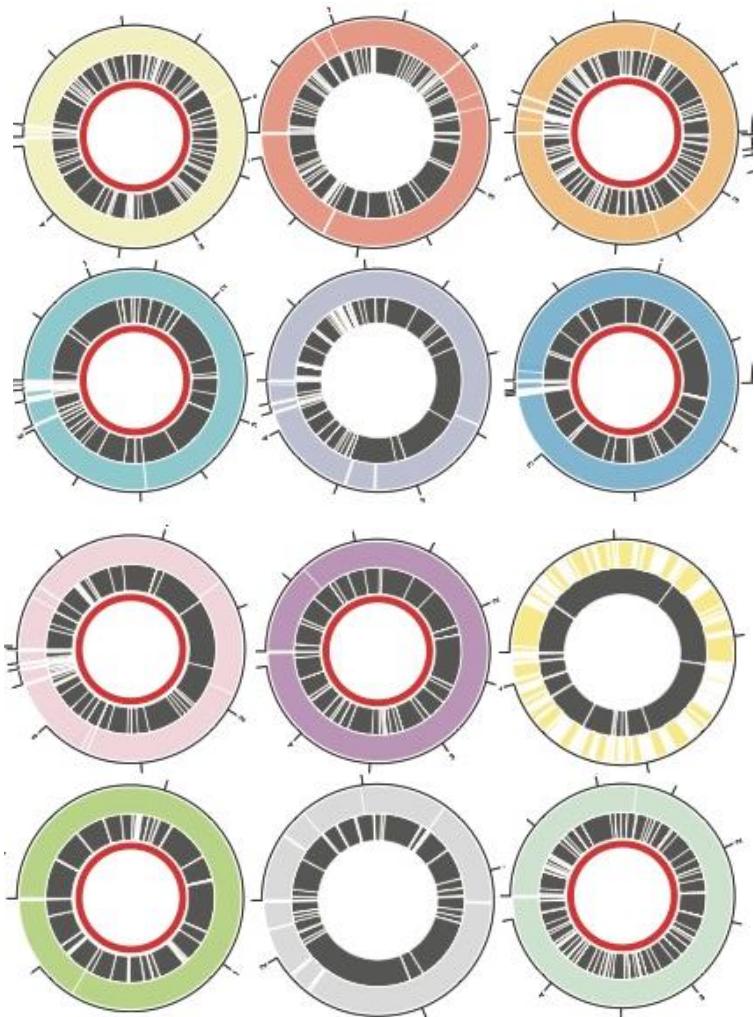
- closed genome
- 3131 genes total
- 5 16S loci



