

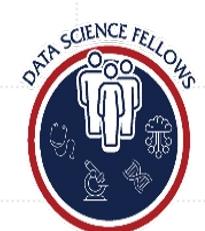


The Microbes around Us

Presented by Nicole Jimenez, PhD

University of Arizona – College of Medicine-
Phoenix

Department of Obstetrics and Gynecology



Journey to Data Science/Microbiome Science



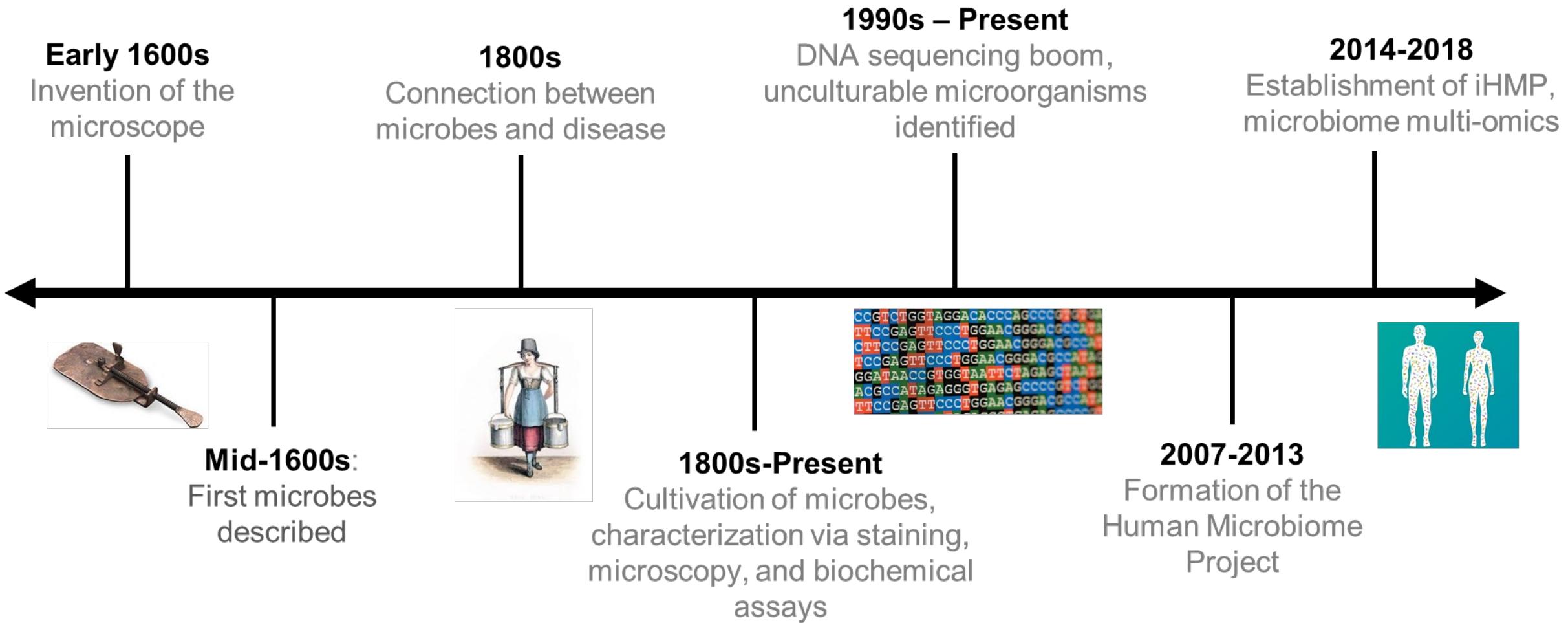
Undergrad Research Experience:

