

UC San Diego

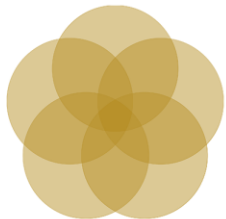
Moore's Cancer Center

ENHANCING MICROBIAL INSIGHTS IN CANCEROUS TISSUE: UNVEILING THE HIDDEN MICROBIOME WITH AN ADVANCED HOST DEPLETION PIPELINE

Caitlin Guccione

University of California, San Diego

July 26th, 2023



**Quantitative
Cancer
Control**

Laboratory · UCSD




What is a microbiome?

microorganisms



tiny organisms
living in all kinds of
environments

microbiome

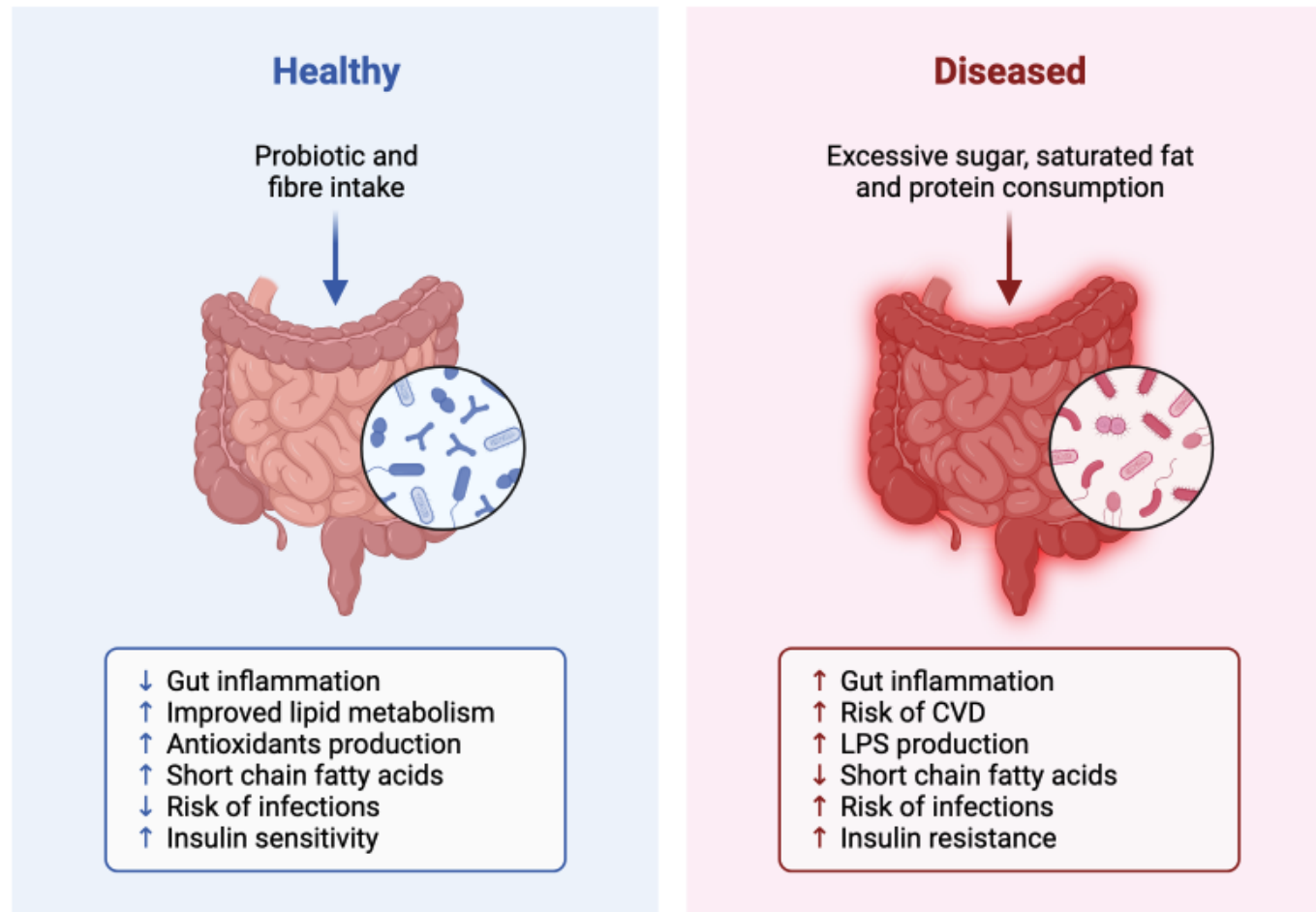


a community of microorganisms
& their role within a specific
environment, considering
environmental conditions &
interactions with each other

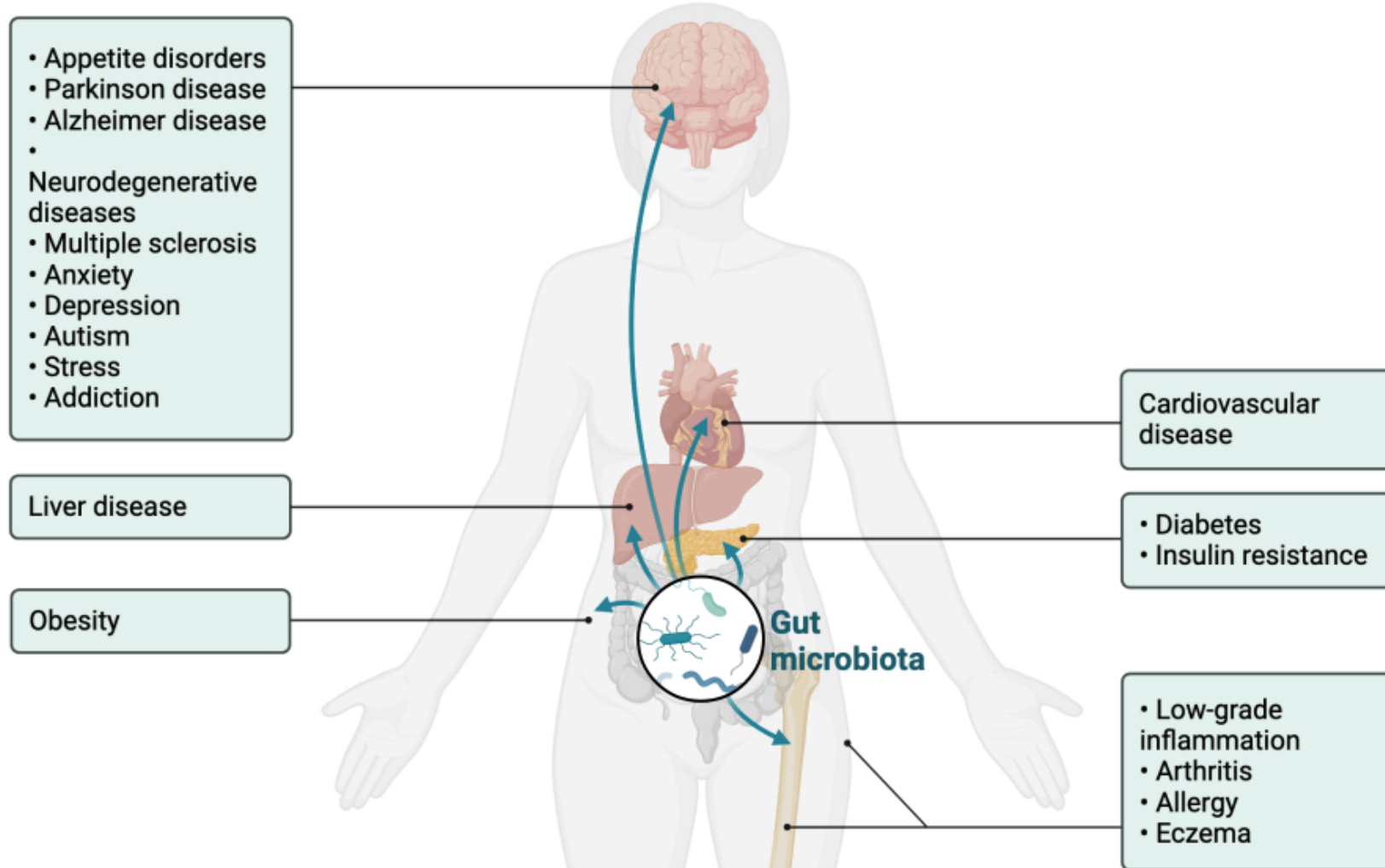
Bacteria ◆ Archaea ◆ Fungi ◆ Protists ◆ Viruses

Human Gut Microbiome

Role of Gut Microbiota in Nutrition and Health



Human Gut Microbiome



Human Oral Microbiome

Components of the oral microbiome:



Viruses



Bacteria



Archaea



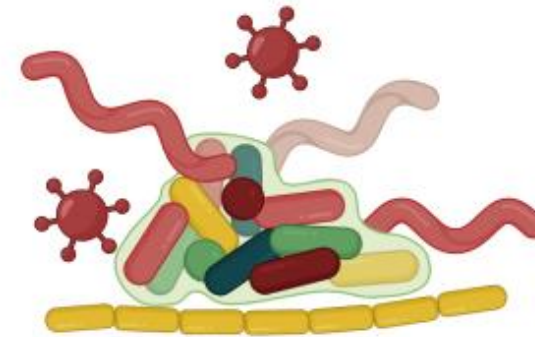
Fungi



Protozoa



Healthy microbiome (Eubiosis)

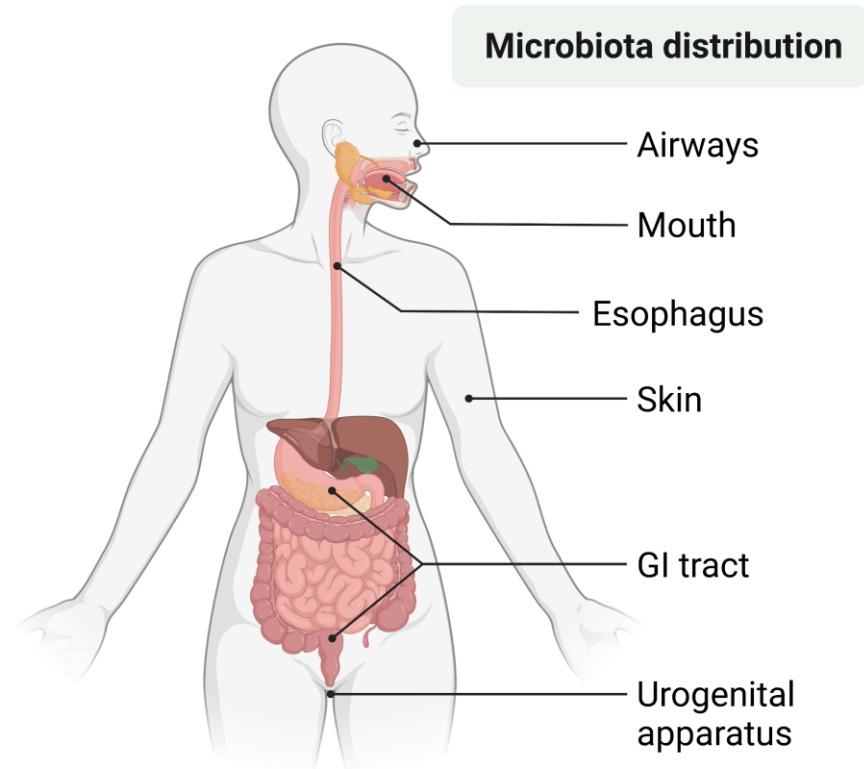
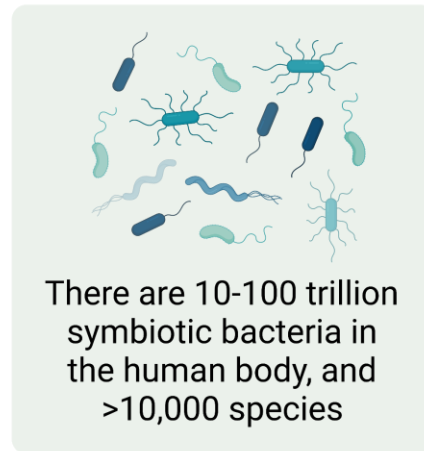


Unbalanced microbiome (Dysbiosis)

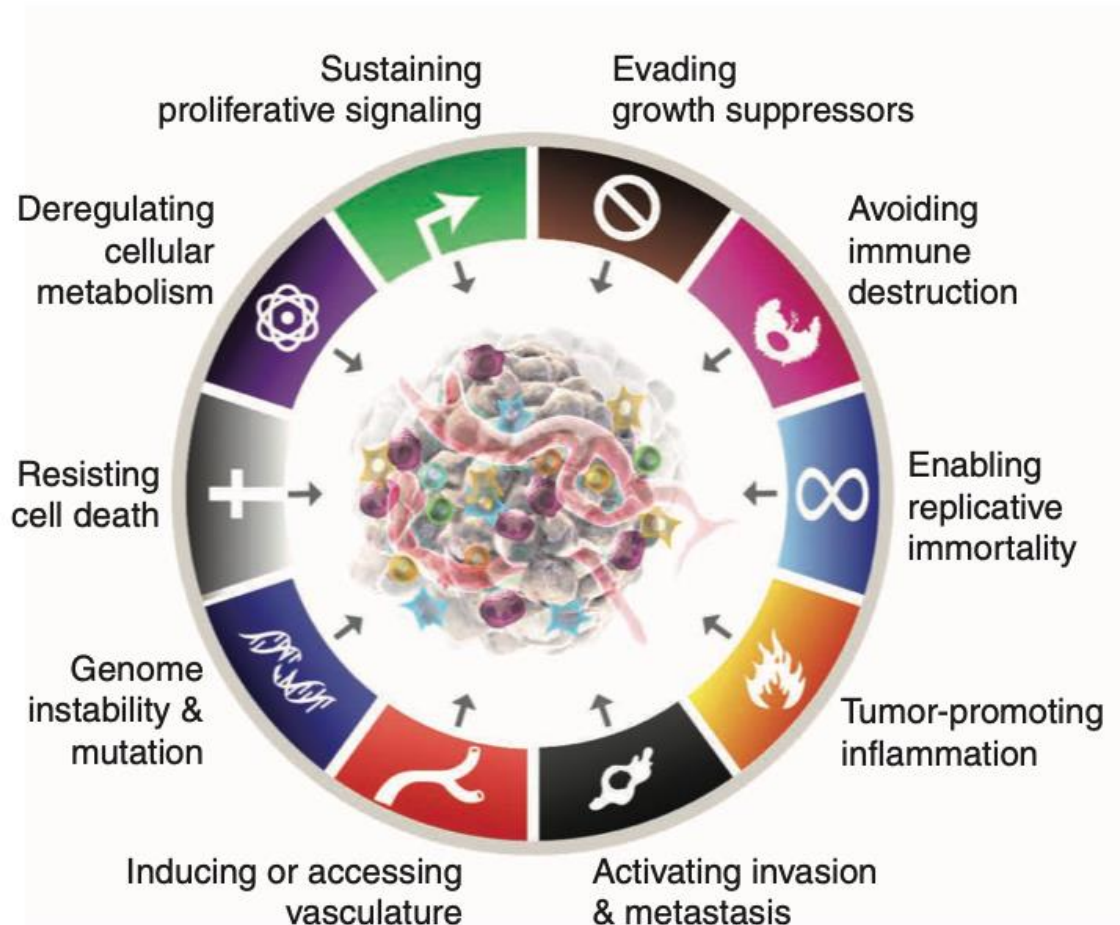
- Periodontitis/Caries & sequelae
- Endocarditis
- Atherosclerosis
- Alzheimer's disease
- Diabetes
- Head and neck cancer

Human Microbiome

More bacteria reside in human body than the actual human cells
The ratio of microbes to human cells is 1.3:1