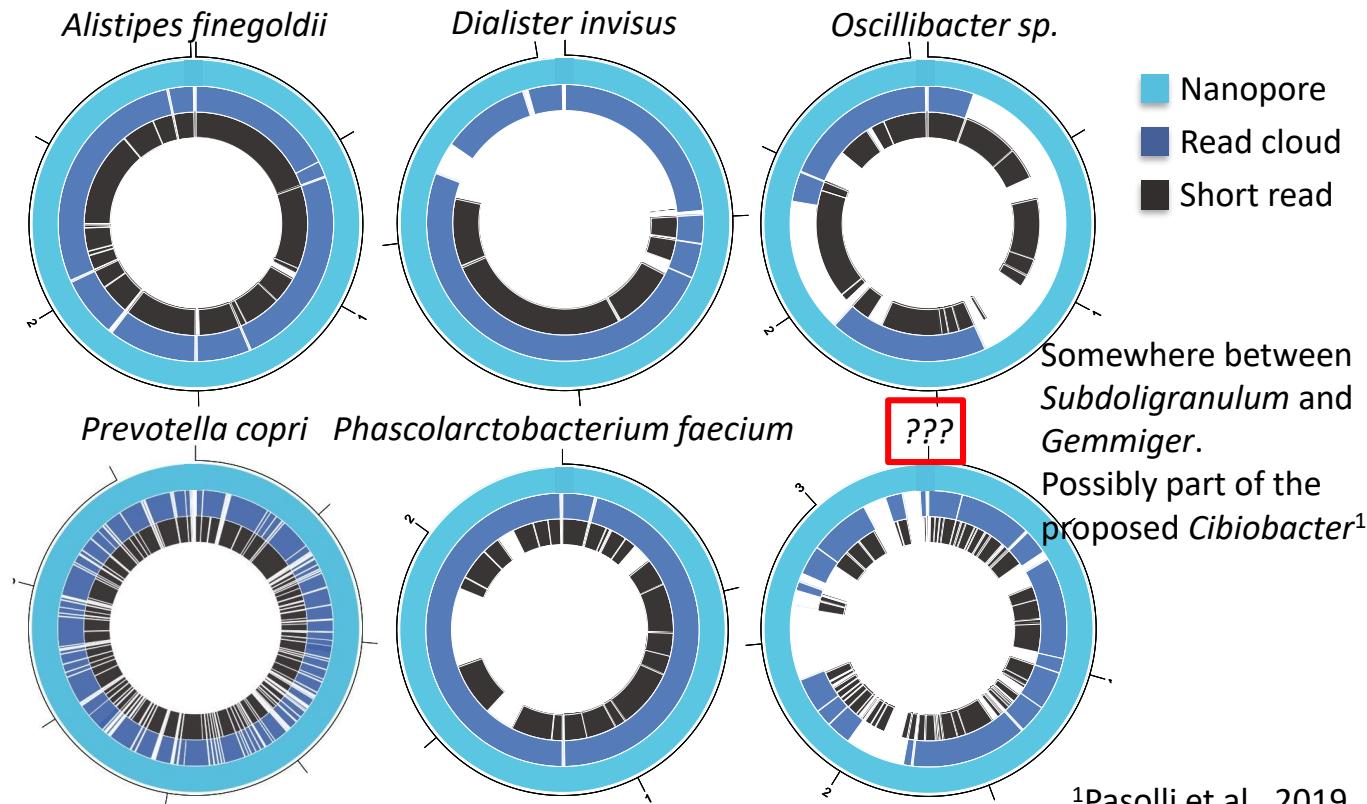


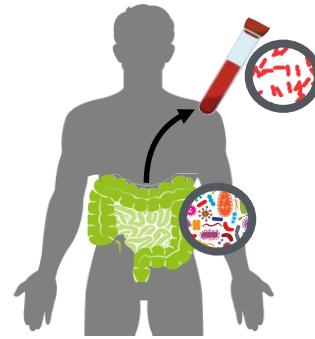
daysoftheyear.com





NANOPORE SEQUENCING CLOSES SEVERAL NOVEL BACTERIAL GENOMES



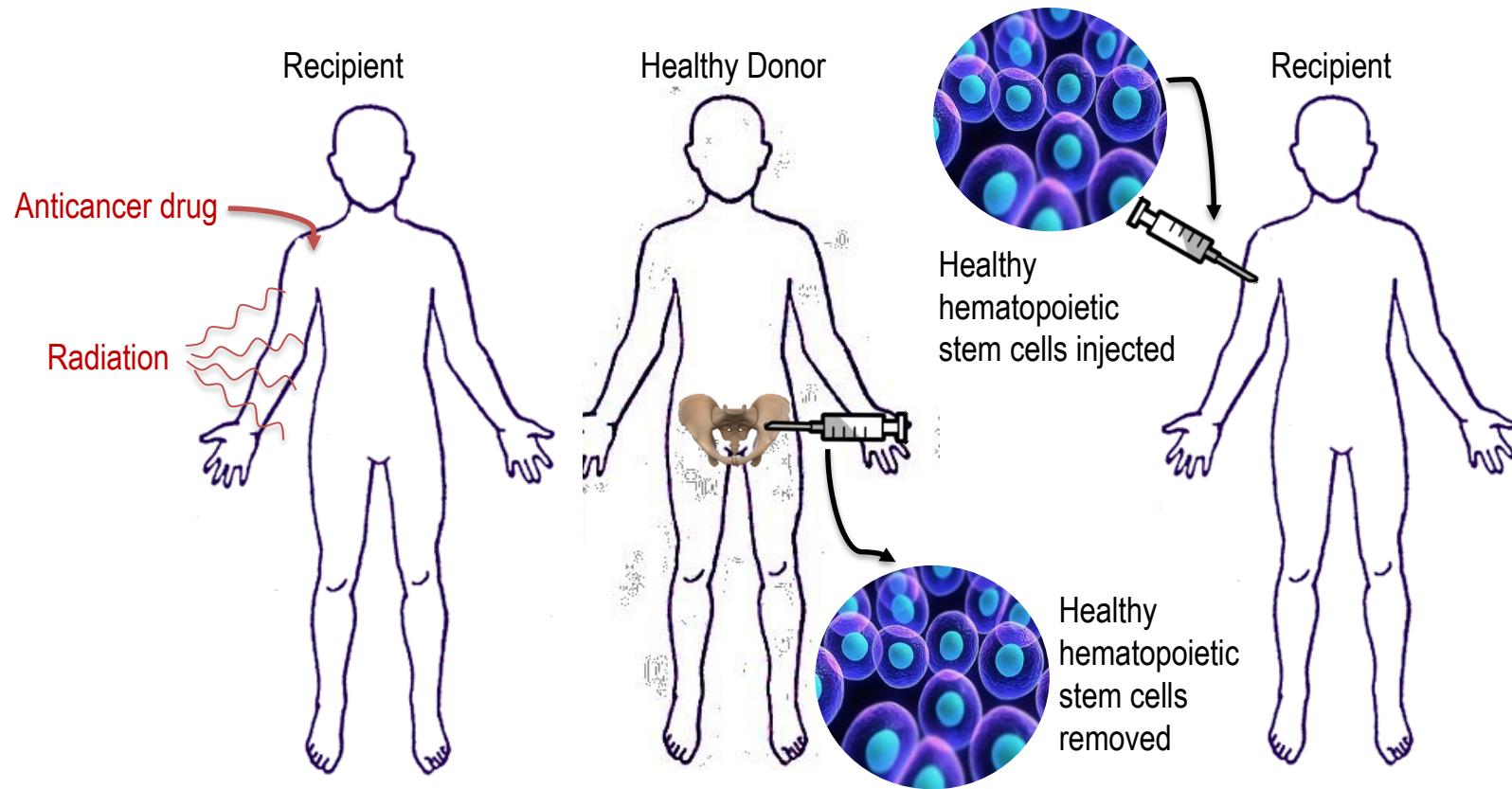


STRAINS MATTER

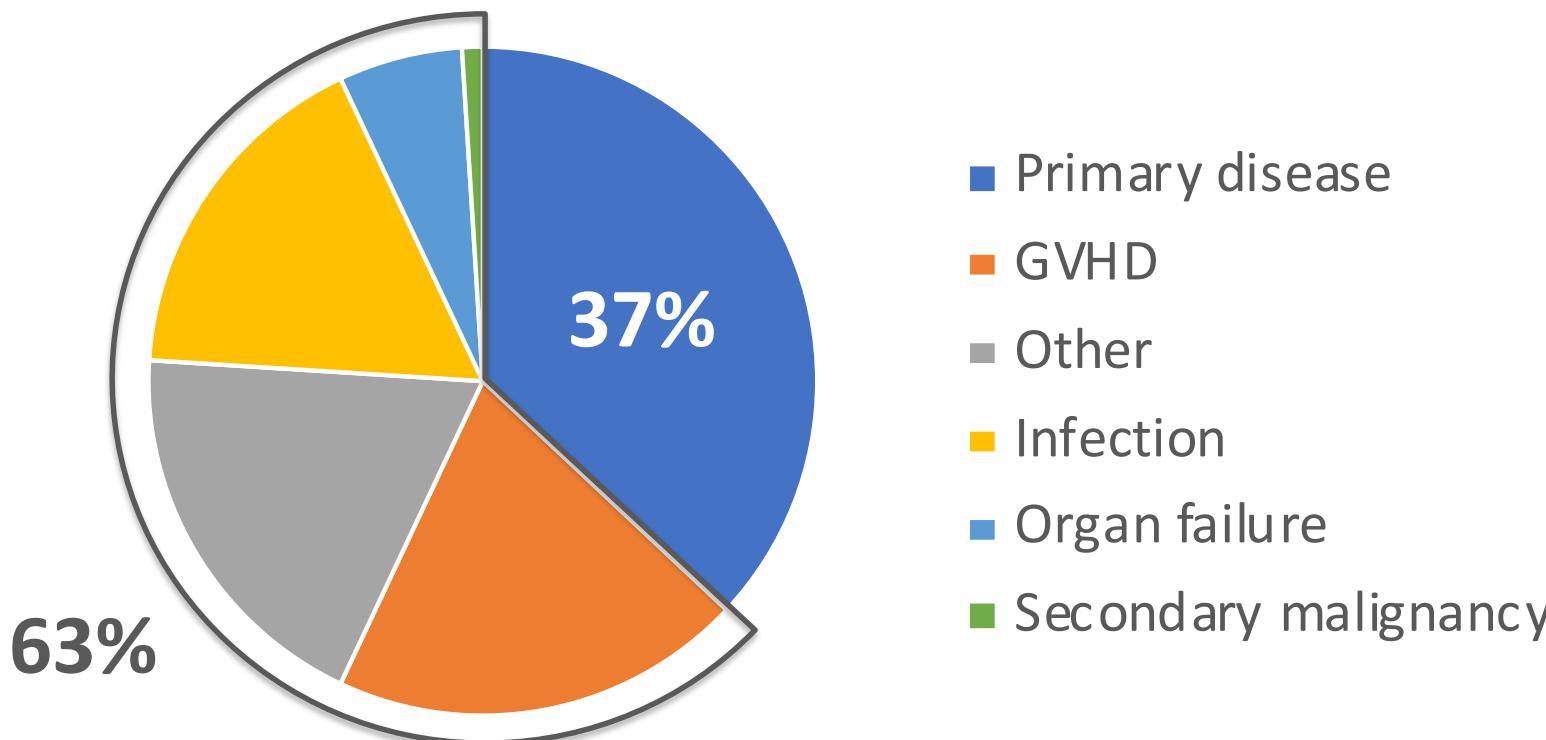
#3

PRECISION CLINICAL MICROBIOLOGY IN THE SEQUENCING ERA...