- Microbiology TA
- Anthropology Research
- Microbial Ecology Research

a alamy stock photo



Technological Advancements Drive Discovery



Adapted from University of Utah, 2015

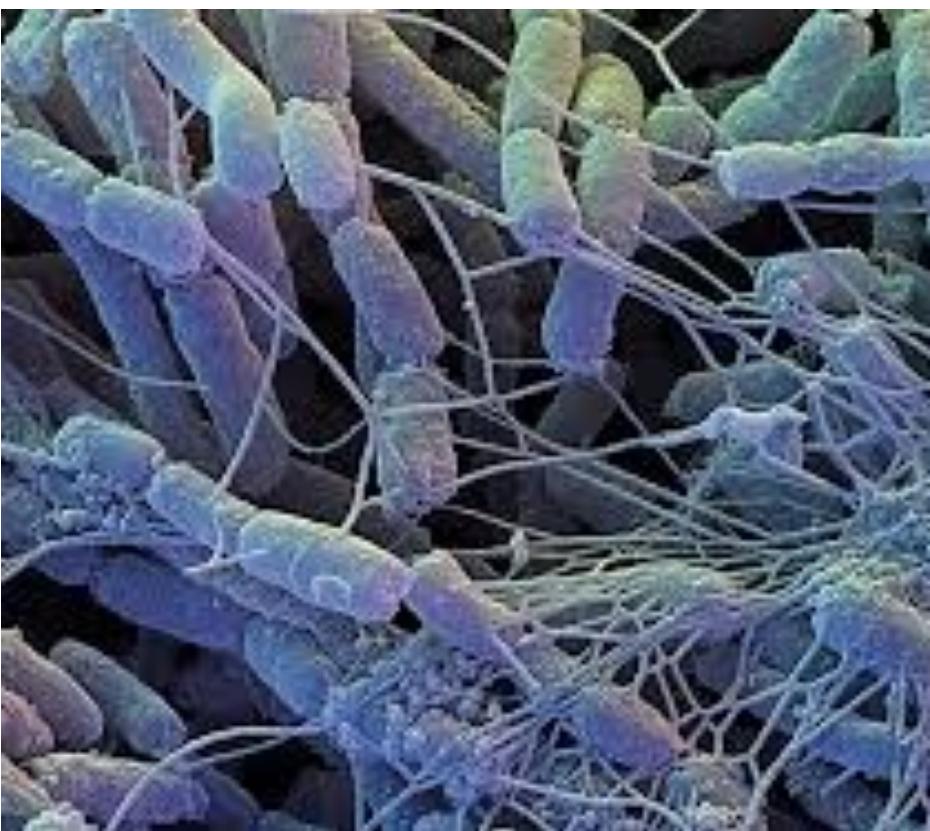
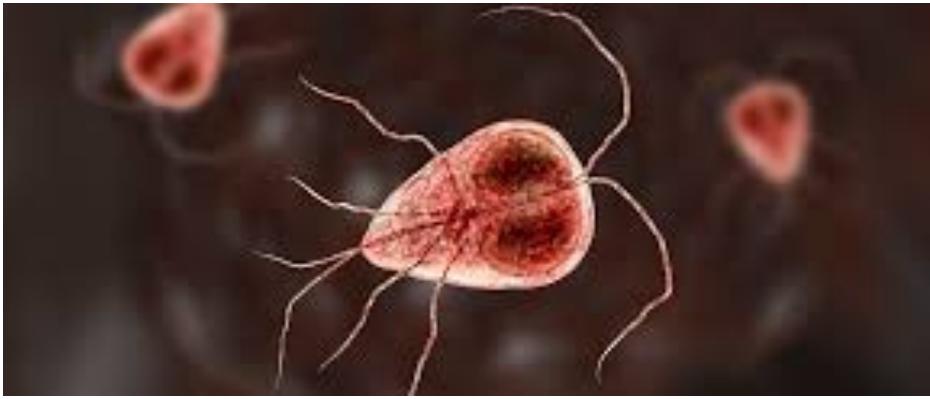