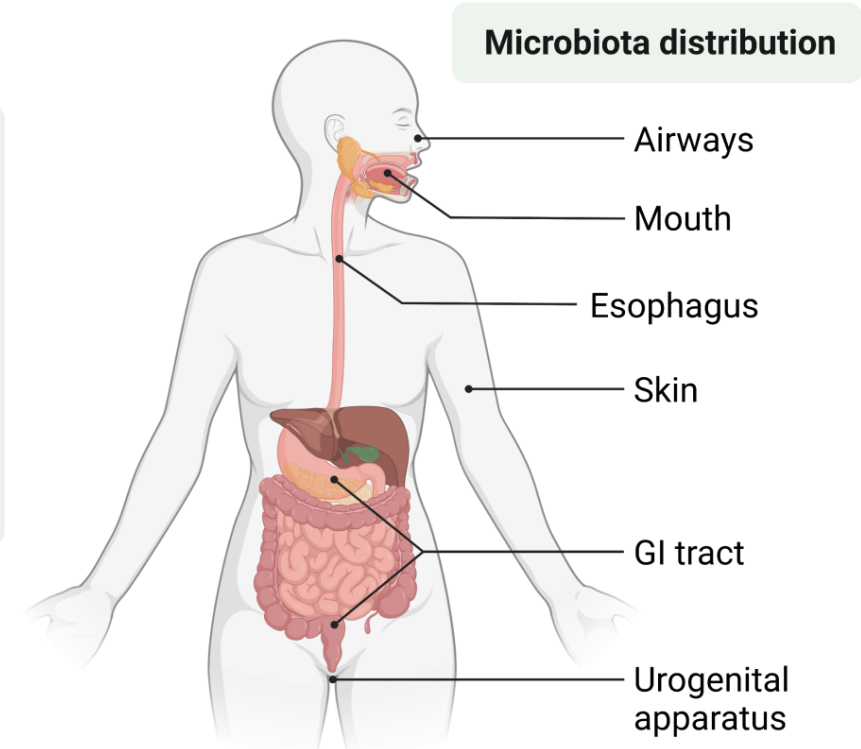
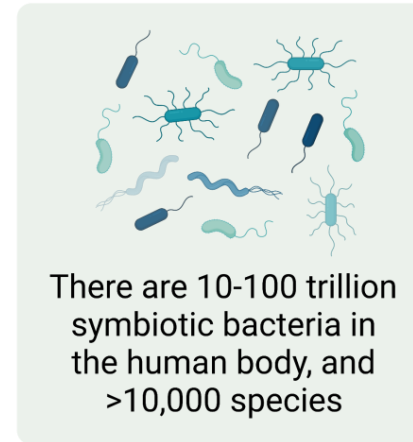


Cancer

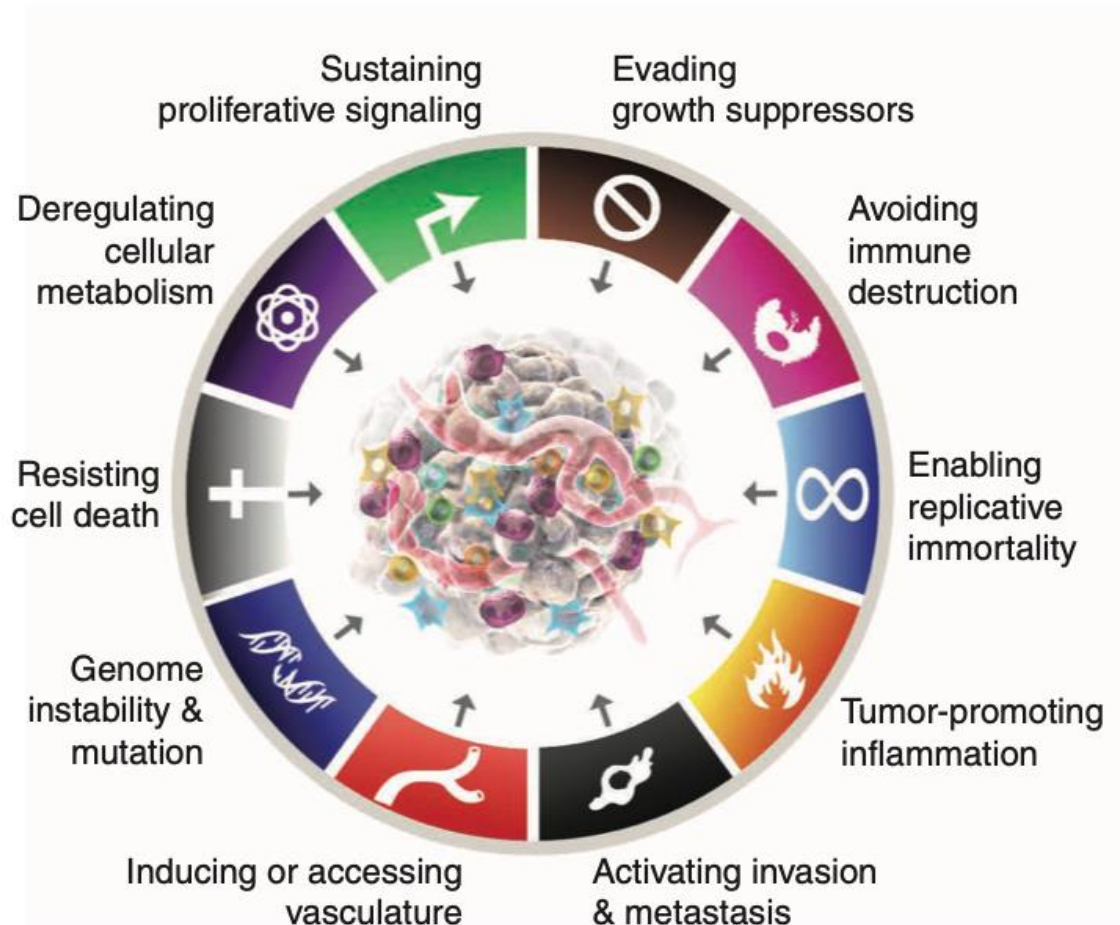


Human Microbiome

More bacteria reside in human body than the actual human cells
The ratio of microbes to human cells is 1.3:1



Cancer

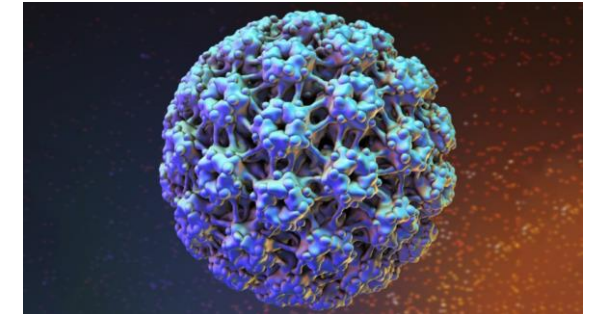


>15% of cancer cases worldwide are due to microbes

Human Microbiome



Helicobacter pylori (H. pylori)



Human papillomavirus (HPV)

Are there even more?

Are there microbes in The Cancer Genome Atlas (TCGA)?

nature

Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [articles](#) > [article](#)

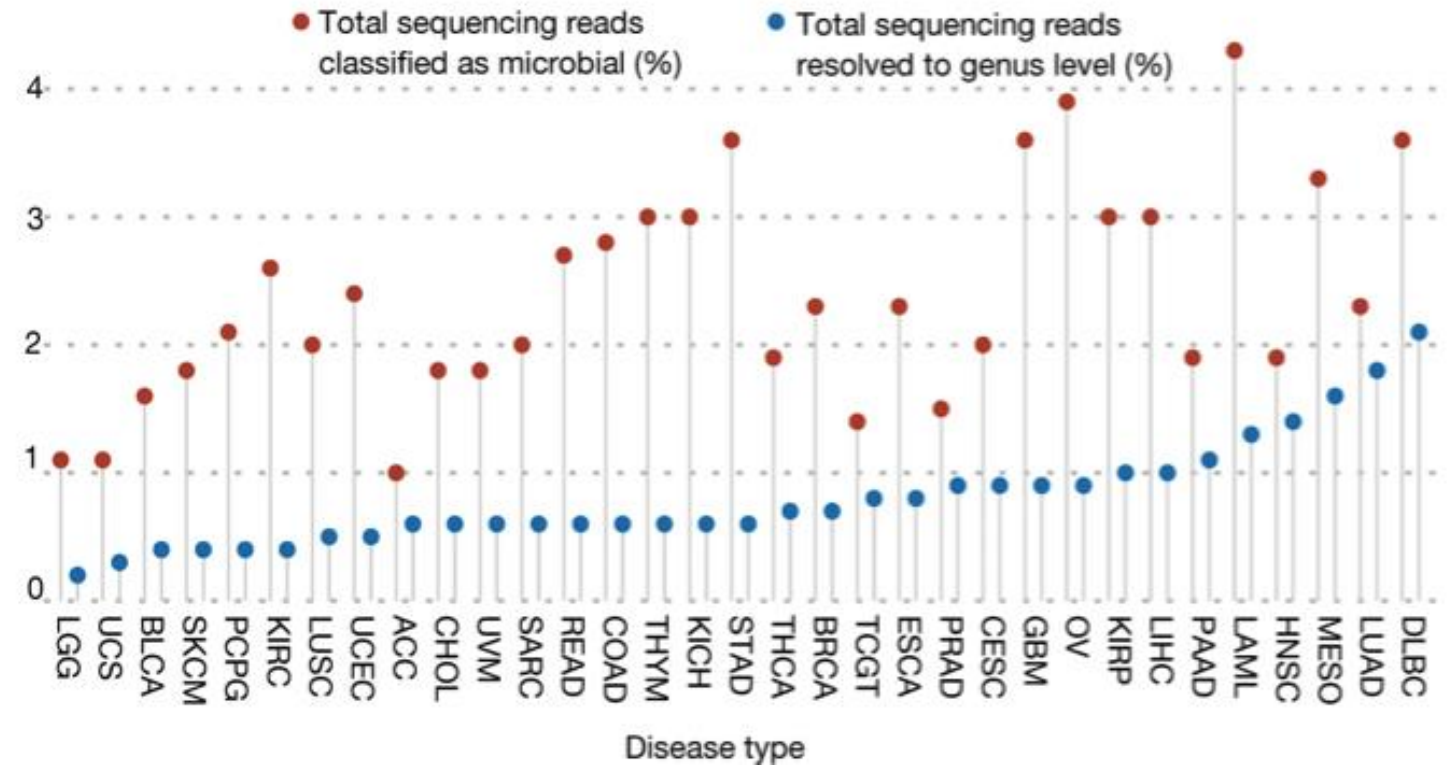
Article | [Published: 11 March 2020](#)

Microbiome analyses of blood and tissues suggest cancer diagnostic approach

[Gregory D. Poore](#), [Evguenia Kopylova](#), [Qiyun Zhu](#), [Carolina Carpenter](#), [Serena Fraraccio](#), [Stephen Wandro](#), [Tomasz Kosciolk](#), [Stefan Janssen](#), [Jessica Metcalf](#), [Se Jin Song](#), [Jad Kanbar](#), [Sandrine Miller-Montgomery](#), [Robert Heaton](#), [Rana McKay](#), [Sandip Pravin Patel](#), [Austin D. Swafford](#) & [Rob Knight](#) ✉

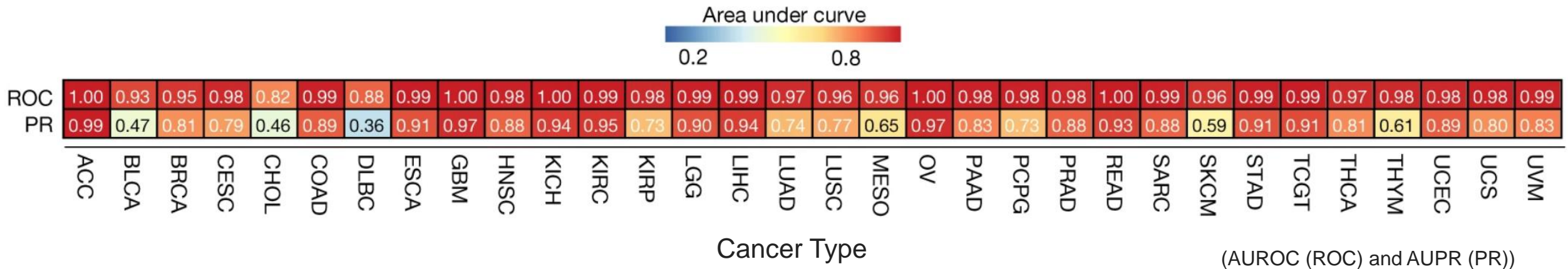
[Nature](#) **579**, 567–574 (2020) | [Cite this article](#)

68k Accesses | **458** Citations | **893** Altmetric | [Metrics](#)



Differences in microbiomes of tumor tissue across cancer types

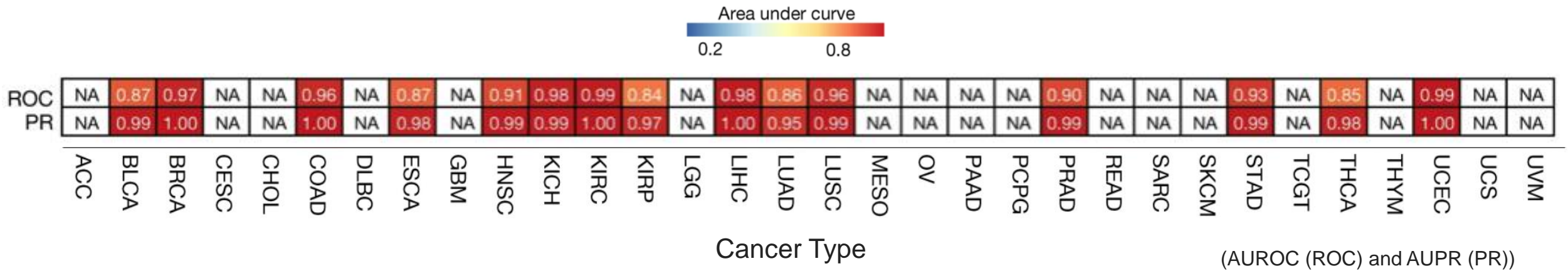
Heatmap of classifier performance comparing one cancer type vs all others



Takeaway: Each cancer type has a unique microbiome

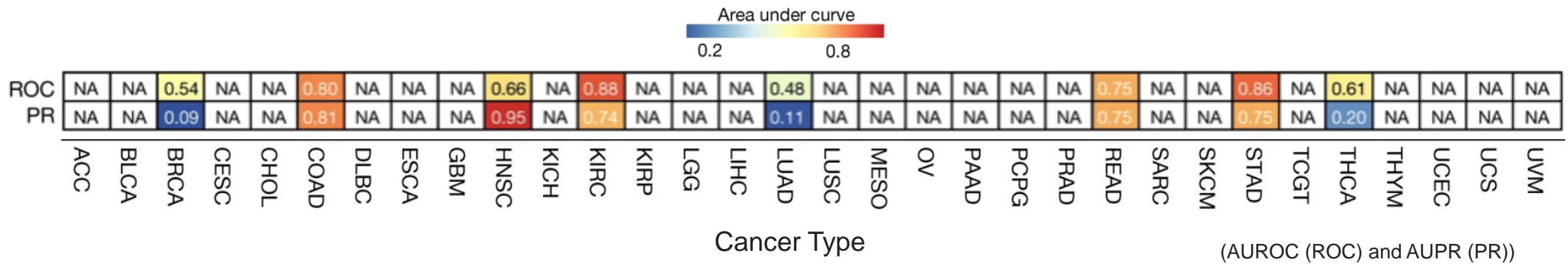
Differences in microbiomes of tumor tissue and normal samples