ALLOGENEIC HEMATOPOIETIC CELL TRANSPLANTATION

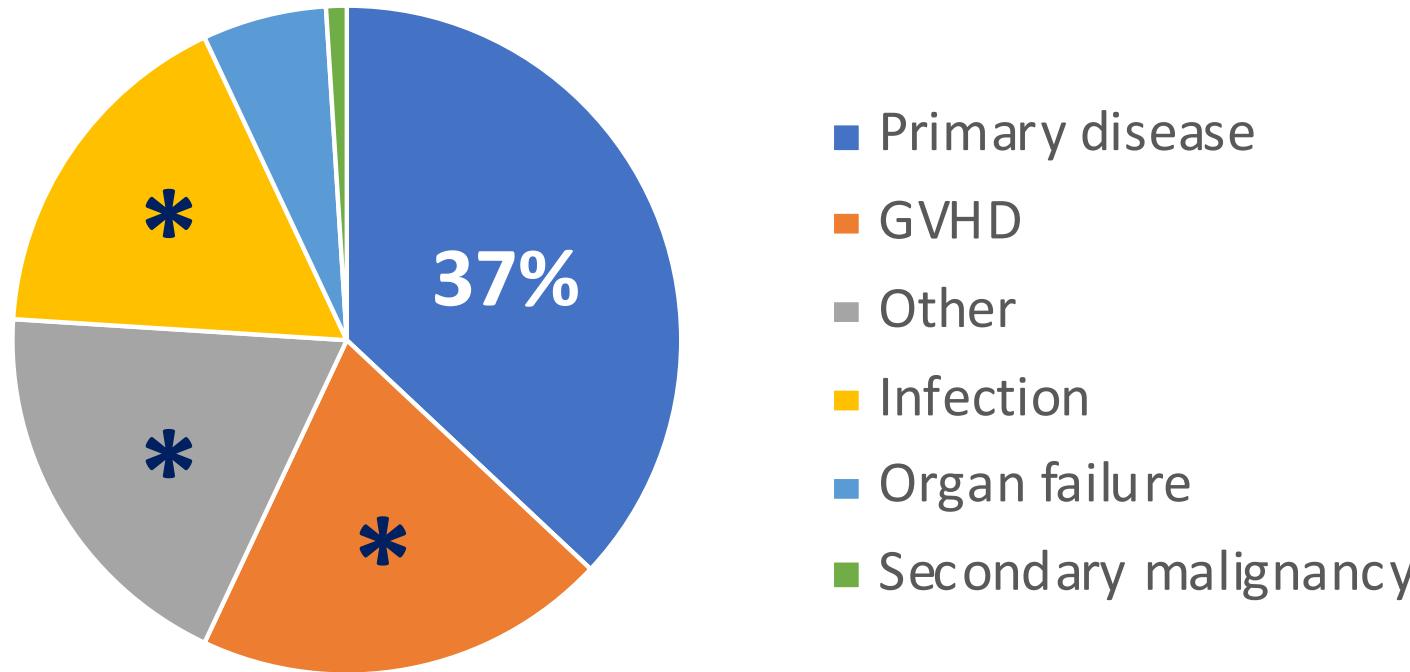


63% DIE OF TREATMENT-RELATED COMPLICATIONS

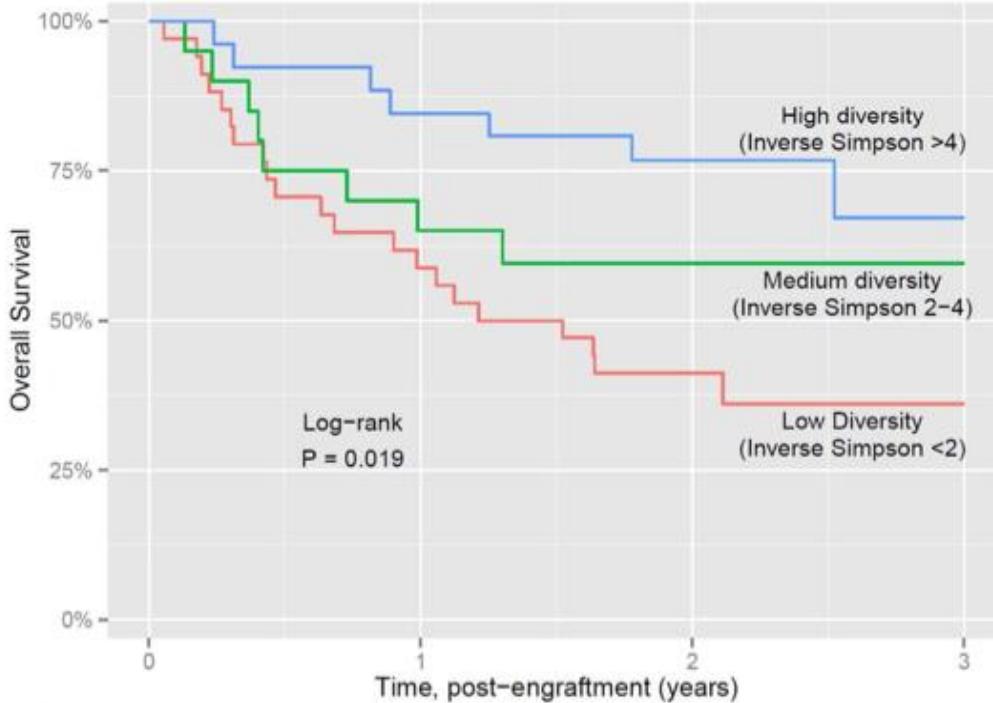


Causes of Death after Unrelated Donor Transplants done in 2011-2012

MANY NON-CANCER HCT OUTCOMES ARE ASSOCIATED WITH CHANGES IN THE MICROBIOME



Causes of Death after Unrelated Donor Transplants done in 2011-2012



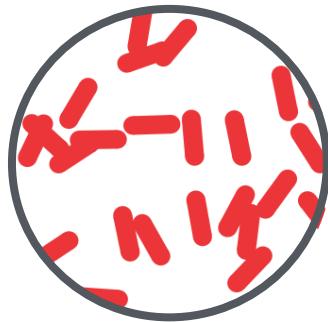
LOW DIVERSITY IS ASSOCIATED WITH POOR SURVIVAL

Taur et al,
Blood, 2014

POST-HCT LOSS OF DIVERSITY IS COMMON

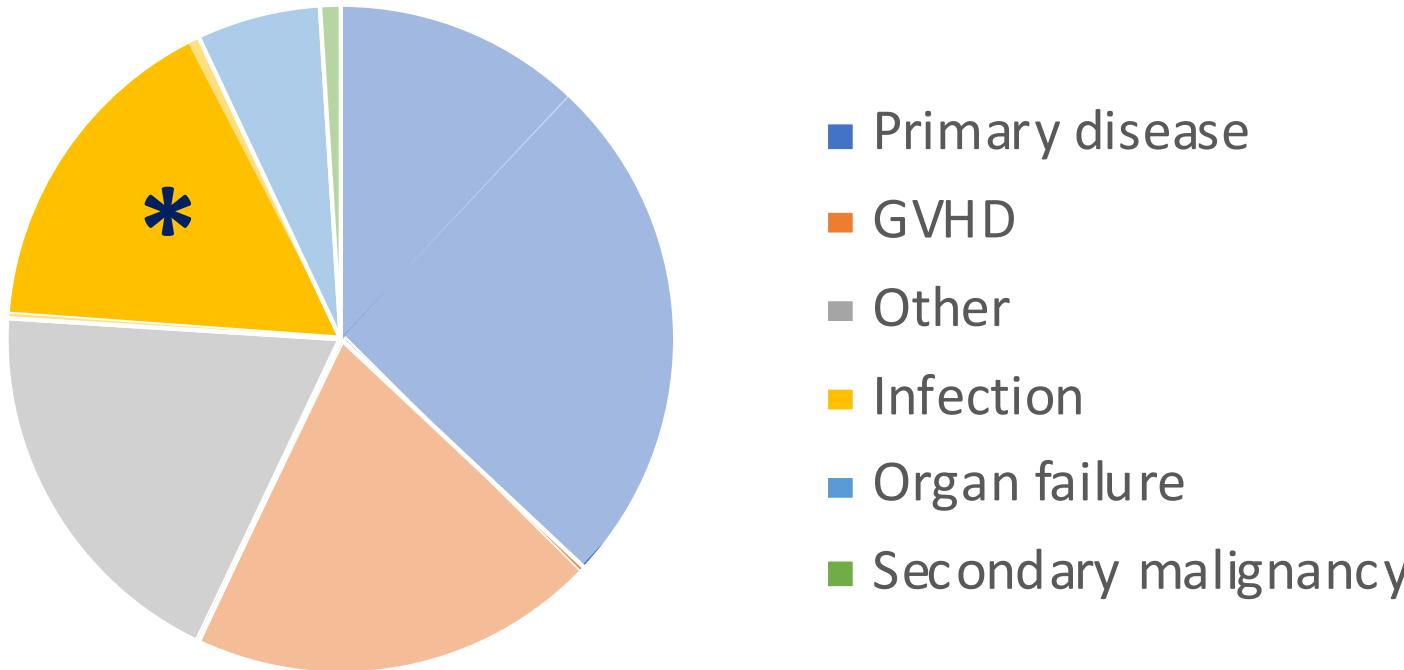


POST-HCT LOSS OF DIVERSITY IS COMMON



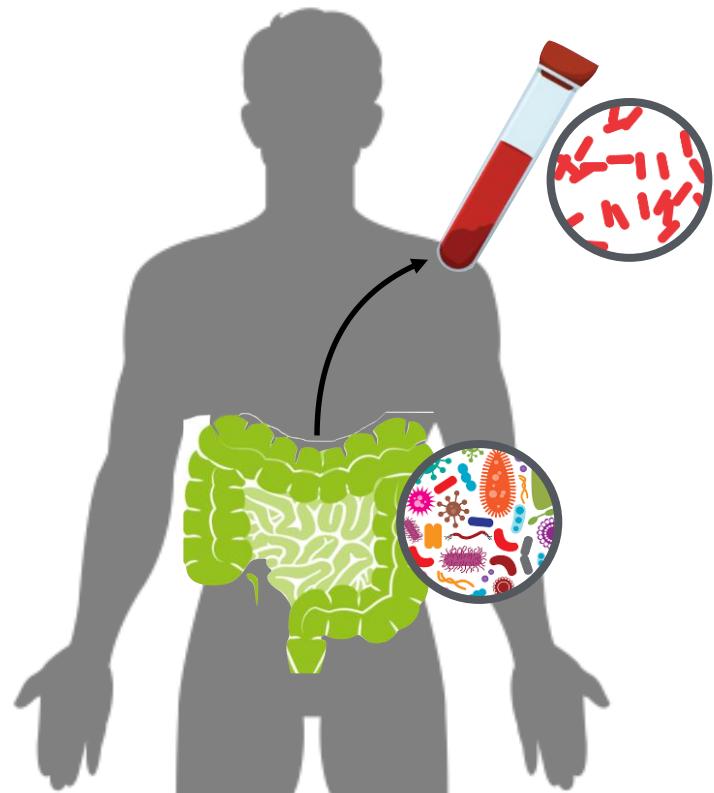
500 sequential HCT patients at Stanford