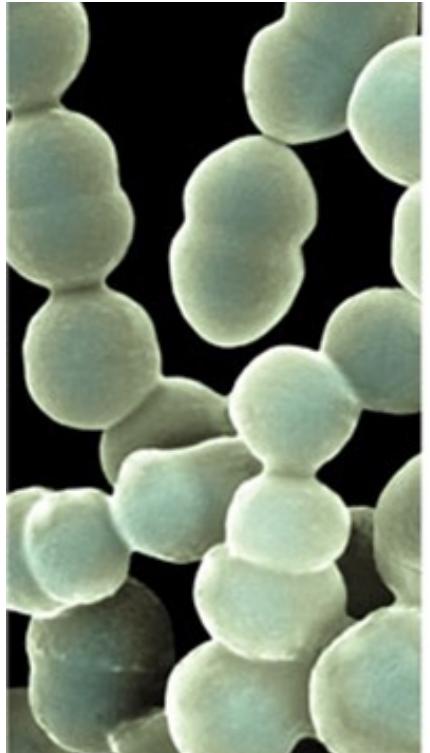
How did we get here looking at microbes on a large scale?

Curiosity of Differences



Classification
based on
physical traits





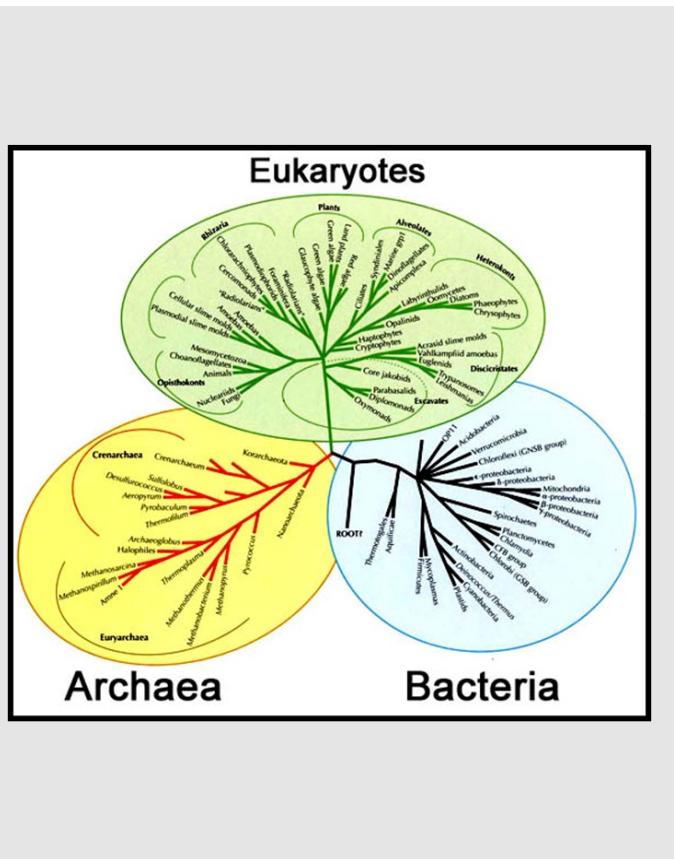
Classification
based on
morphology
doesn't work
well for
microorganisms

Is there a way to classify micro-organisms based on evolutionary
lineage?