Heatmap of classifier performance comparing cancer vs normal



Differences in microbiomes between **stage I** and **stage IV** cancers

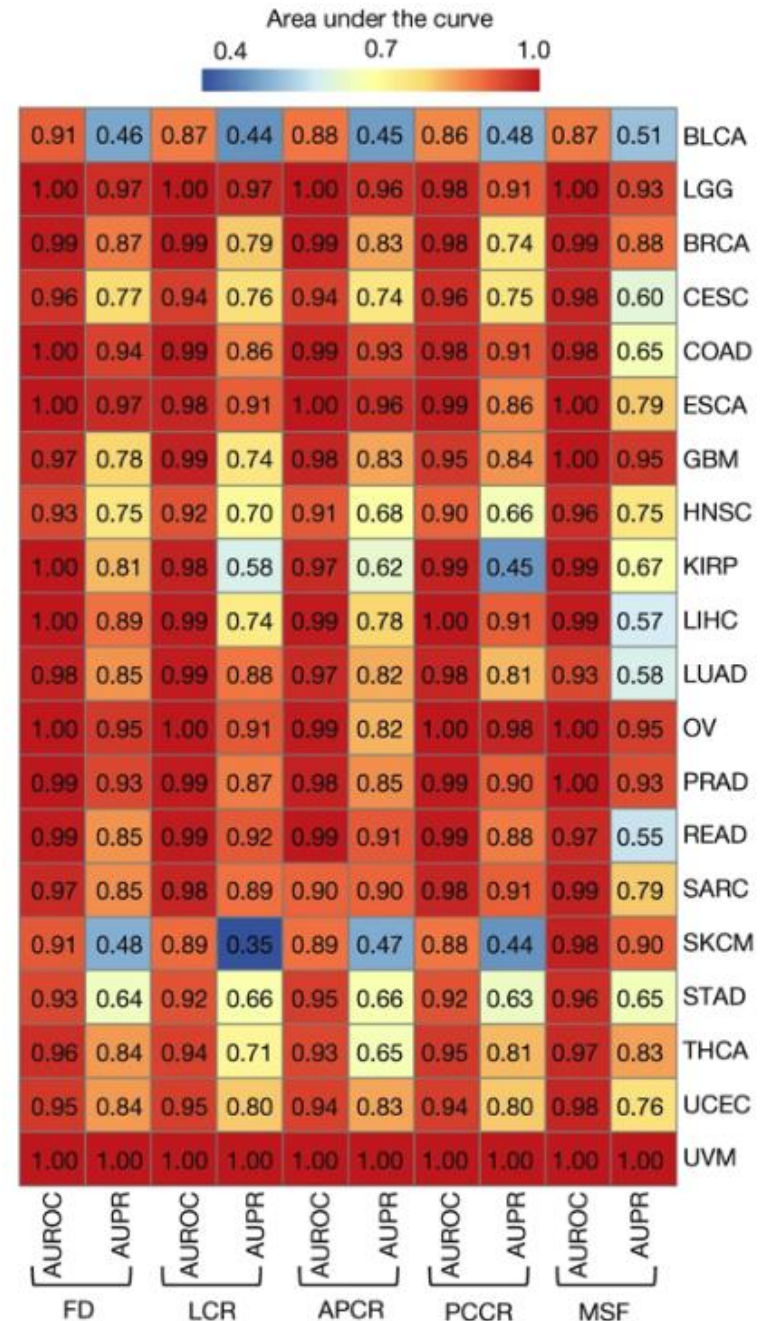
Heatmap of classifier performance comparing Stage I and stage IV cancers



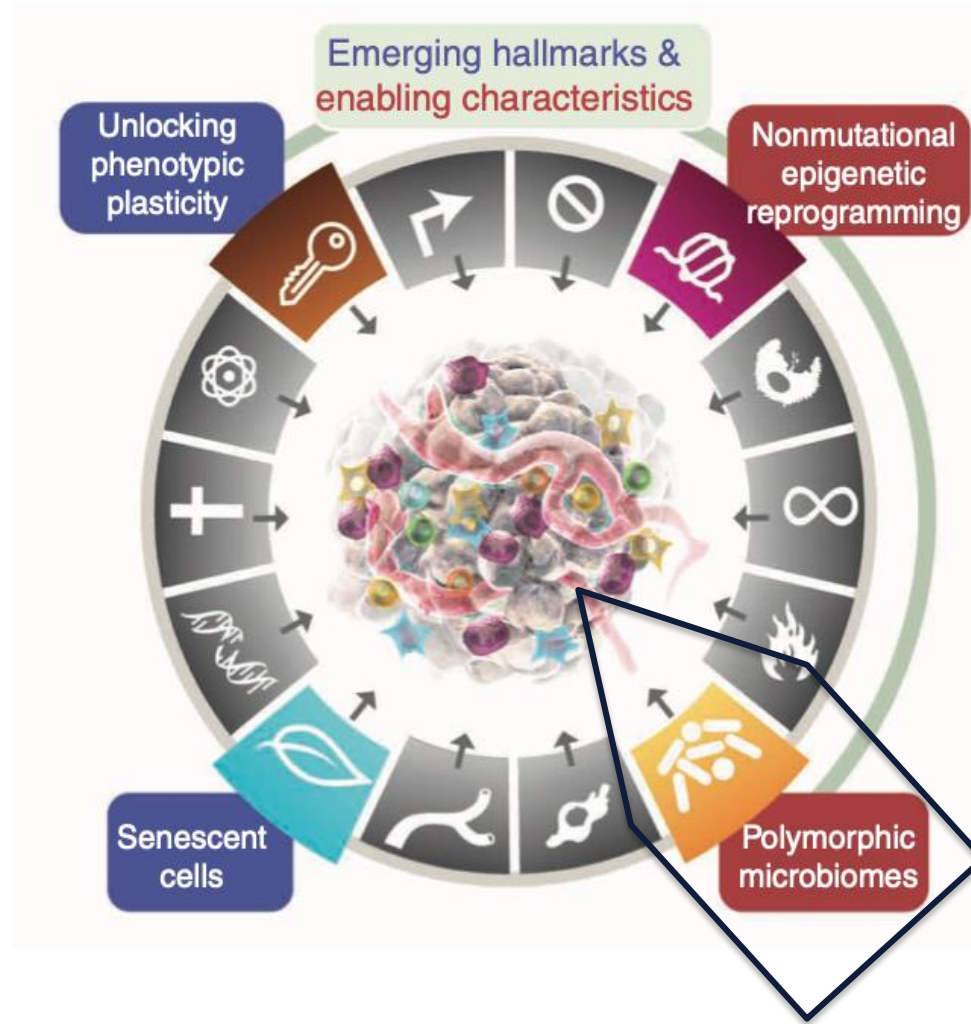
Differences in microbiomes found in the **blood** of one cancer type vs. all others

Takeaway: We can identify cancer type in patients using microbes found in the blood

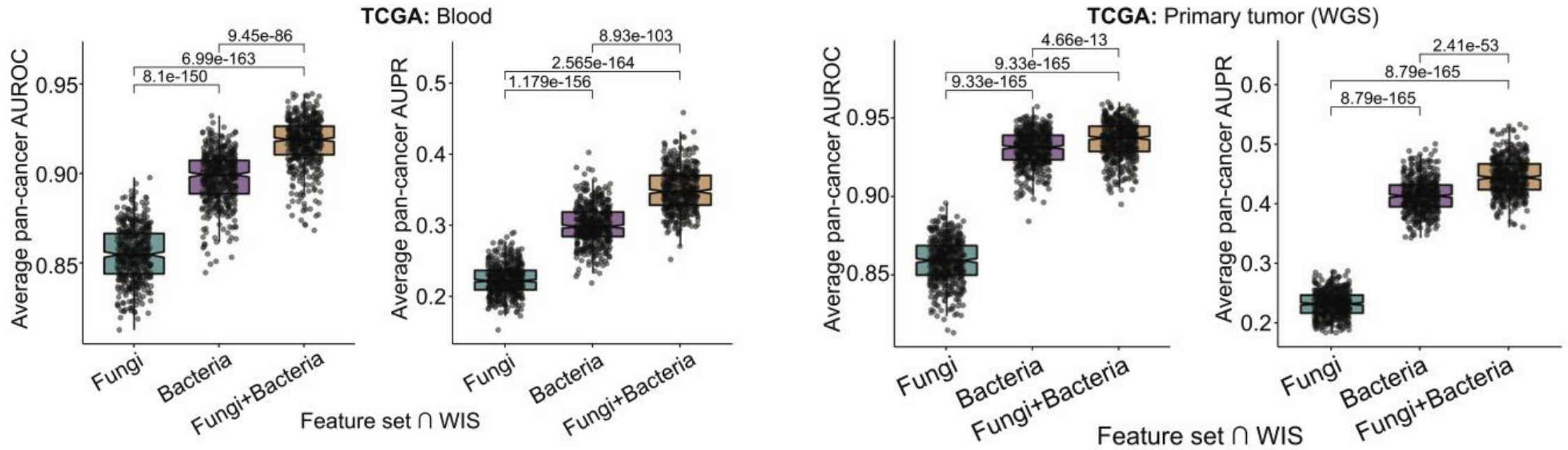
(AUROC (ROC) and AUPR (PR))



Cancer Microbiome

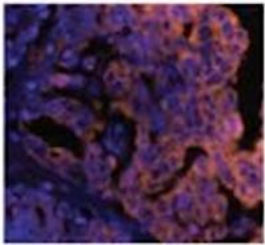


TCGA Microbiome + Fungal Reads

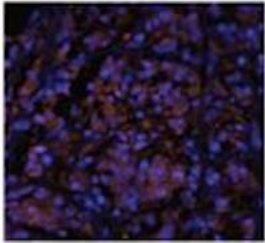


Previous Studies in Cancer Microbiome

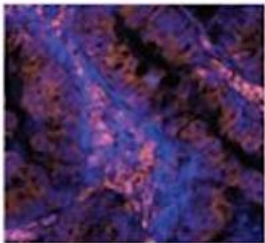
Ovarian
Cancer



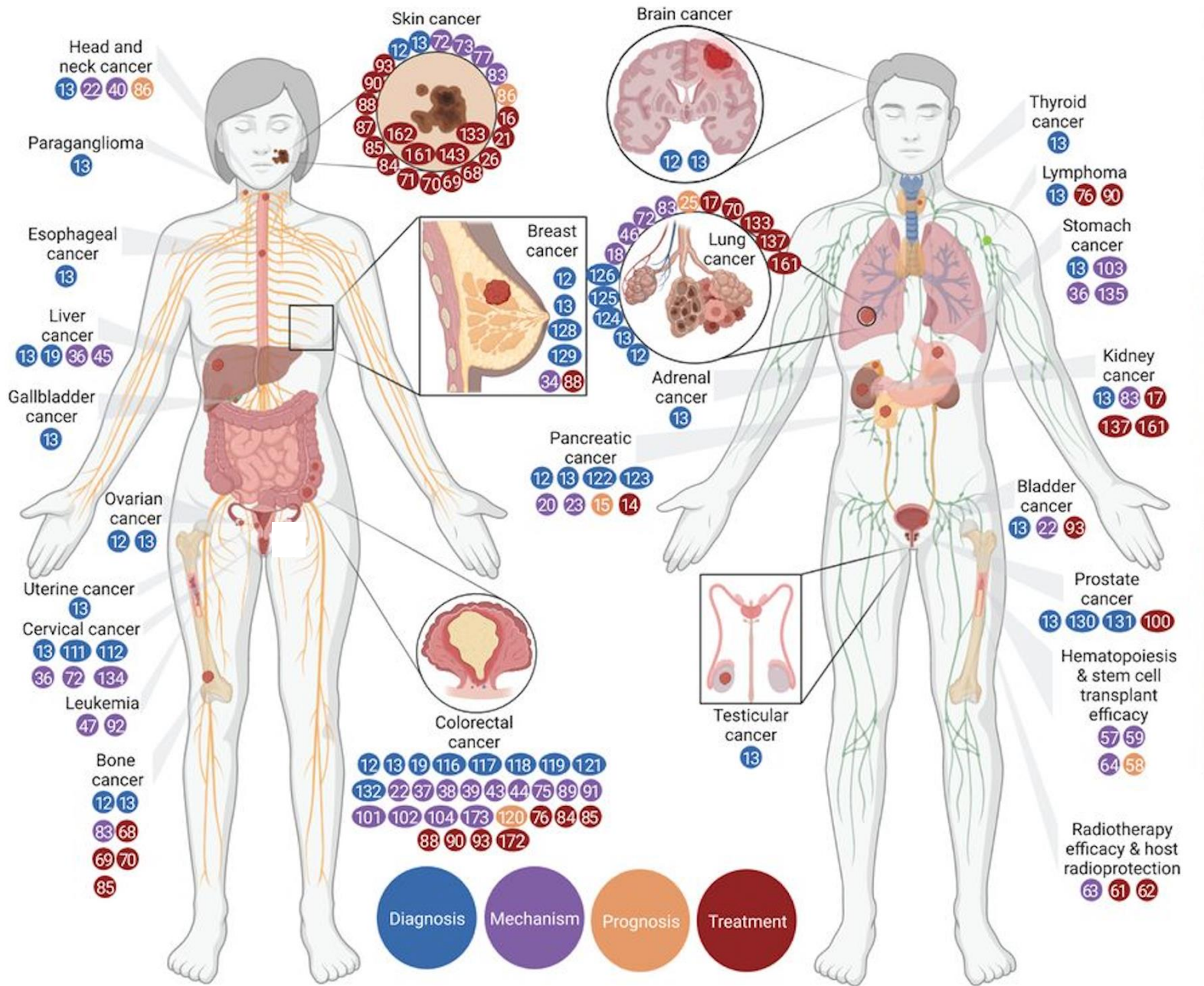
Melanoma



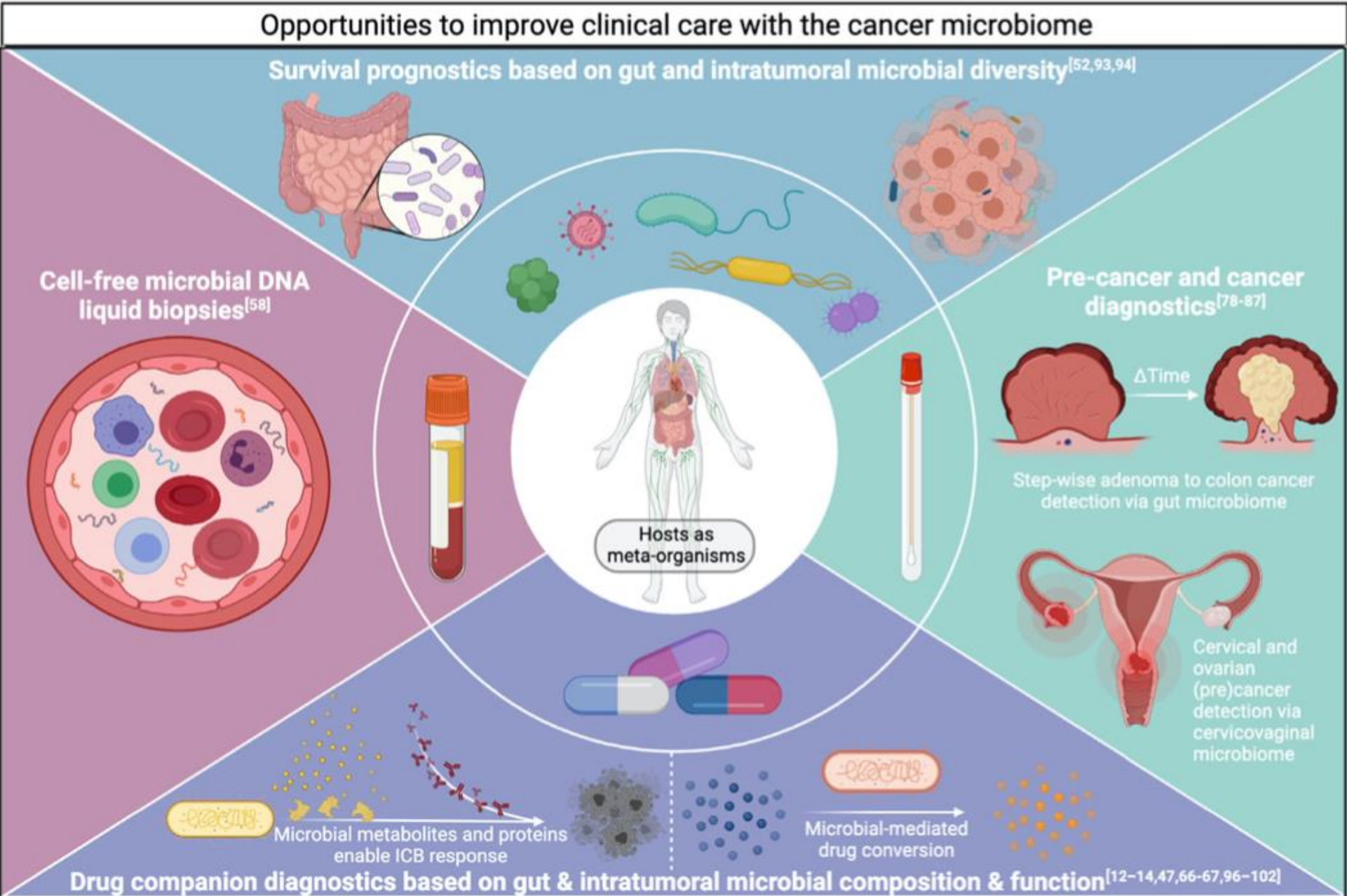
Lung Cancer



Immunofluorescence
for bacterial 16S rRNA



Future Directions of Cancer Microbiome

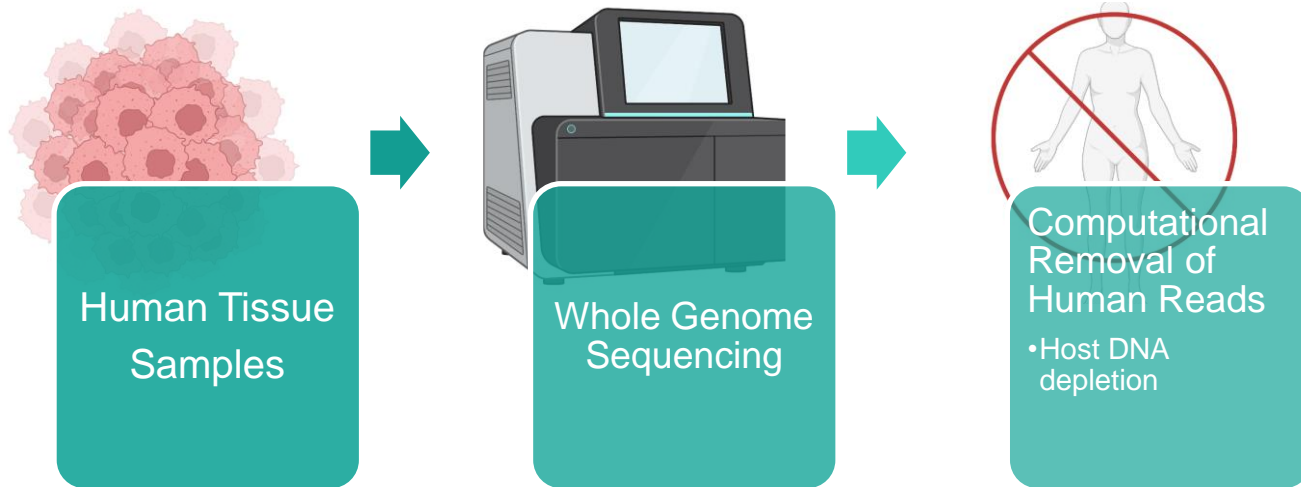


G. Sepich-Poore*, C. Guccione* et al. *BioEssays* (2022)



Define a pipeline that extracts microbial reads from whole genome sequencing samples of tumor tissue