>1200 stool samples collected
>40% of subjects are “oligodominated”

THE INTESTINAL MICROBIOME MAY BE A RESERVOIR FOR PATHOGENS IN HCT



Causes of Death after Unrelated Donor Transplants done in 2011-2012

WHERE DO PATHOGENS COME FROM?



HOSPITAL ACQUIRED?

-or-

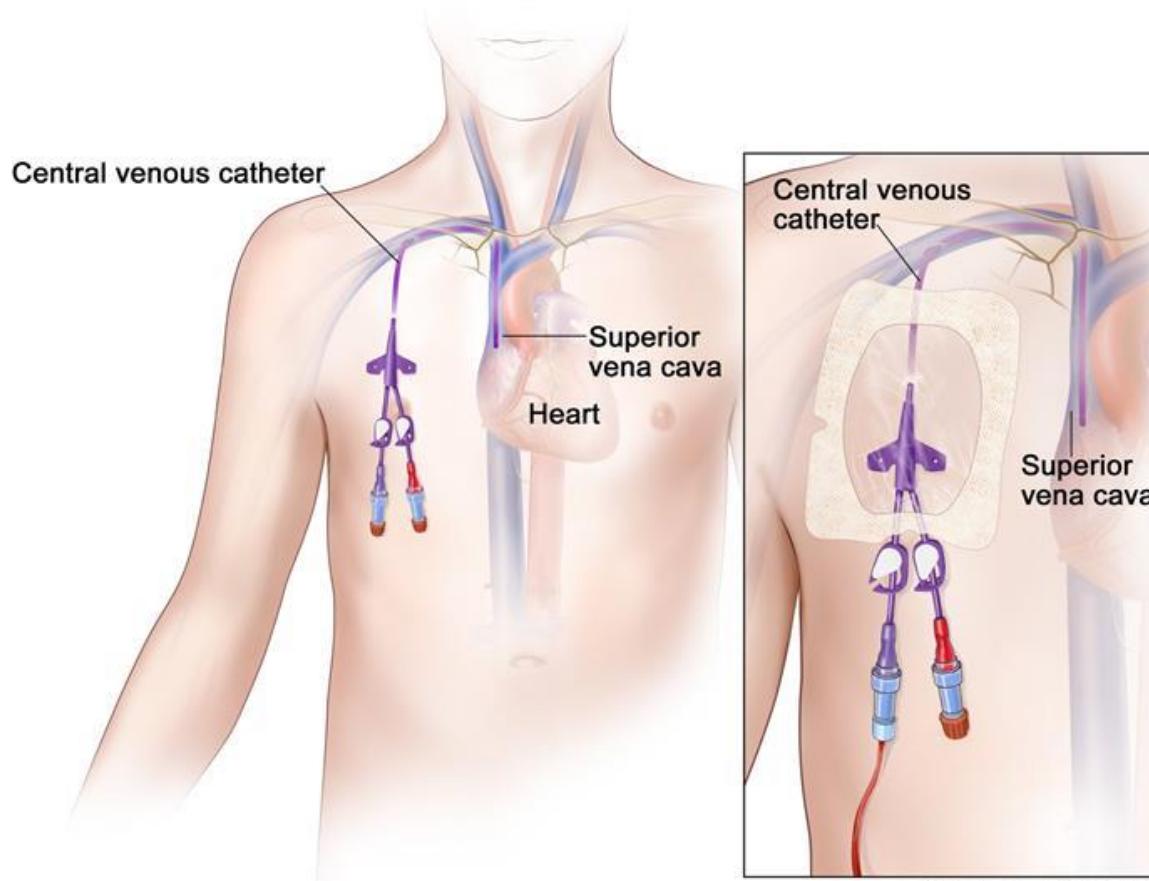
INTERNAL SOURCE?

Mystery Case

39F with B-cell acute lymphoblastic leukemia underwent HCT. 37 days after HCT, she developed fevers and chills. A blood culture showed growth of *Staphylococcus epidermidis*.

Where did this infection come from?

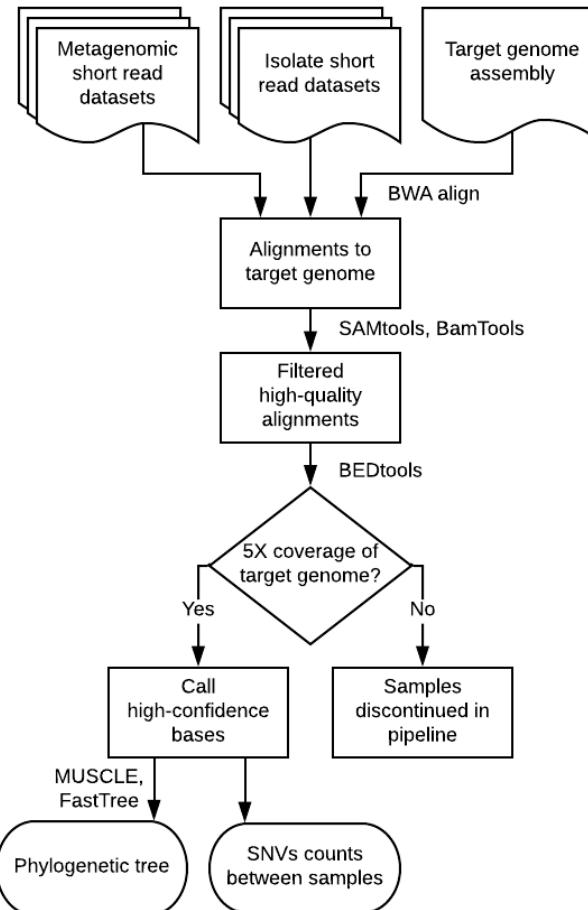
Central Venous Catheter



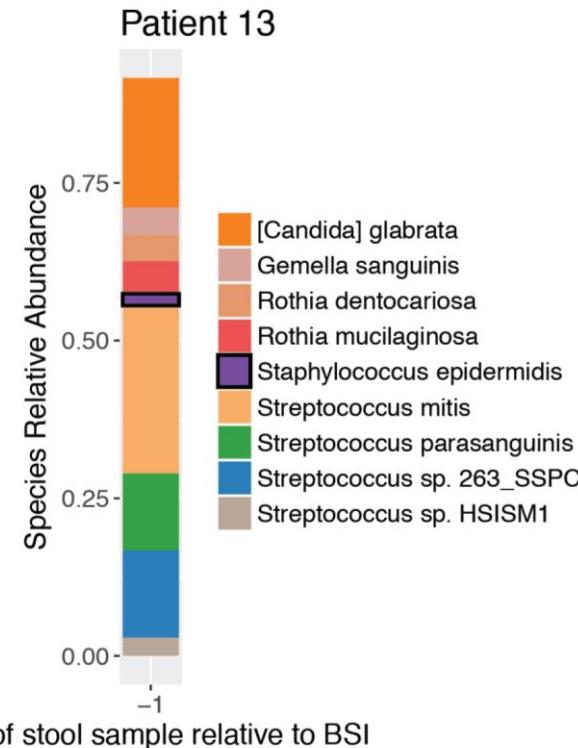
© 2015 Terese Winslow LLC
U.S. Govt. has certain rights

StrainSifter

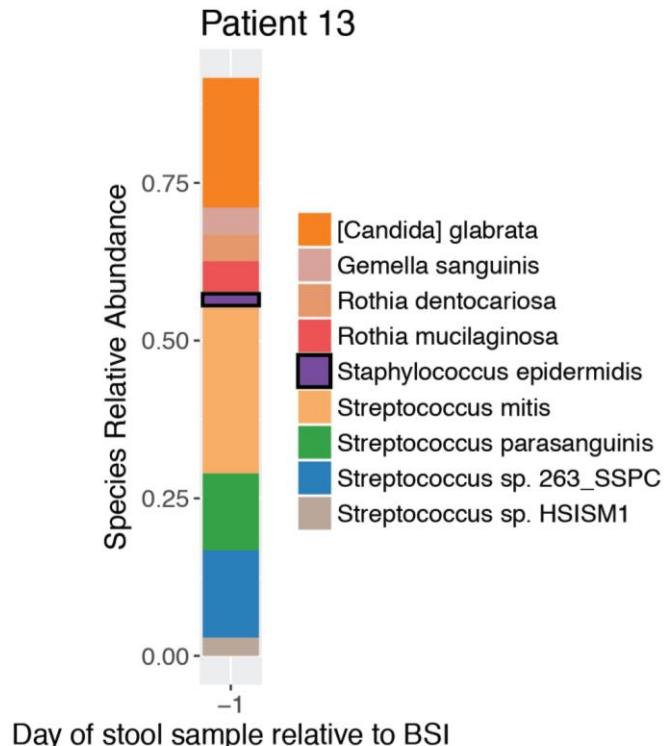
- A straightforward bioinformatic tool
- Identifies highly concordant microbial strains between samples
- Compares metagenomes to bacterial isolate sequencing



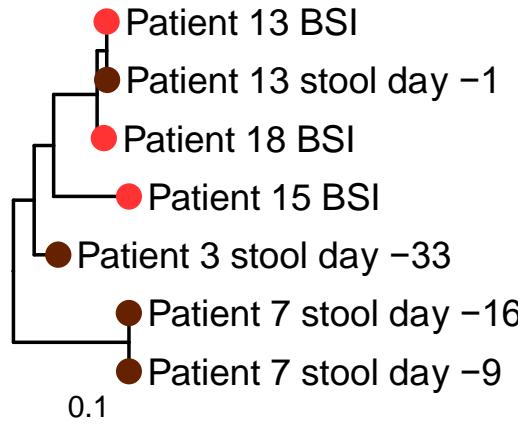
S. epidermidis infection likely originated from the gut microbiome – not the skin