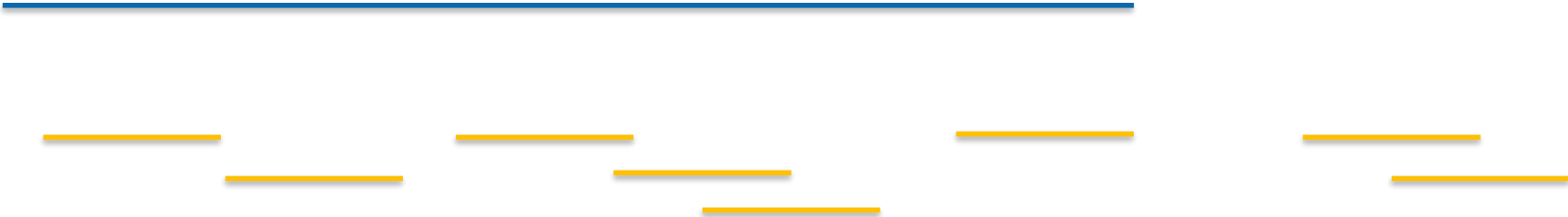
Going from whole genome sequencing to microbial taxonomy



Computational Removal of Human Reads

Human Reference: GRCh38

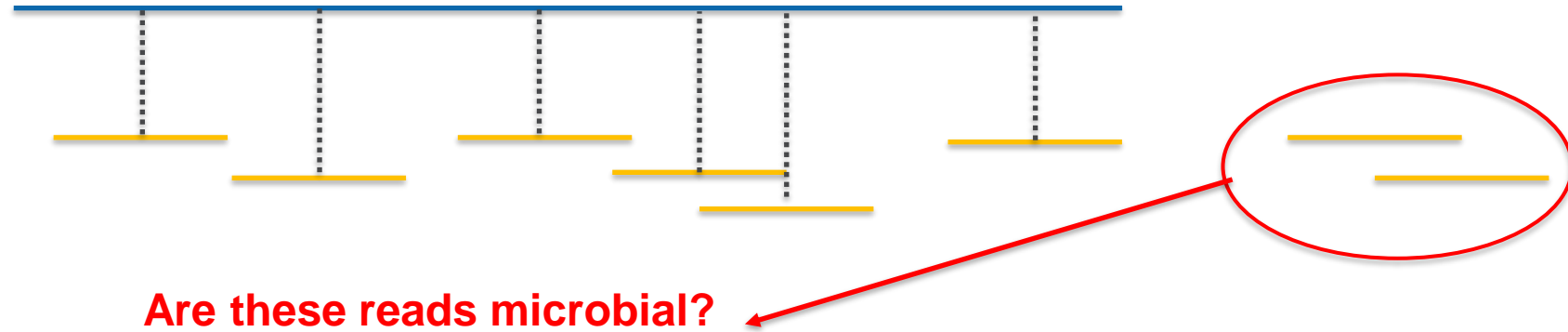
Experimental Reads



Computational Removal of Human Reads

Human Reference: GRCh38

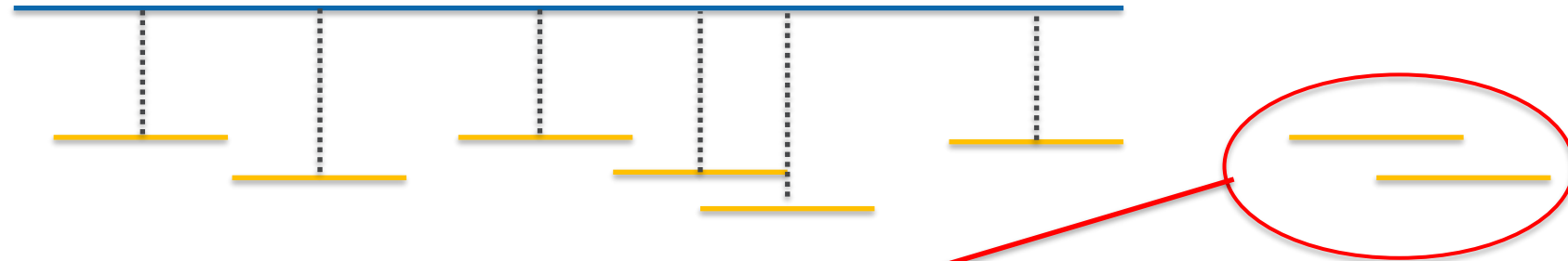
Experimental Reads



Computational Removal of Human Reads

Human Reference: GRCh38

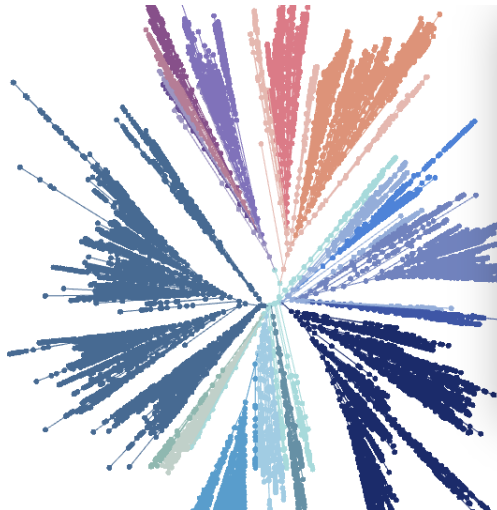
Experimental Reads



Are these reads microbial?

Microbial Reference

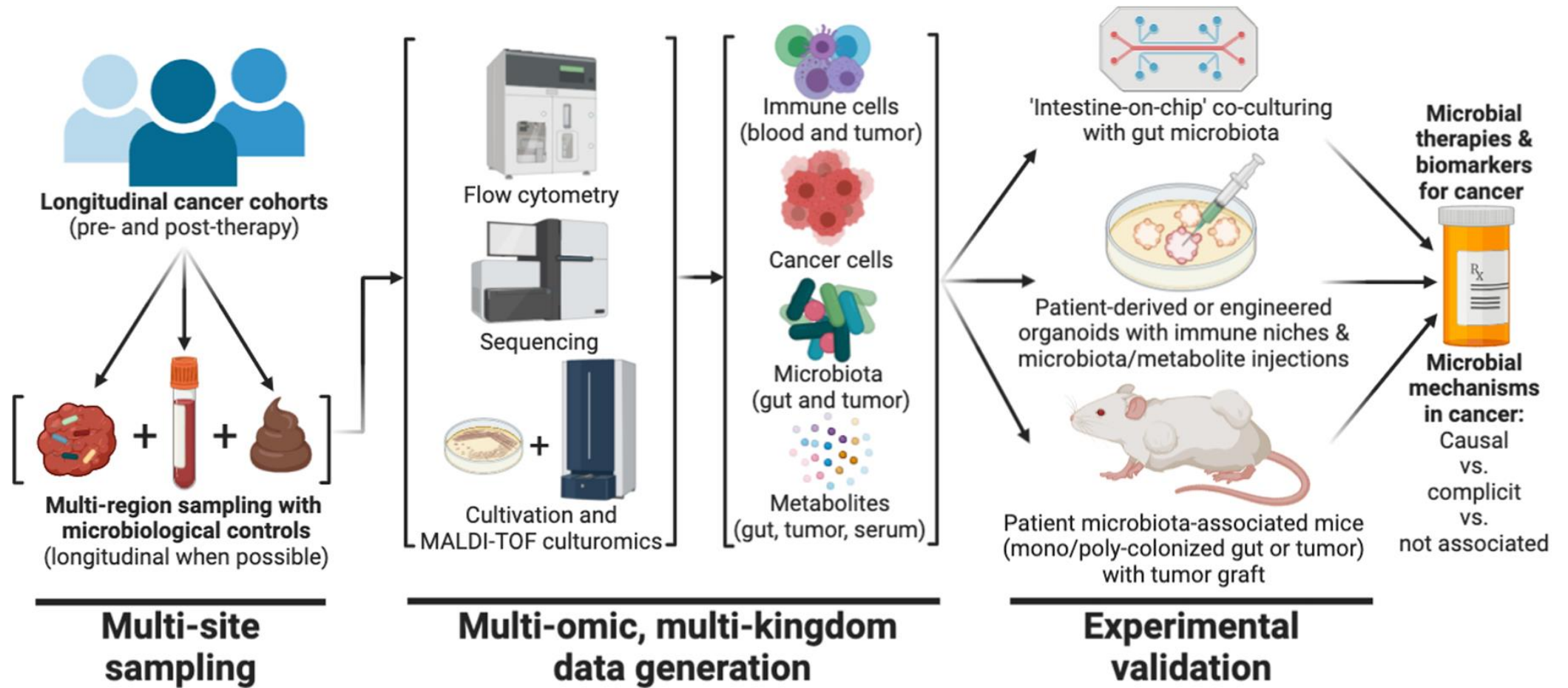
- preset
- Archaea
- Asgard
- TACK
- Crenarchaeota
- Euryarchaeota
- CPR
- Microgenomates



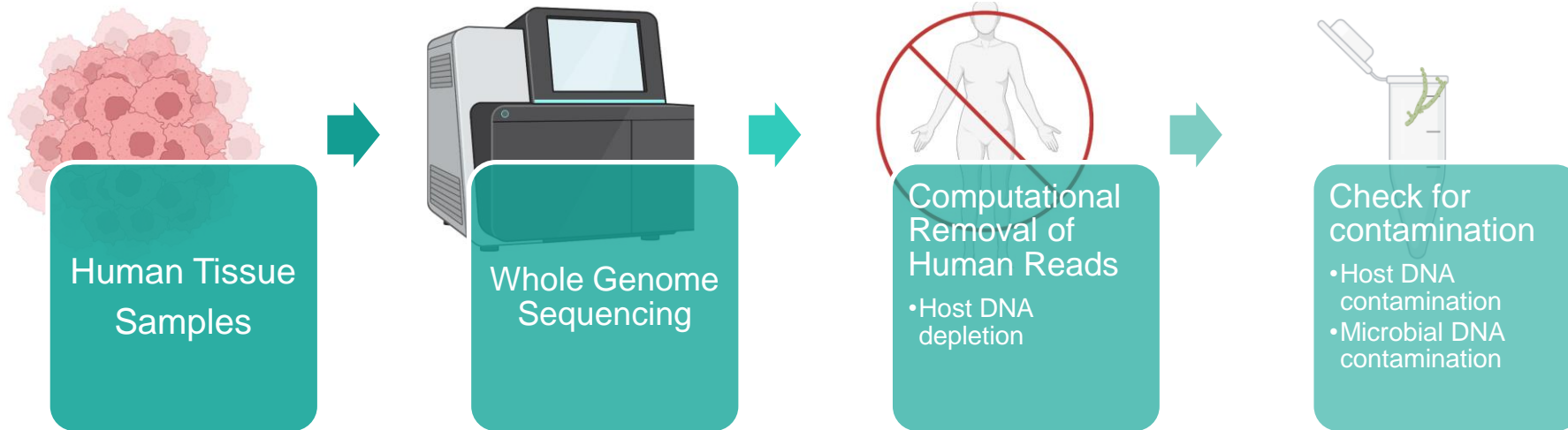
	Sample1	Sample2	Sample3	Sample4	Sample5
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__g_s_haloarchaeon 3A1_DGR	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Haloarculaceae;g__Haloarcula;s__Haloarcula vallismortis	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Haloarculaceae;g__Halomicrobium;s__Halomicrobium mukohataei	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Haloarculaceae;g__Halorhabdus;s__Halorhabdus tiamatea	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Haloarculaceae;g__Halorientalis;s__Halorientalis persicus	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Haloarculaceae;g__Natronomonas;s__Natronomonas moolapensis	0	0	2	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Halobacteriaceae;g__Halarchaeum;s__Halarchaeum acidiphilum	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Halobacteriaceae;g__Halobacterium;s__Halobacterium hubeiense	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Halobacteriaceae;g__Halobacterium;s__Halobacterium salinarum	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Halobacteriaceae;g__Halobacterium;s__Halobacterium sp. DL1	0	0	5	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferales;f__Haloferales;g__Halobellus;s__Halobellus rufus	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferales;f__Haloferales;g__Haloferales;s__Haloferales sp. S829	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferales;f__Haloferales;g__Haloferales;s__Haloferales pallidum	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferales;f__Haloferales;g__Haloplanus;s__Haloplanus natans	0	4	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferales;f__Haloferales;g__Haloquadratum;s__Haloquadratum walsbyi	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferales;f__Haloferales;g__Halopenitus;s__Halopenitus sp. DYS4	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferales;f__Haloferales;g__Halorubrum;s__Halorubrum californiense	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferales;f__Haloferales;g__Halorubrum;s__Halorubrum halophilum	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferales;f__Haloferales;g__Halorubrum;s__Halorubrum kocunii	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferales;f__Haloferales;g__Halorubrum;s__Halorubrum lipolyticum	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferales;f__Haloferales;g__Halorubrum;s__Halorubrum sp. BV1	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferales;f__Haloferales;g__Halorubrum;s__Halorubrum tebenquichense	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Natrialbales;f__Natrialbales;g__Halostagnicola;s__Halostagnicola sp. A56	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Natrialbales;f__Natrialbales;g__Natriema;s__Natriema pellibrum	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Natrialbales;f__Natrialbales;g__Natronorubrum;s__Natronorubrum tibetense	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Methanomicrobiales;f__Methanomicrobiales;g__Methanolacinia;s__Methanolacinia	443	2565	2276	1717	1337
k__Archaea;p__Euryarchaeota;c__Methanomicrobiales;f__Methanomicrobiales;g__Methanosarcina;s__Methanosarcina m	0	0	0	0	0
k__Bacteria;p__c__o__f__g_s_bacteria symbiont Bfo2 of Frankiniella occidentalis	0	0	0	0	0