S. epidermidis infection likely originated from the gut microbiome – not the skin

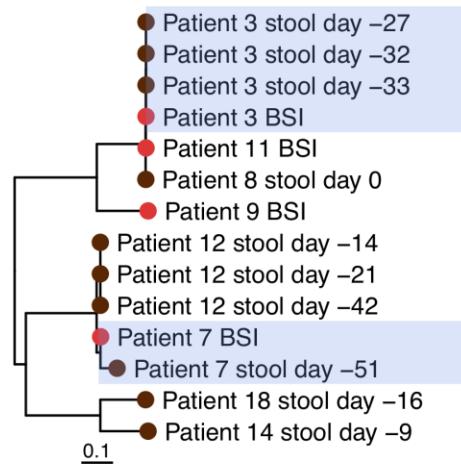


Staphylococcus epidermidis



DIVERSE BACTERIAL INFECTIONS ORIGINATE FROM THE GUT

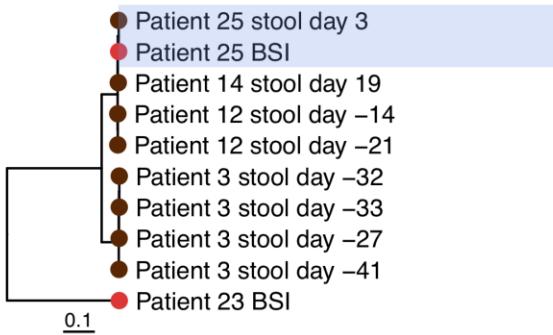
Escherichia coli



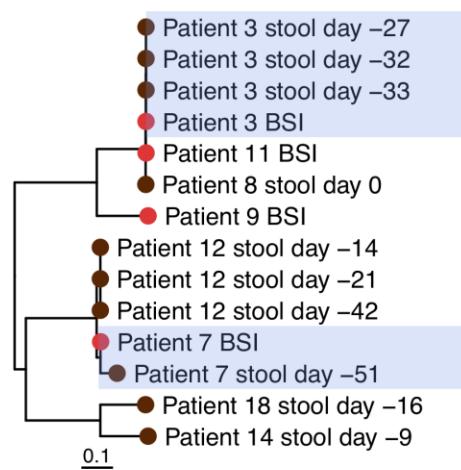
- Bloodstream isolate (BSI)
- Stool

DIVERSE BACTERIAL INFECTIONS ORIGINATE FROM THE GUT

Enterococcus faecium



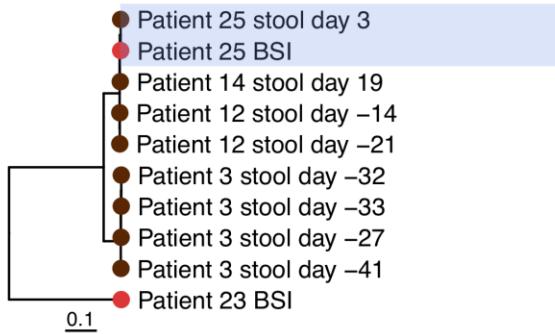
Escherichia coli



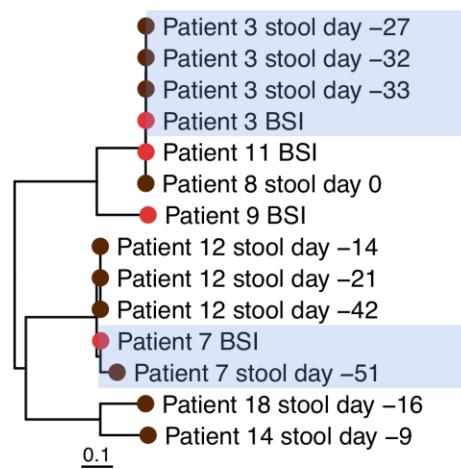
— Bloodstream isolate (BSI)
● Stool

DIVERSE BACTERIAL INFECTIONS ORIGINATE FROM THE GUT

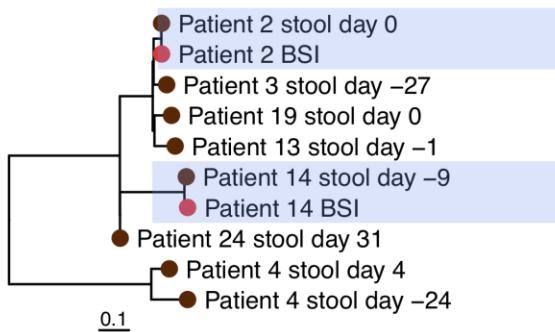
Enterococcus faecium



Escherichia coli



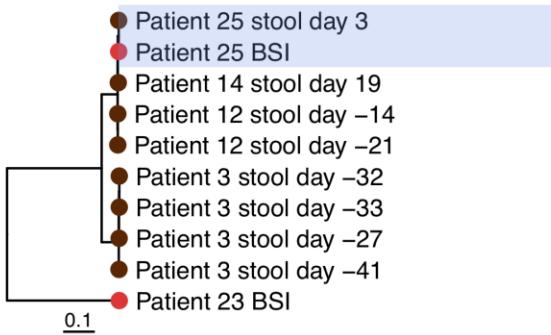
Klebsiella pneumoniae



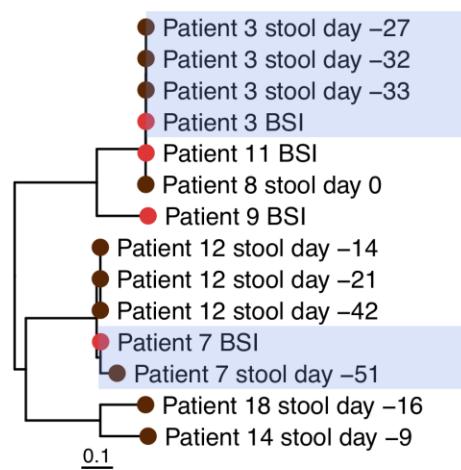
● Bloodstream isolate (BSI)
● Stool

DIVERSE BACTERIAL INFECTIONS ORIGINATE FROM THE GUT

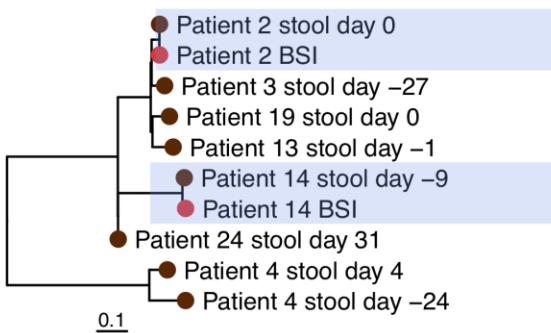
Enterococcus faecium



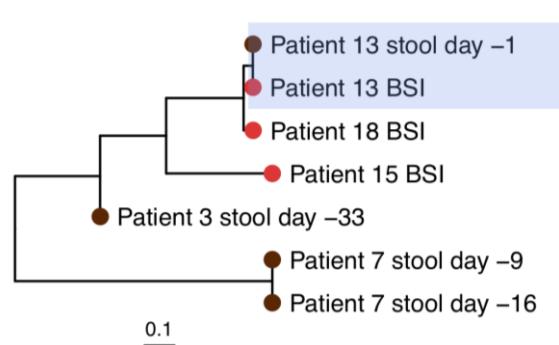
Escherichia coli



Klebsiella pneumoniae



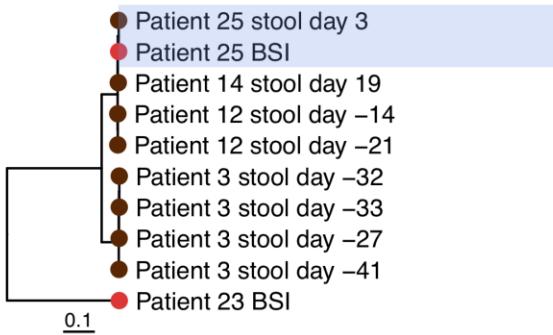
Staphylococcus epidermidis



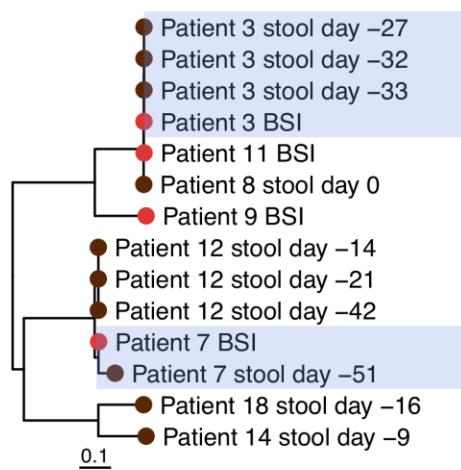
● Bloodstream isolate (BSI)
● Stool