Taxonomy Table

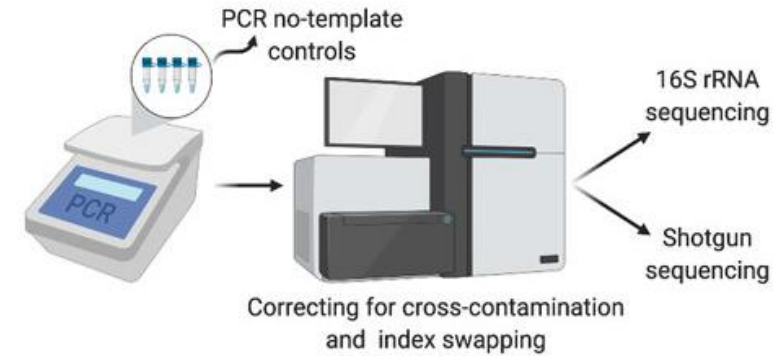
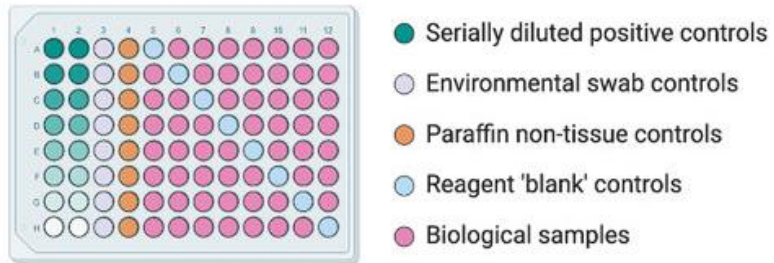
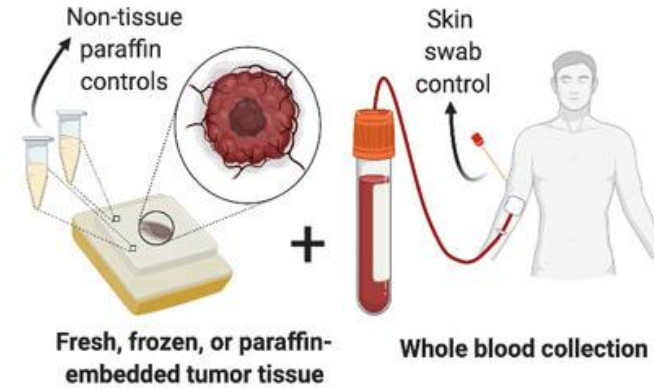
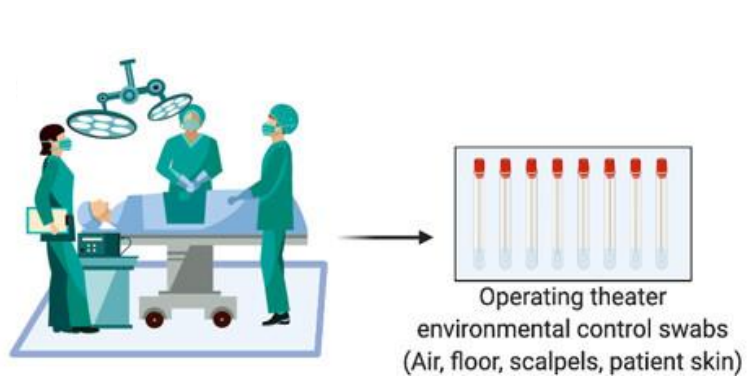
Is there a better way to study the cancer microbiome?



Going from whole genome sequencing to microbial taxonomy



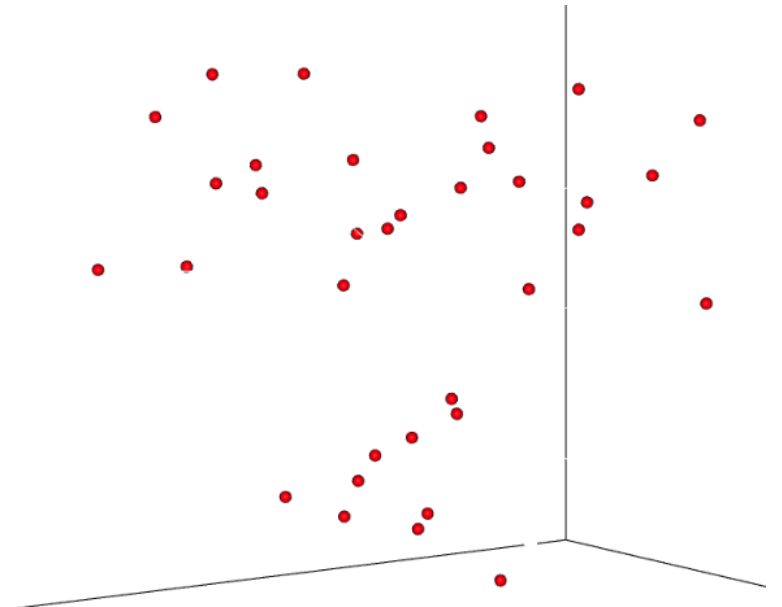
How can we check for contamination?



How can we check for contamination computationally?

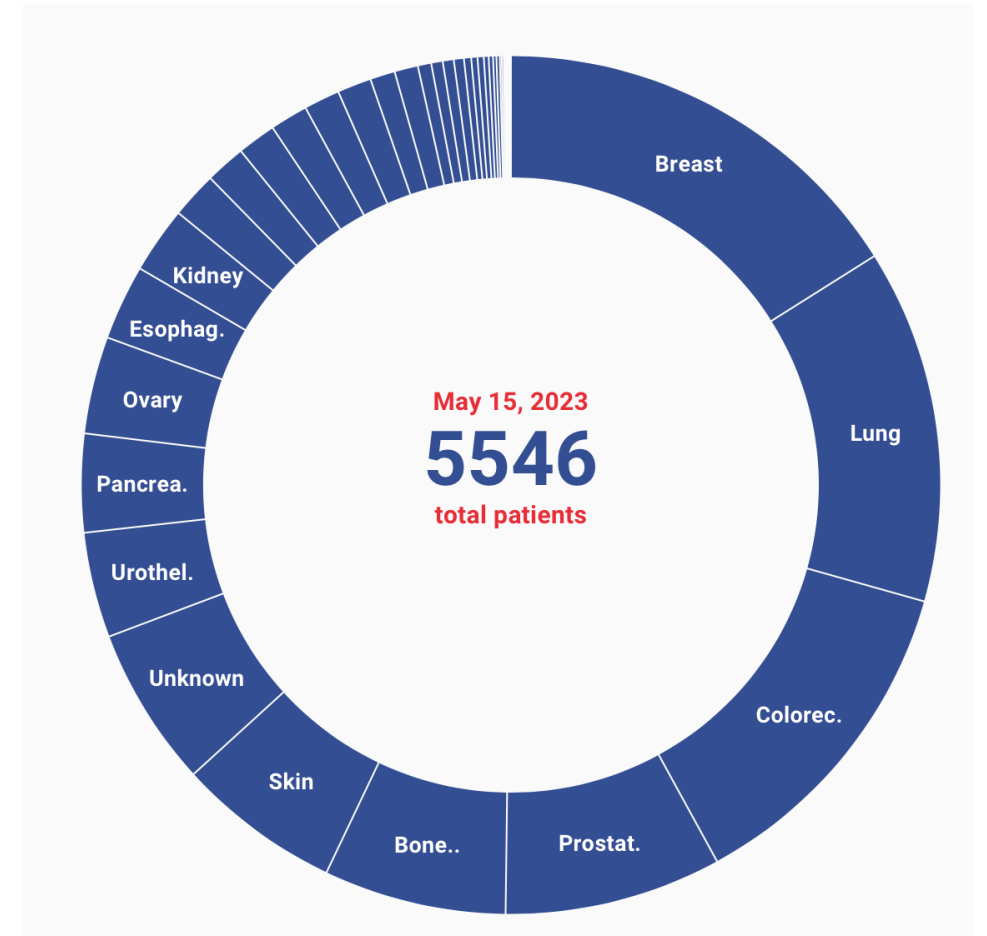
- Determine relationships between samples using Robust Aitchison PCA (RPCA)
- Uses matrix completion so zero values do not influence the ordination

	Sample1	Sample2	Sample3	Sample4	Sample5
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__g__;s__haloarchaeon 3A1_DGR	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Haloarculaceae;g__Haloarcula;s__Haloarcula vallismortis	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Haloarculaceae;g__Haloarcula;s__Halomicrobium mukohataei	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Haloarculaceae;g__Halorhabdus;s__Halorhabdus tiamatea	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Haloarculaceae;g__Halorientalis;s__Halorientalis persicus	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Haloarculaceae;g__Natronomonas;s__Natronomonas moolapensis	0	0	2	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Halobacteriaceae;g__Halarchaeum;s__Halarchaeum acidiphilum	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Halobacteriaceae;g__Halobacterium;s__Halobacterium hubeiense	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Halobacteriaceae;g__Halobacterium;s__Halobacterium salinarum	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Halobacteriales;f__Halobacteriaceae;g__Halobacterium;s__Halobacterium sp. DL1	0	0	5	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferacales;f__Haloferacaceae;g__Halobellus;s__Halobellus rufus	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferacales;f__Haloferacaceae;g__Haloferax;s__Haloferax sp. SB29	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferacales;f__Haloferacaceae;g__Halogeometricum;s__Halogeometricum pallidum	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferacales;f__Haloferacaceae;g__Haloplanus;s__Haloplanus natans	0	4	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferacales;f__Haloferacaceae;g__Haloquadratum;s__Haloquadratum walsbyi	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferacales;f__Halorubraceae;g__Halopenitus;s__Halopenitus sp. DYS4	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferacales;f__Halorubraceae;g__Halorubrum;s__Halorubrum californiense	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferacales;f__Halorubraceae;g__Halorubrum;s__Halorubrum halophilum	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferacales;f__Halorubraceae;g__Halorubrum;s__Halorubrum kocurii	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferacales;f__Halorubraceae;g__Halorubrum;s__Halorubrum lipolyticum	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferacales;f__Halorubraceae;g__Halorubrum;s__Halorubrum sp. BV1	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Haloferacales;f__Halorubraceae;g__Halorubrum;s__Halorubrum tebenquichense	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Natrialbales;f__Natrialbaceae;g__Halostagnicola;s__Halostagnicola sp. A56	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Natrialbales;f__Natrialbaceae;g__Natrinema;s__Natrinema pellibrum	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Halobacteria;o__Natrialbales;f__Natrialbaceae;g__Natronorubrum;s__Natronorubrum tibetense	0	0	0	0	0
k__Archaea;p__Euryarchaeota;c__Methanomicrobia;o__Methanomicrobiales;f__Methanomicrobiaceae;g__Methanolacinia;s__Methanolacinia	443	2565	2276	1717	1337
k__Archaea;p__Euryarchaeota;c__Methanomicrobia;o__Methanosarcinales;f__Methanosarcinaceae;g__Methanosarcina;s__Methanosarcina m	0	0	0	0	0
k__Bacteria;p__c__;o__f__;g__;s__bacteria symbiont BFo2 of Frankiniella occidentalis	0	0	0	0	0

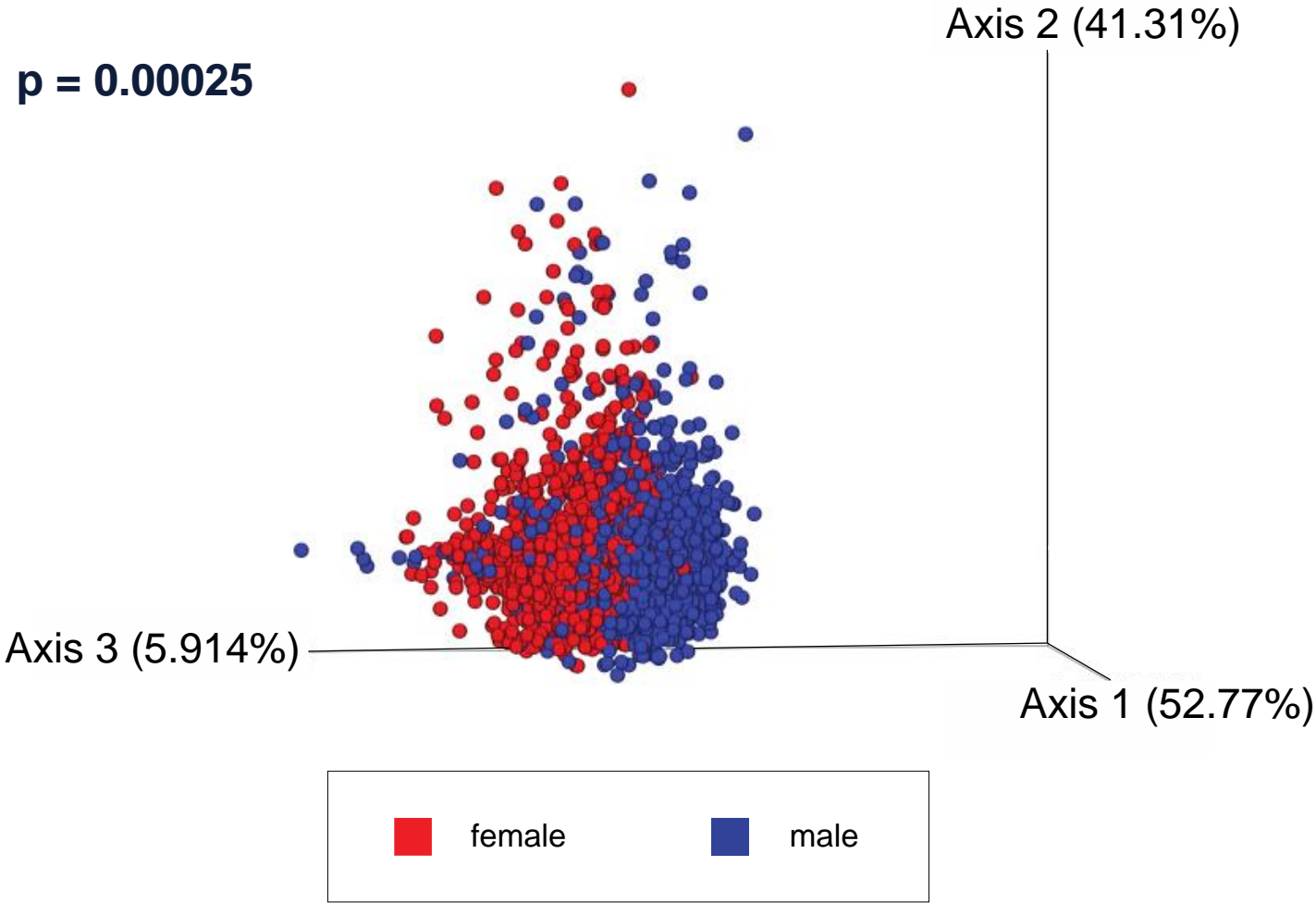


Hartwig Medical Foundation (HMF) Dataset

- Contains over 5,000 metastatic human tissue samples
- Whole Genome Sequencing (WGS)
- Resulting in low-depth microbial reads

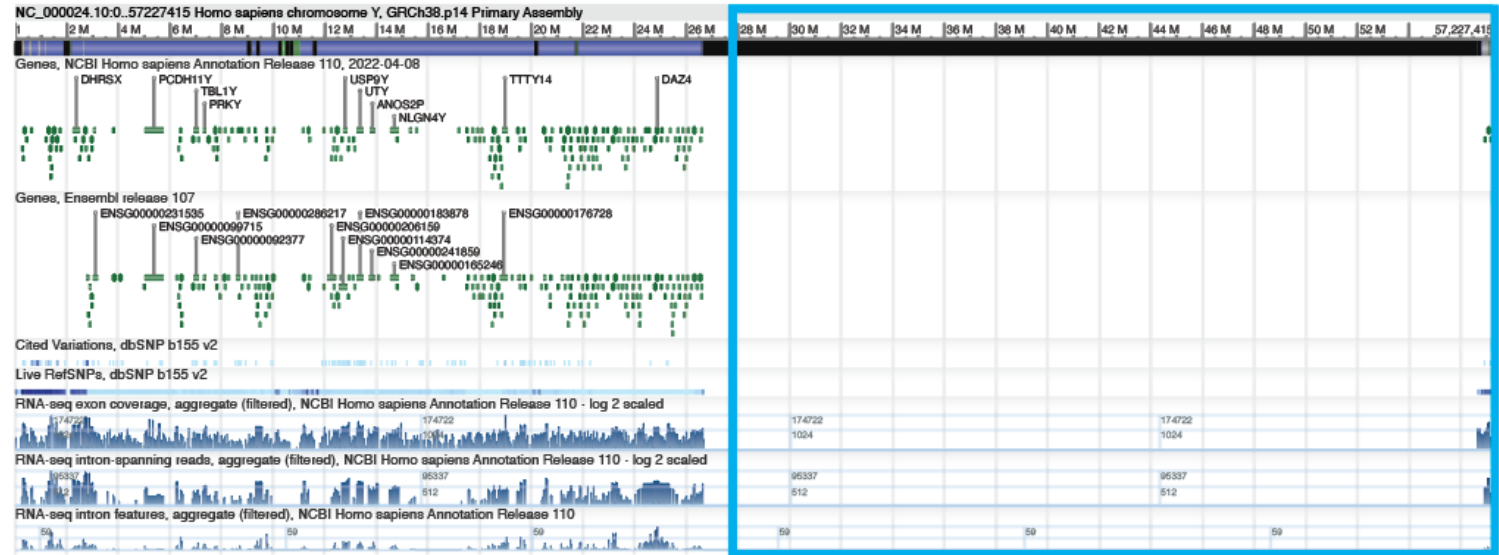


WGS of HMF data with GRCh38 host depletion shows splitting across sex in RPCA plot



GRCh38

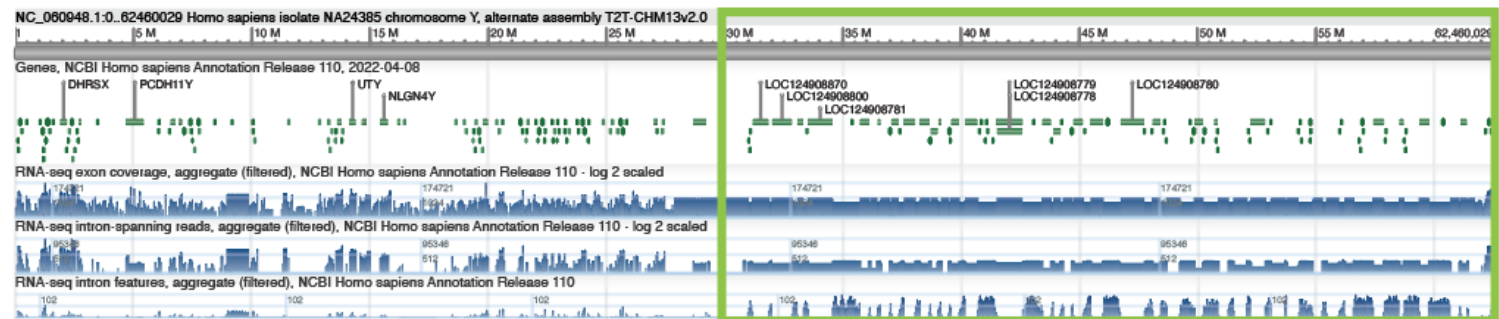
Released 2013



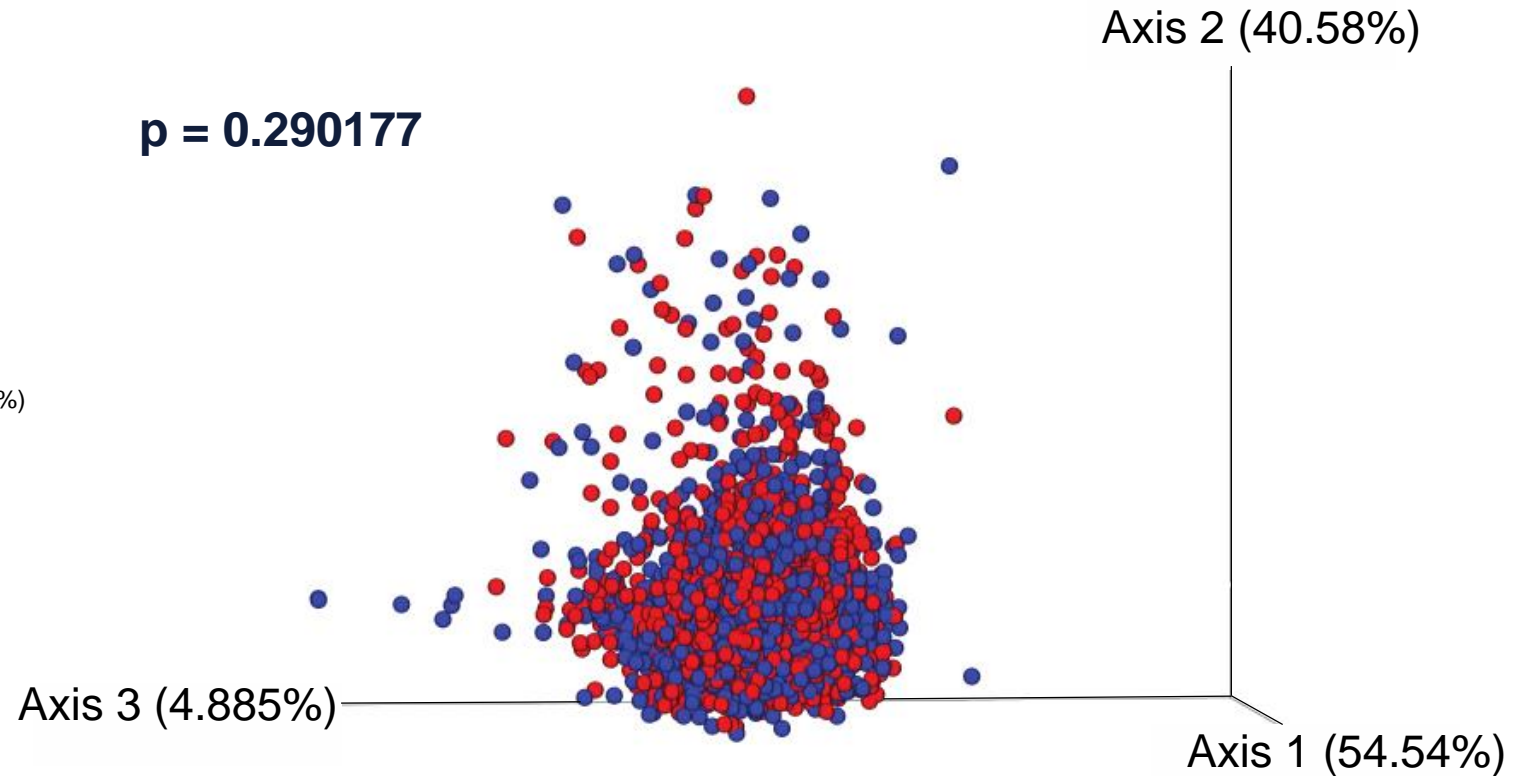
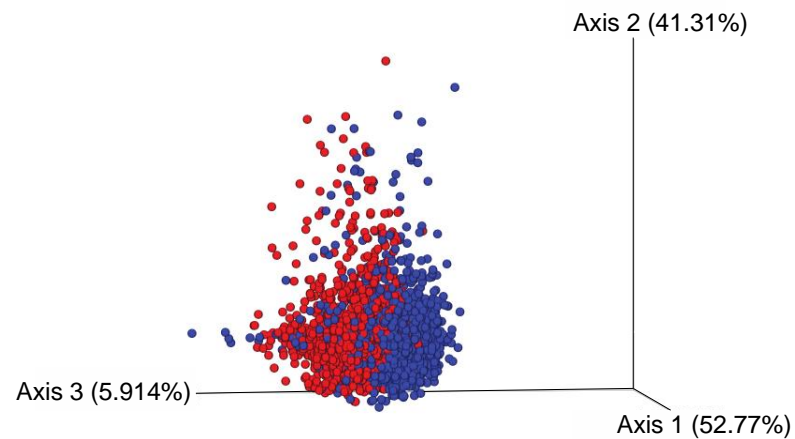
Improved mapping of human Y chromosome reference

T2T-CHM13

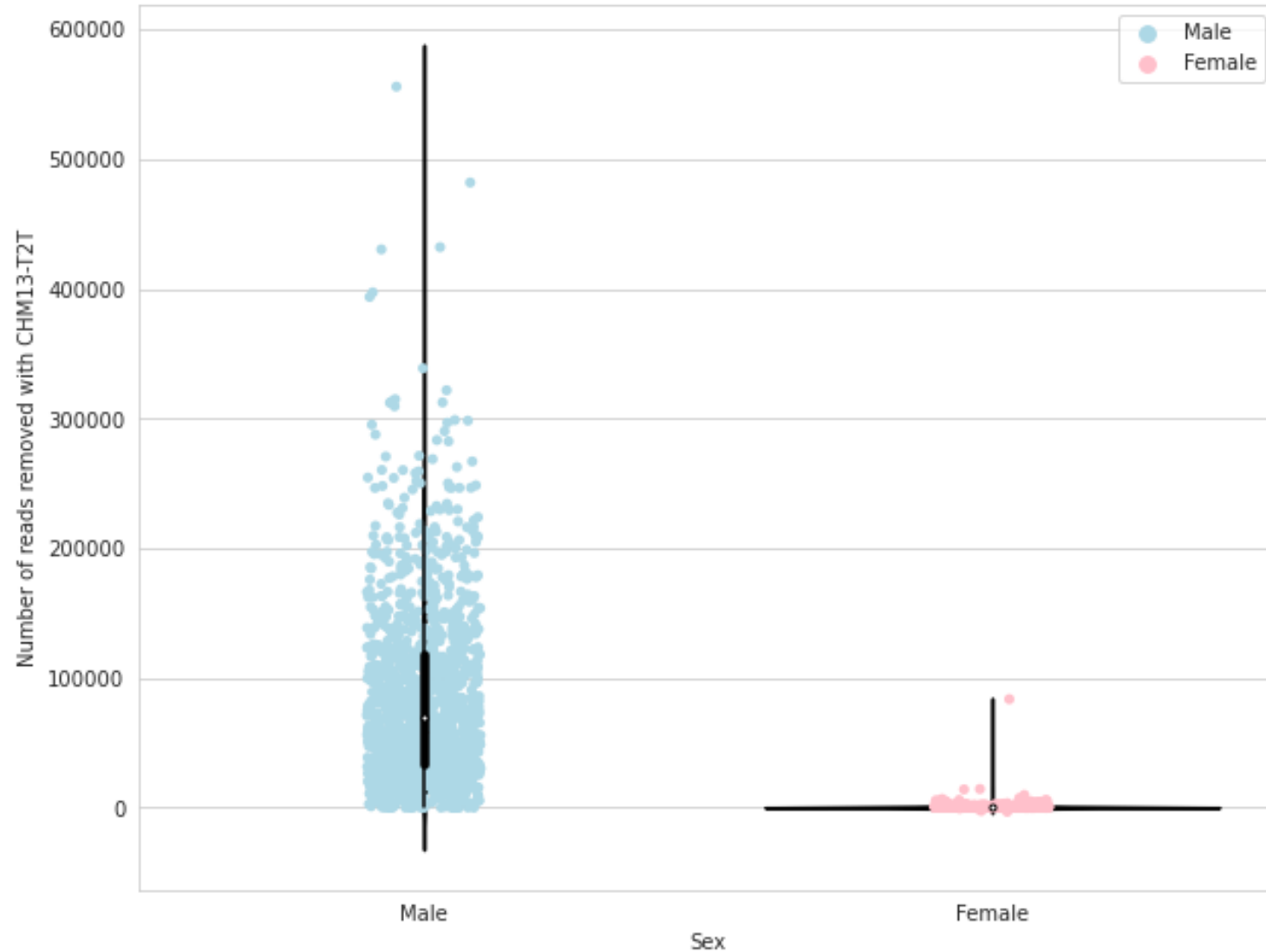
Released 2022



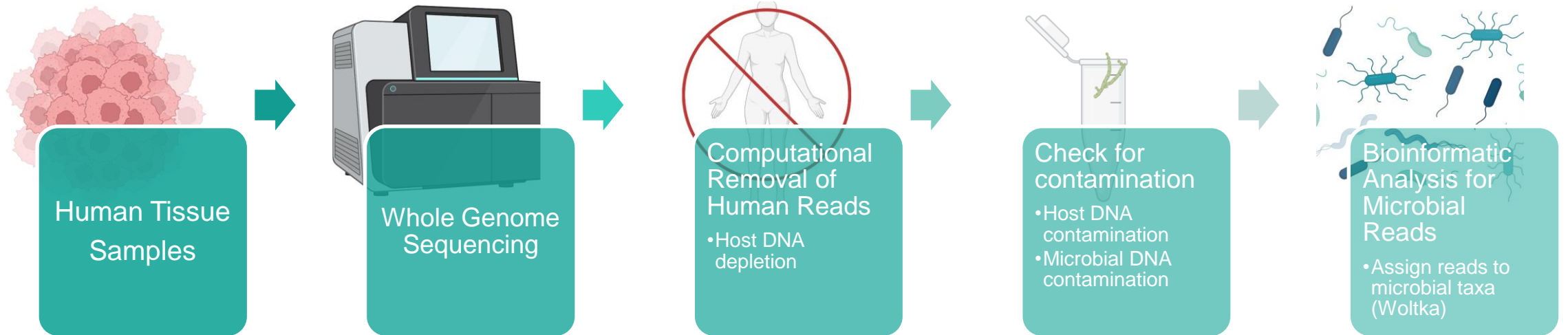
WGS of HMF data with GRCh38 + T2T-CHM13 host depletion shows no splitting across sex in RPCA plot



Male samples have many more reads removed using CHM13-T2T compared with females

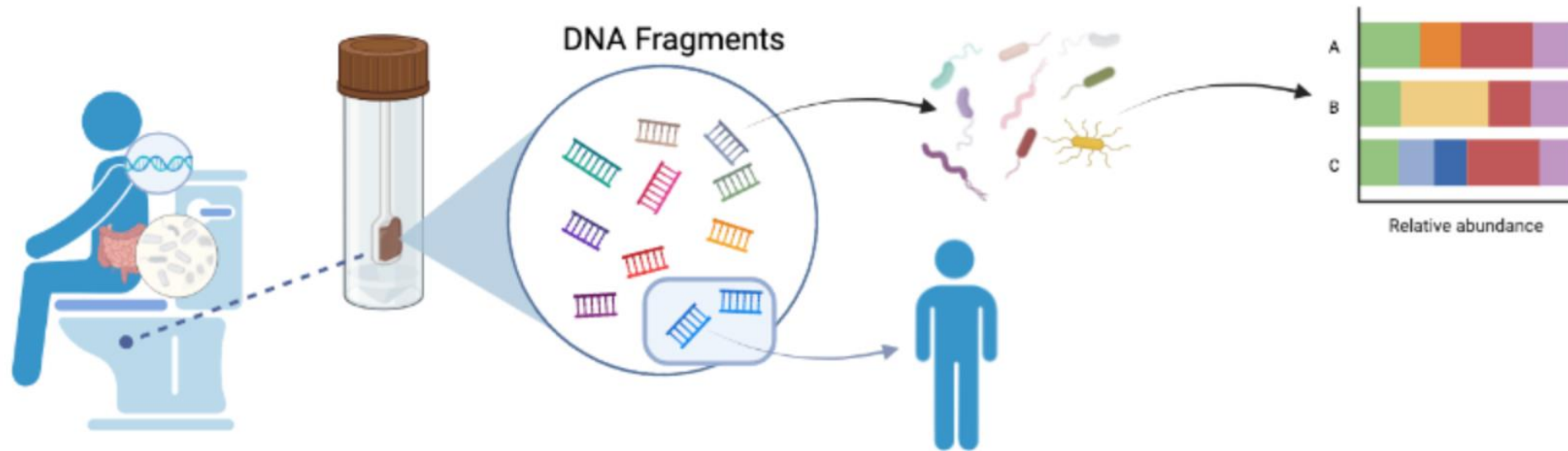


Going from whole genome sequencing to microbial taxonomy



Using fecal microbiome studies to check our host depletion pipeline

Typical fecal microbiome study workflow:



Fecal samples are often publicly accessible because they are thought to contain such small amounts of human reads especially after host filtering