DIVERSE BACTERIAL INFECTIONS ORIGINATE FROM THE GUT

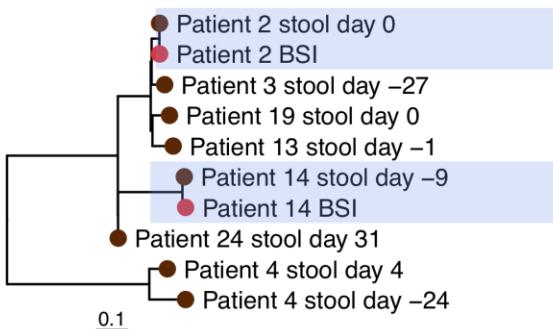
Enterococcus faecium



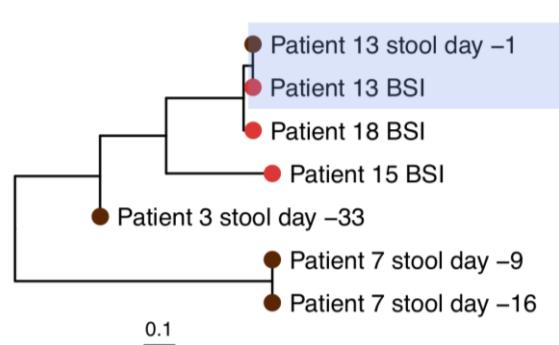
Escherichia coli



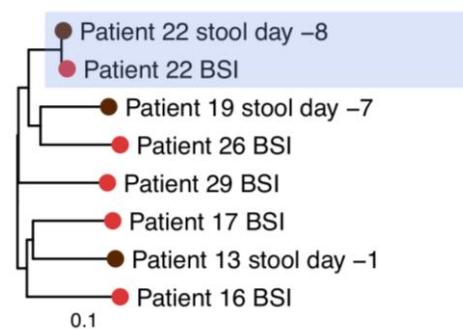
Klebsiella pneumoniae



Staphylococcus epidermidis



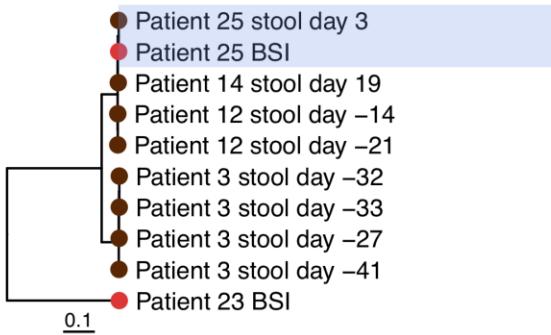
Streptococcus mitis



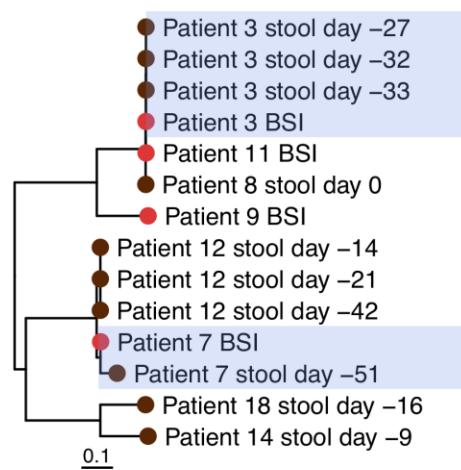
● Bloodstream isolate (BSI)
● Stool

DIVERSE BACTERIAL INFECTIONS ORIGINATE FROM THE GUT

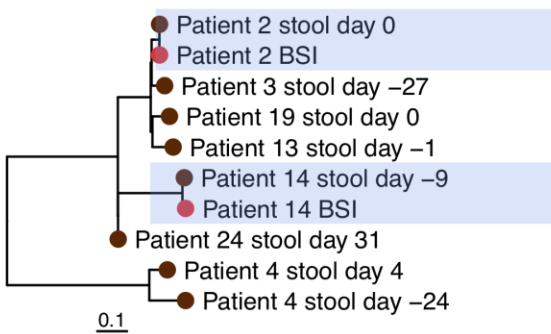
Enterococcus faecium



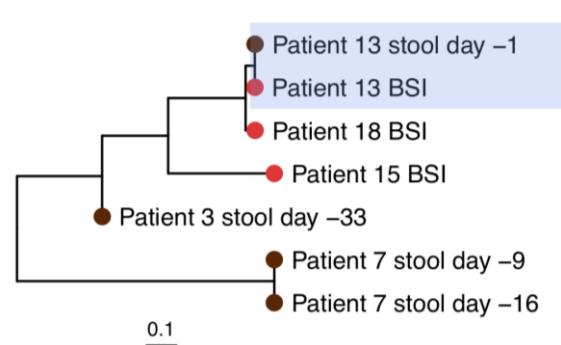
Escherichia coli



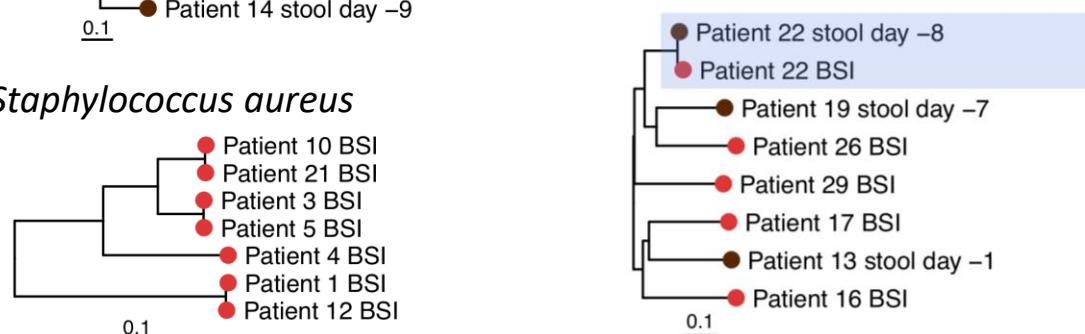
Klebsiella pneumoniae



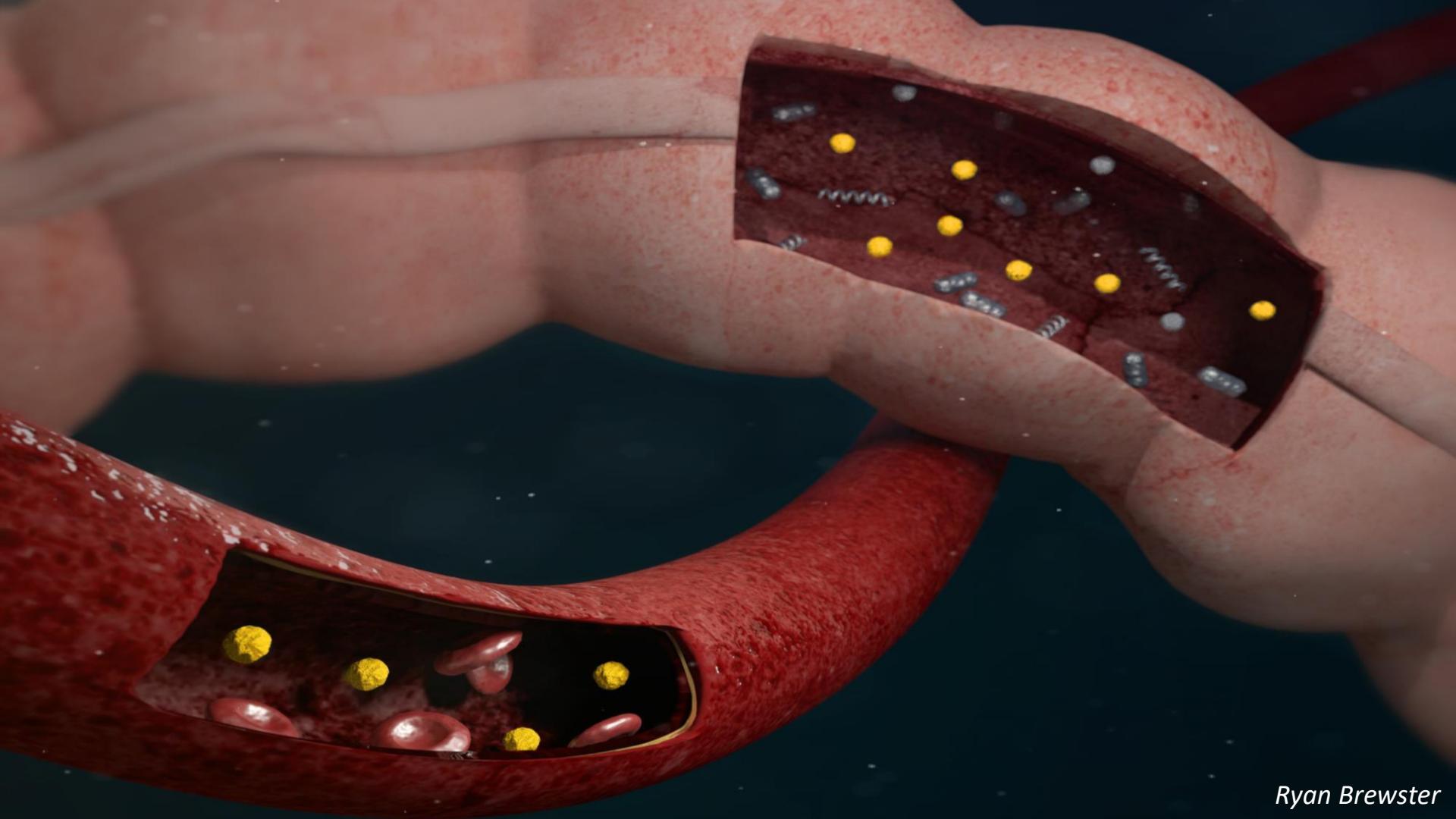
Staphylococcus epidermidis



Streptococcus mitis



● Bloodstream isolate (BSI)
● Stool

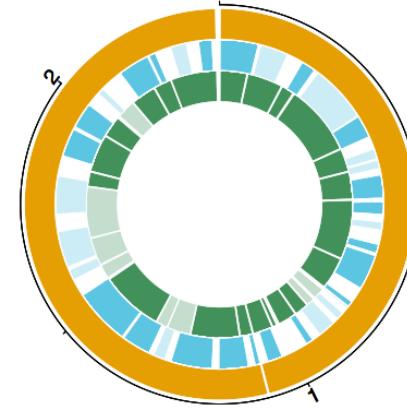
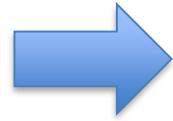
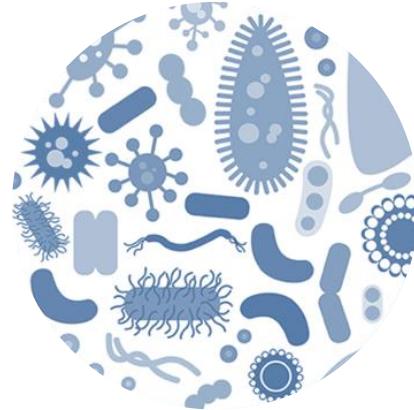


Ryan Brewster

#1

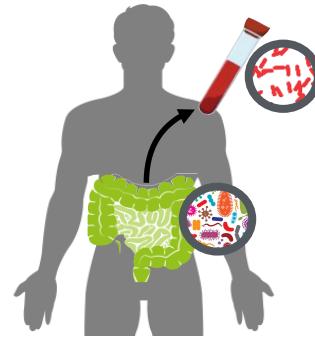
**WE CAN SAVE CANCER PATIENT LIVES
BY MANIPULATING THE MICROBIOME**





#2

GENOMES FROM METAGENOMES: ILLUMINATING THE DARK MATTER



STRAINS MATTER

#3

PRECISION CLINICAL MICROBIOLOGY IN THE SEQUENCING ERA...