Current methods are not removing all the host data

nature microbiology

Explore content ▾ About the journal ▾ Publish with us ▾

nature > nature microbiology > articles > article

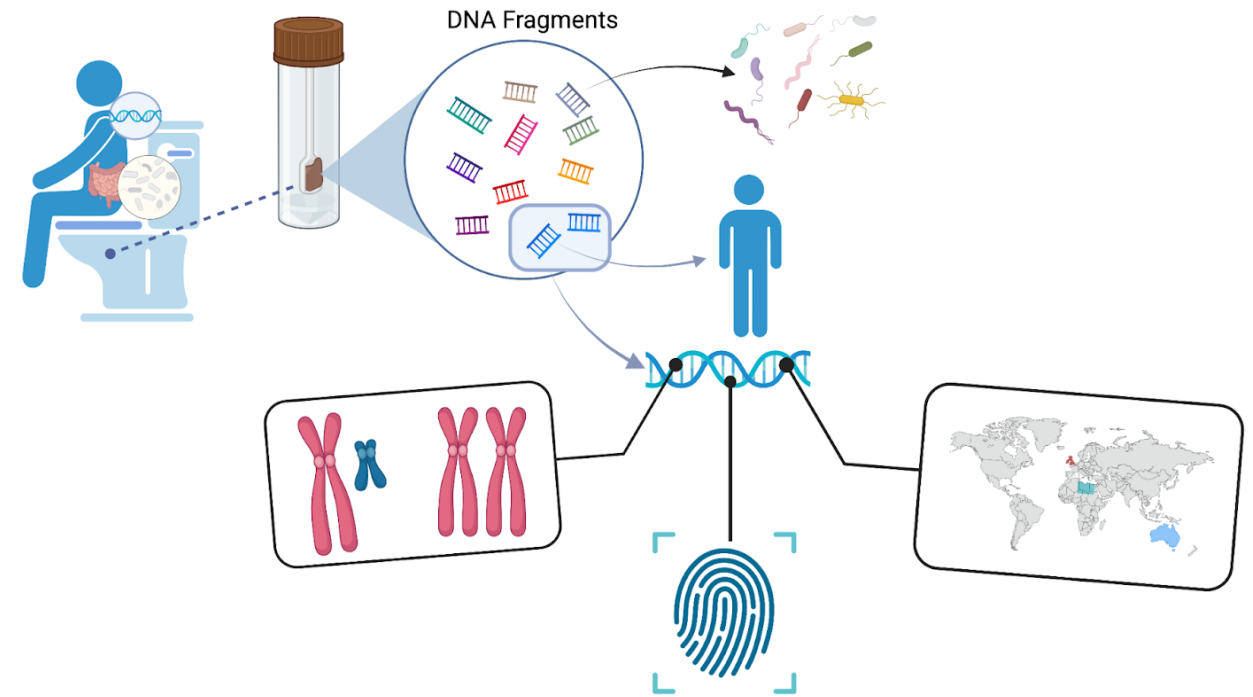
Article | [Open Access](#) | Published: 15 May 2023

Reconstruction of the personal information from human genome reads in gut metagenome sequencing data

[Yoshihiko Tomofuji](#) , [Kyuto Sonehara](#), [Toshihiro Kishikawa](#), [Yuichi Maeda](#), [Kotaro Ogawa](#), [Shuhei Kawabata](#), [Takuro Nii](#), [Tatsusada Okuno](#), [Eri Oguro-Igashira](#), [Makoto Kinoshita](#), [Masatoshi Takagaki](#), [Kenichi Yamamoto](#), [Takashi Kurakawa](#), [Mayu Yagita-Sakamaki](#), [Akiko Hosokawa](#), [Daisuke Motooka](#), [Yuki Matsumoto](#), [Hidetoshi Matsuoka](#), [Maiko Yoshimura](#), [Shiro Ohshima](#), [Shota Nakamura](#), [Hidenori Inohara](#), [Haruhiko Kishima](#), [Hideki Mochizuki](#), ... [Yukinori Okada](#)  [+ Show authors](#)

Nature Microbiology (2023) | [Cite this article](#)

89 Altmetric | [Metrics](#)



Using the pangenome for host depletion

nature

Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [articles](#) > article

Article | [Open Access](#) | [Published: 10 May 2023](#)

A draft human pangenome reference

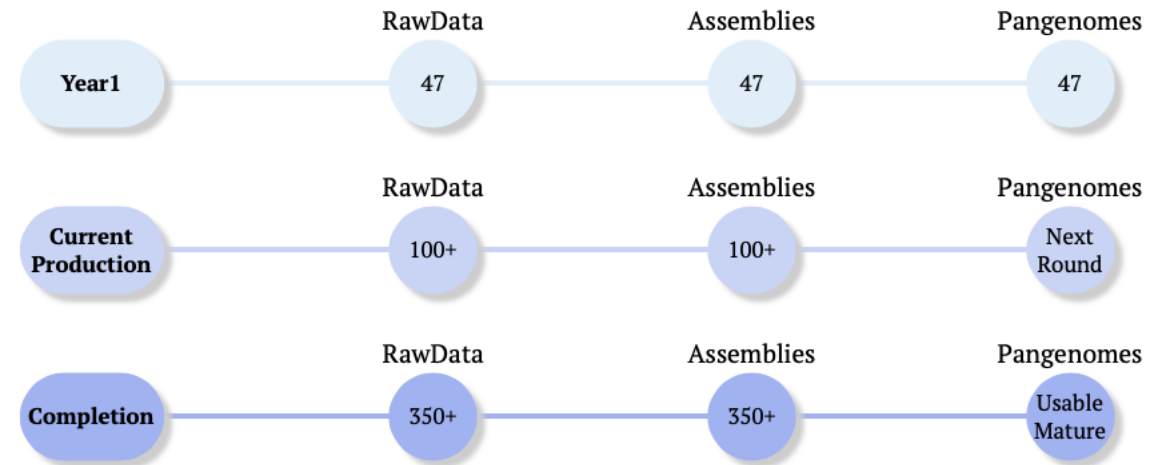
[Wen-Wei Liao](#), [Mobin Asri](#), [Jana Ebler](#), [Daniel Doerr](#), [Marina Haukness](#), [Glenn Hickey](#), [Shuangjia Lu](#), [Julian K. Lucas](#), [Jean Monlong](#), [Haley J. Abel](#), [Silvia Buonaiuto](#), [Xian H. Chang](#), [Haoyu Cheng](#), [Justin Chu](#), [Vincenza Colonna](#), [Jordan M. Eizenga](#), [Xiaowen Feng](#), [Christian Fischer](#), [Robert S. Fulton](#), [Shilpa Garg](#), [Cristian Groza](#), [Andrea Guarracino](#), [William T. Harvey](#), [Simon Heumos](#), ... [Benedict Paten](#)  [+ Show authors](#)

[Nature](#) **617**, 312–324 (2023) | [Cite this article](#)

141k Accesses | **19** Citations | **3196** Altmetric | [Metrics](#)

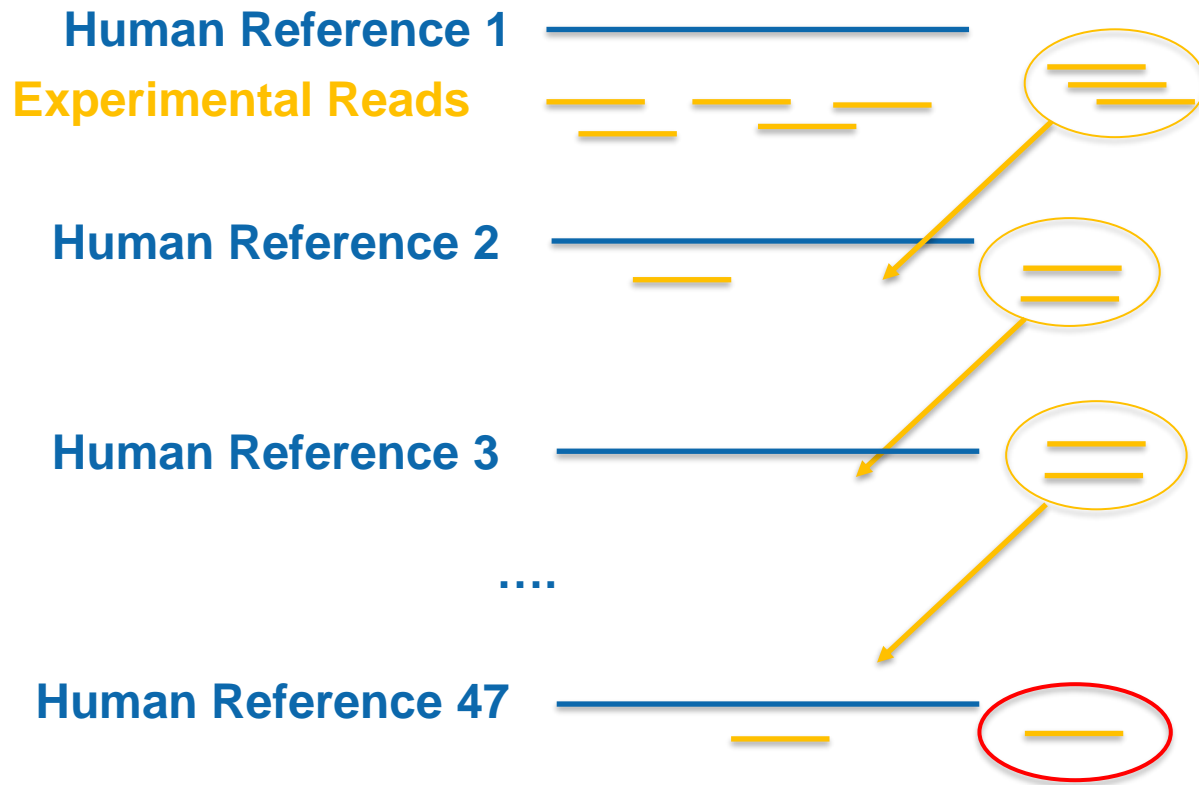


Milestones



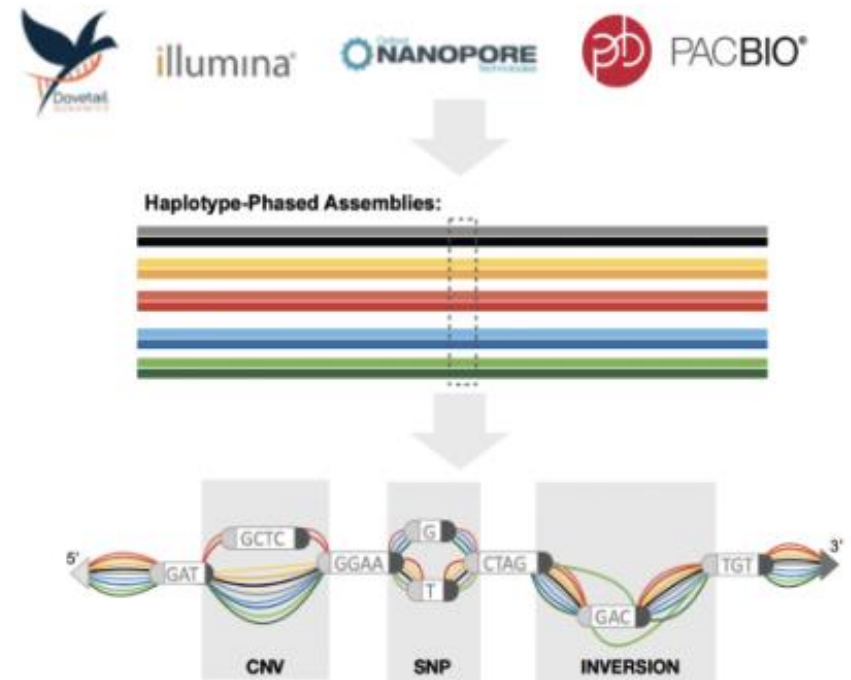
How can we use the pangenome to host deplete?

Current method:



Are these microbial?

Developing method:



Liao et al. *Nature* (2023)



Target Dataset:
Characterizing Several
Childhood Cancers



CANCER GENOMICS CLOUD
SEVEN BRIDGES

UC San Diego
School of Medicine

Target Dataset

- Therapeutically Applicable Research to Generate Effective Treatments (TARGET) in childhood cancer
- Contains over 13,692 applicable samples that can be translated to low-depth microbial reads
- Various cancer types including:
 - Acute Lymphoblastic Leukemia
 - Acute Myeloid Leukemia
 - Neuroblastoma
 - Osteosarcoma
 - Rhabdoid Tumors
 - High-Risk Wilms Tumors



The Therapeutically Applicable Research to Generate Effective Treatments program is characterizing several childhood cancers.



Childhood Cancer Microbiome Questions

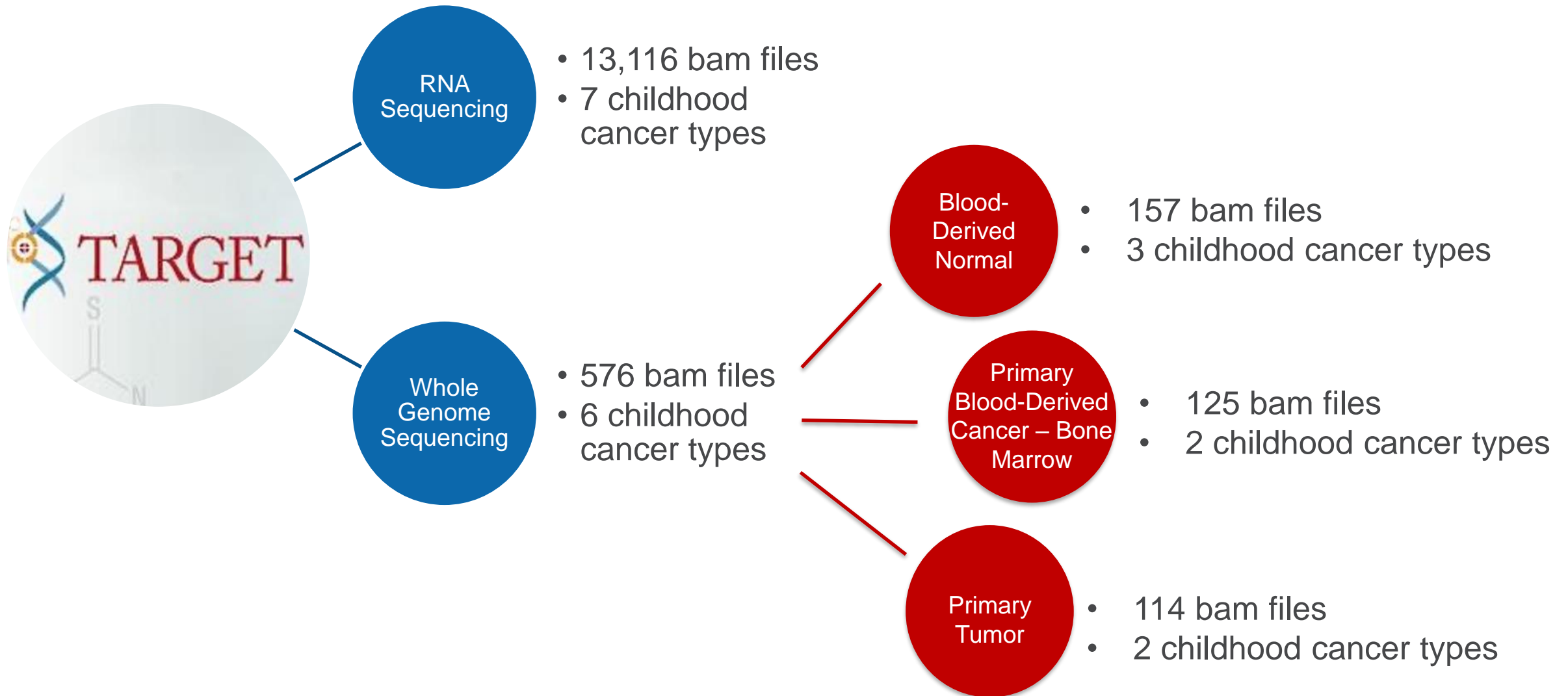
Using Target from CGC

- Can we see differences comparing one cancer type vs. all others using tumor tissue?
- Can we see differences comparing one cancer type vs. all others using 'normal' blood?
- Are the microbes represented in childhood cancers similar to those in corresponding adult cancers?

Using Additional Data Resources

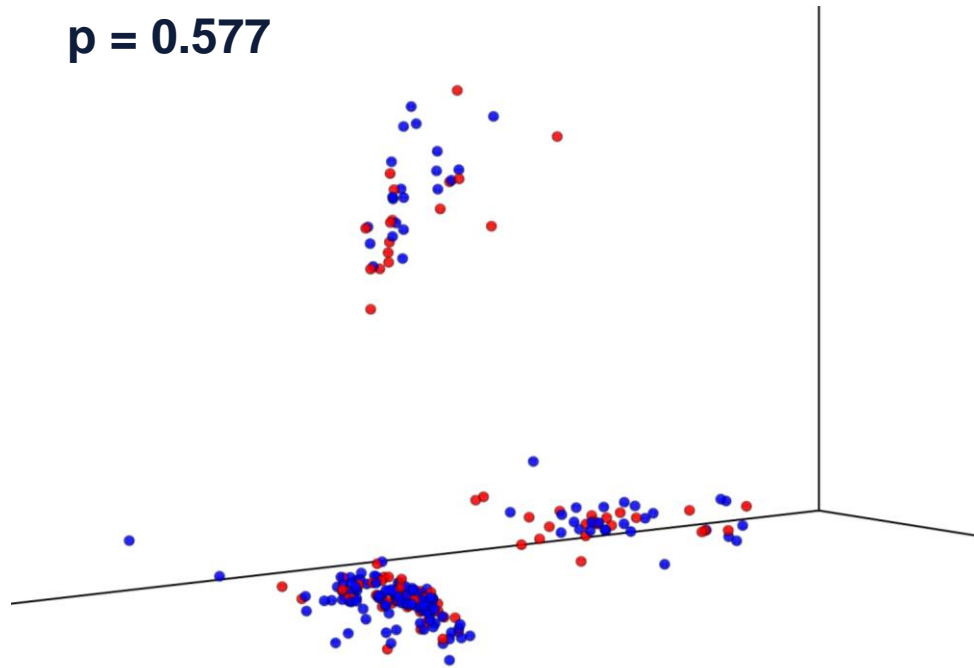
- Can we validate these findings across a larger subset of childhood cancers?
- Can we validate these findings across another dataset with the same cancer types?

Target Dataset Breakdown



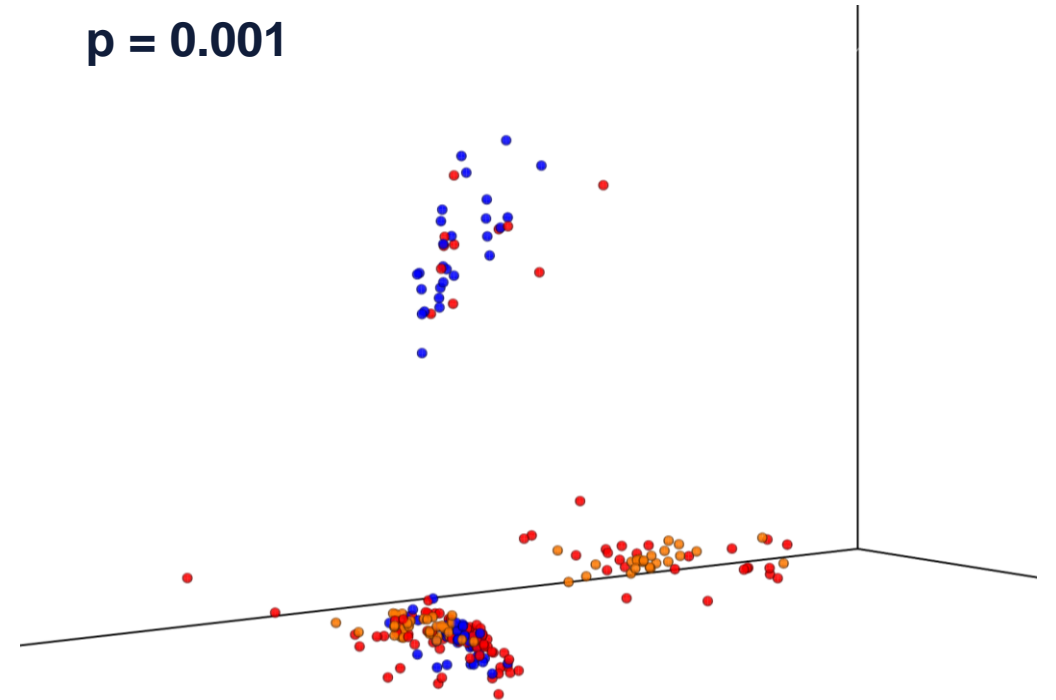
Target Dataset WGS: Checking for Contamination

$p = 0.577$



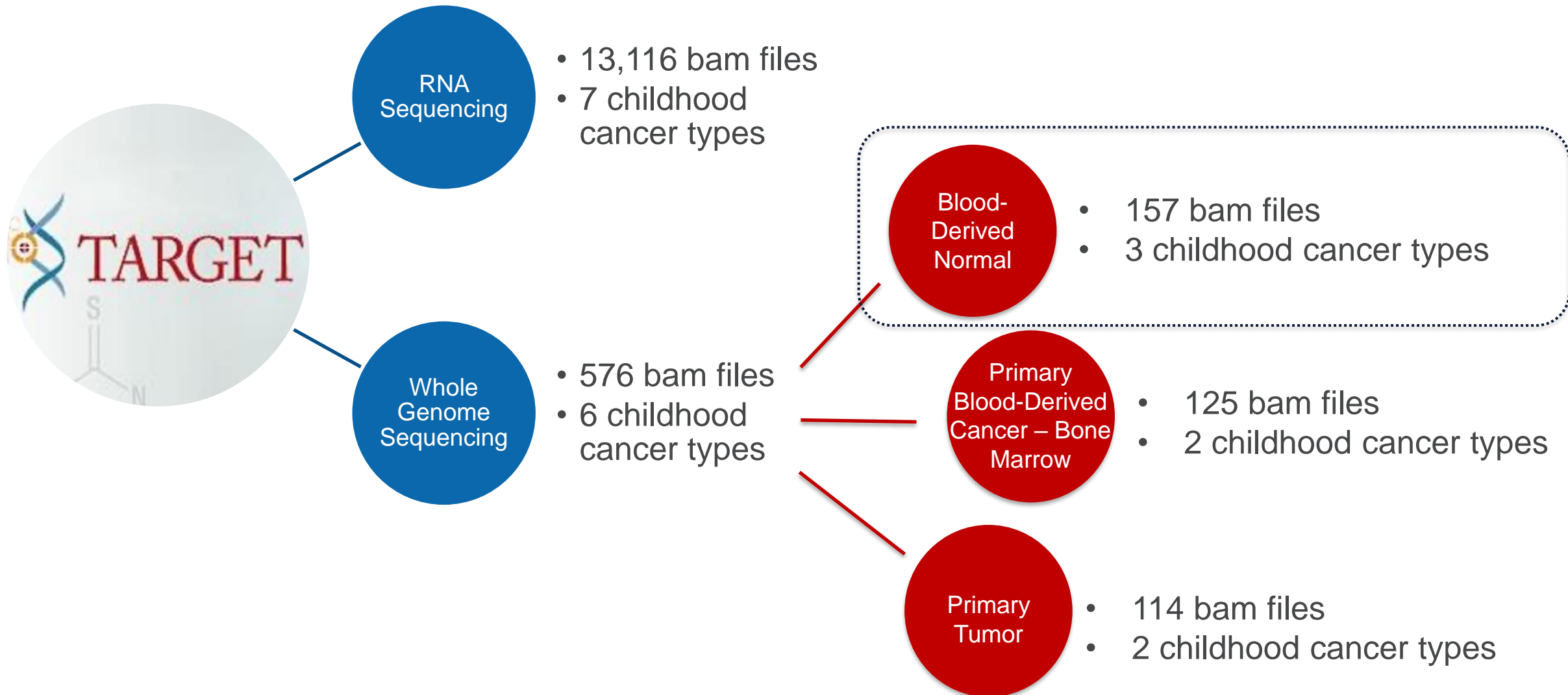
- Female
- Male
- Unknown

$p = 0.001$

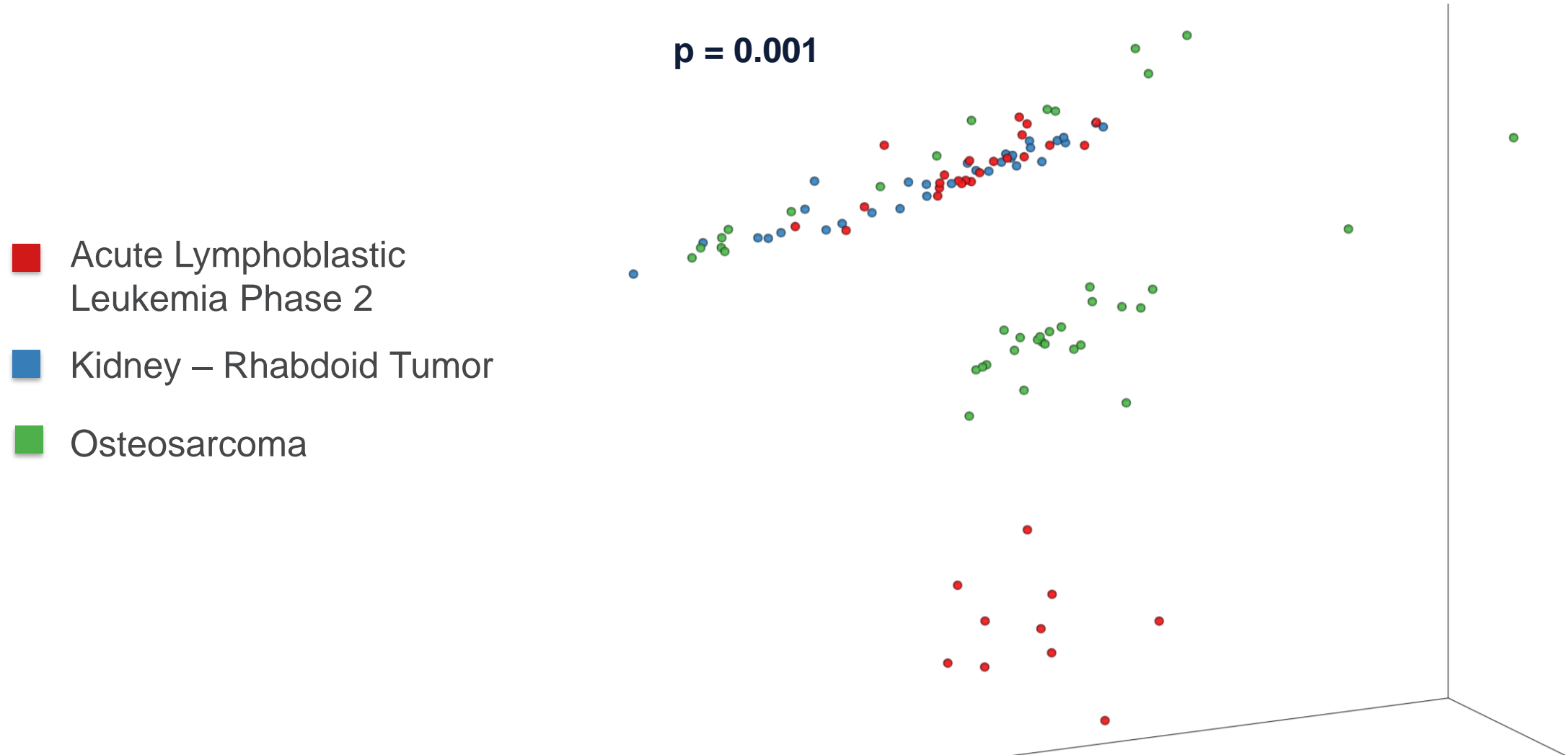


- Blood-Derived Normal
- Primary Blood-Derived Cancer – Bone Marrow
- Primary Tumor

Target Dataset Breakdown



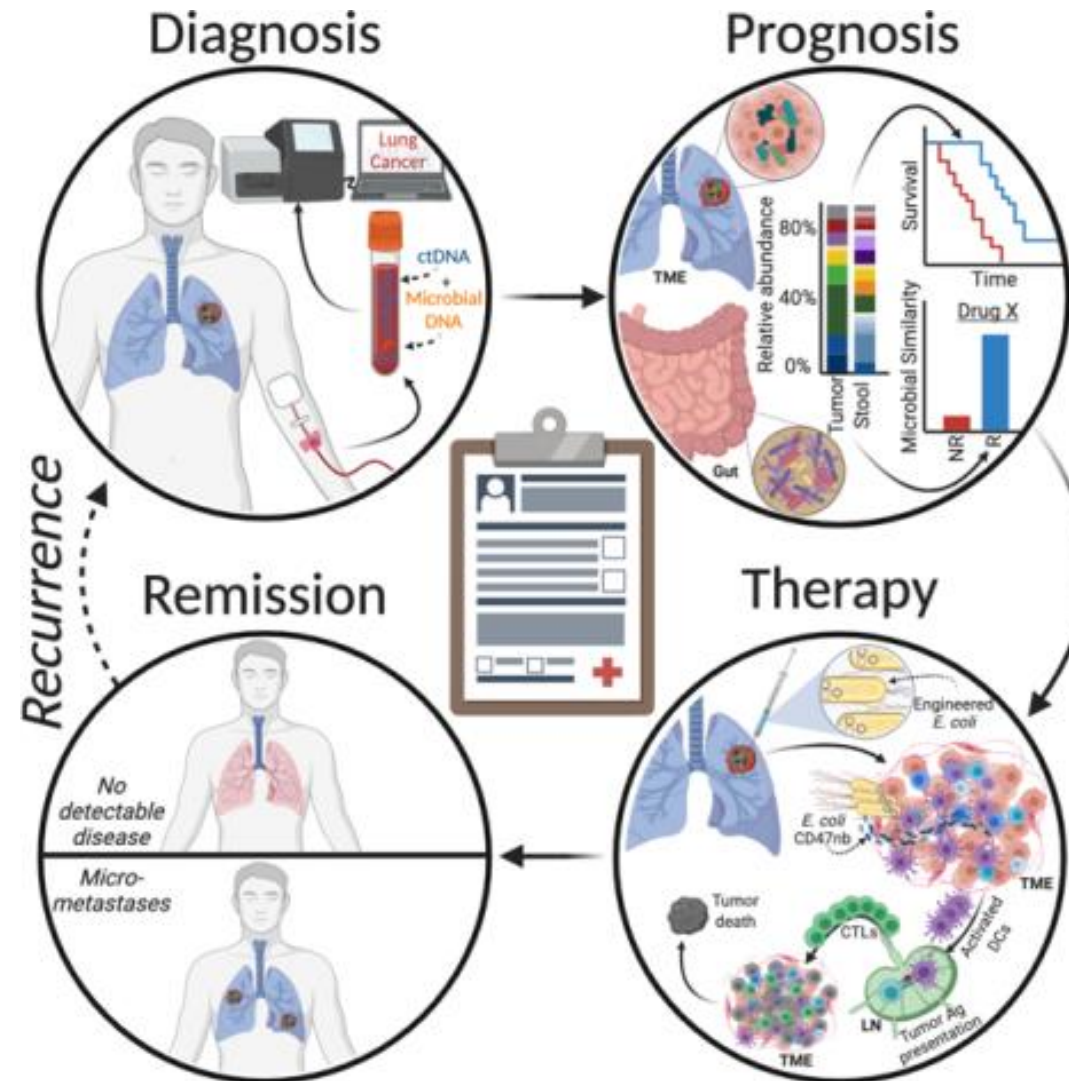
Target Dataset WGS: Blood-Derived Normal



Next steps for Target Data

- Finish analysis of RNA-Sequencing datasets
- Combine analysis of WGS + RNA-Seq as done in Poore et al. (2020)
- Run random forest machine-learning models:
 - Can we see differences comparing one cancer type vs. all others using tumor tissue?
 - Can we see differences comparing one cancer type vs. all others using 'normal' blood?
- Are the microbes represented in childhood cancers similar to those in corresponding adult cancers?

Next step for Cancer Microbiome



Acknowledgments

- Quantitative Cancer Control Lab
 - Dr. Kit Curtius

- Knight Lab
 - Dr. Rob Knight
 - Dr. Antonio Gonzalez
 - **Dr. Greg Sepich-Poore**
 - Dr. Cameron Martino

- Data and Resources:
 - **Cancer Genomics Cloud**
 - Seattle Barrett's Esophagus Cohort
 - Reid Lab at Fred Hutchinson Cancer Research Center
 - Dr. Amir Zarrinpar's Lab
 - Dr. Ludmil Alexandrov's Lab