G CCACTATATGGAGAGACCCCTGATCTATTCTAGAGCTTTCC

GA TAGAGATTTCCCCATATGCCCTATTGGGGCTCACACGCTCA

AGAGATTTCCCCATATGCCCTATTGGGGCTCACACGCTCACGTCA

IMPROVING GLOBAL REPRESENTATION IN MICROBIOME RESEARCH

Soweto Developmental
Pathways to Health
Research Unit

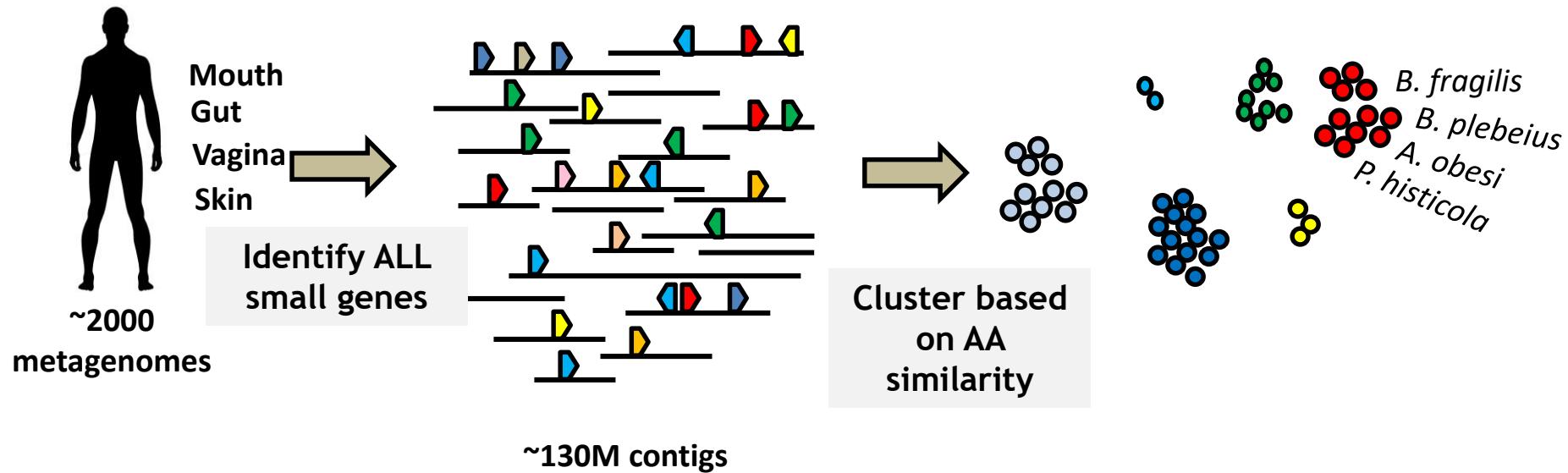


*With Scott Hazelhurst, Ovokeraye Voduaan,
Venesa Sahibdeen and colleagues*

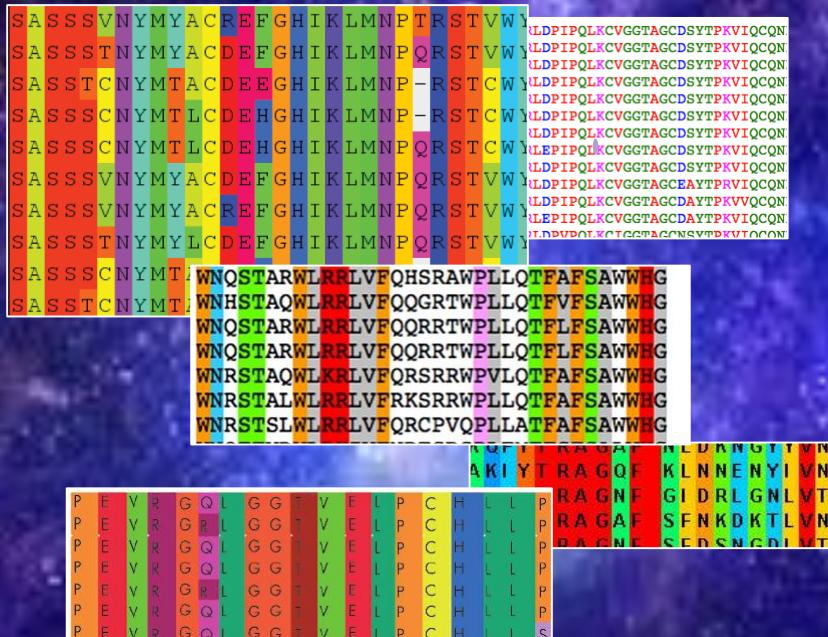
Agincourt Health and
Demographic Surveillance
Site



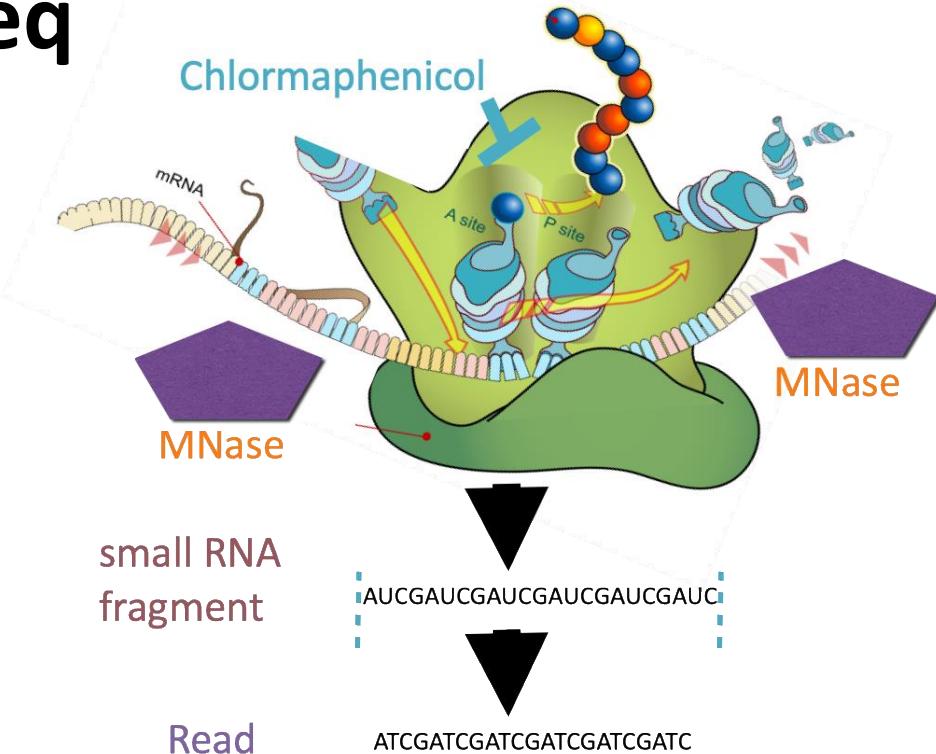
A COMPUTATIONAL APPROACH IDENTIFIES “REAL” SMALL PROTEINS



IDENTIFIED >4,000 NOVEL SMALL PROTEIN FAMILIES → MANY OF TOTALLY UNKNOWN FUNCTION!



MEASURING TRANSLATION – MetaRibo-Seq



The Lab

Ben Siranosian

Brayon Fremin

Chris Severyn

Eli Moss

Eunice Yang

Fiona Tamburini

Hila Sberro

Jessica Ribado

Karen Andrade

Matt Durrant

Michelle Li

Paulina Chamely

Ryan Brewster

Soumaya Zlitni

Summer Vance

Tessa Andermann

Collaborators

Serafim Batzoglou

Andrew Rezvani

Sally Arai

Courtney Greene

Hanlee Ji

Stephen Montgomery

Mike Snyder

Elizabeth Hohmann

Anshul Kundaje

Rob Negrin

Federico Simonetta

David Miklos

Rebecca Culver

Harmony Folse

Edgar Asiimwe

Matt Buckley

Scott Hazelhurst

Venessa Sahibdeen

Ovokeraye Oduaran

Shane Norris

Steve Tollman

Michele Ramsay



Recent "grads"

Gavin Sherlock (visiting prof)

Katia Tkachenko

Joyce Kang

Alex Bishara

Our Patients