Carl Woese & 16S/18S rRNA

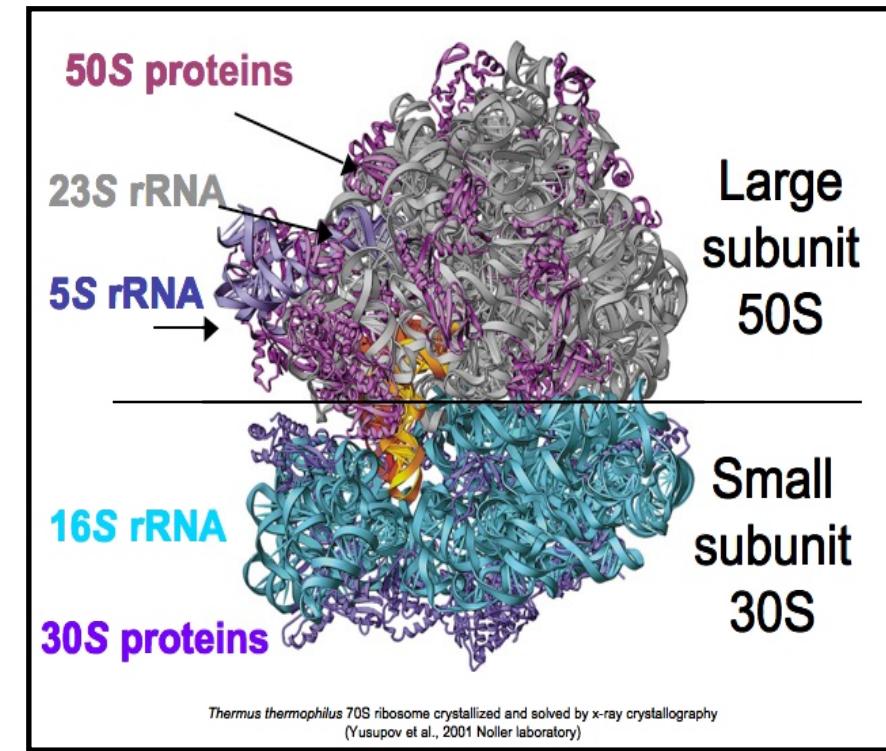
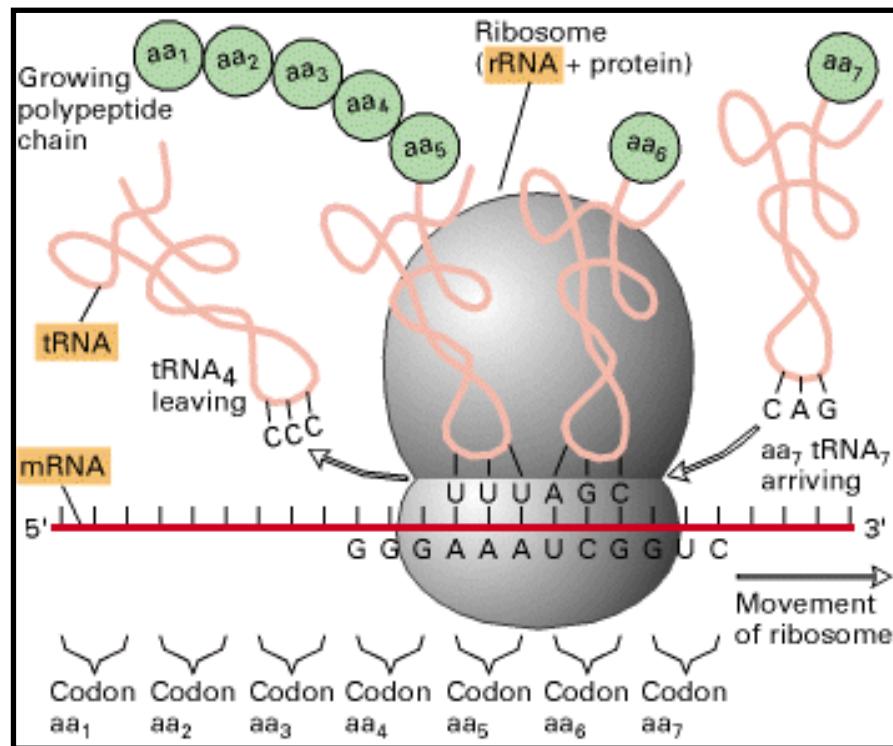


NARA/U. of Illinois 306-PS-E-77—5743



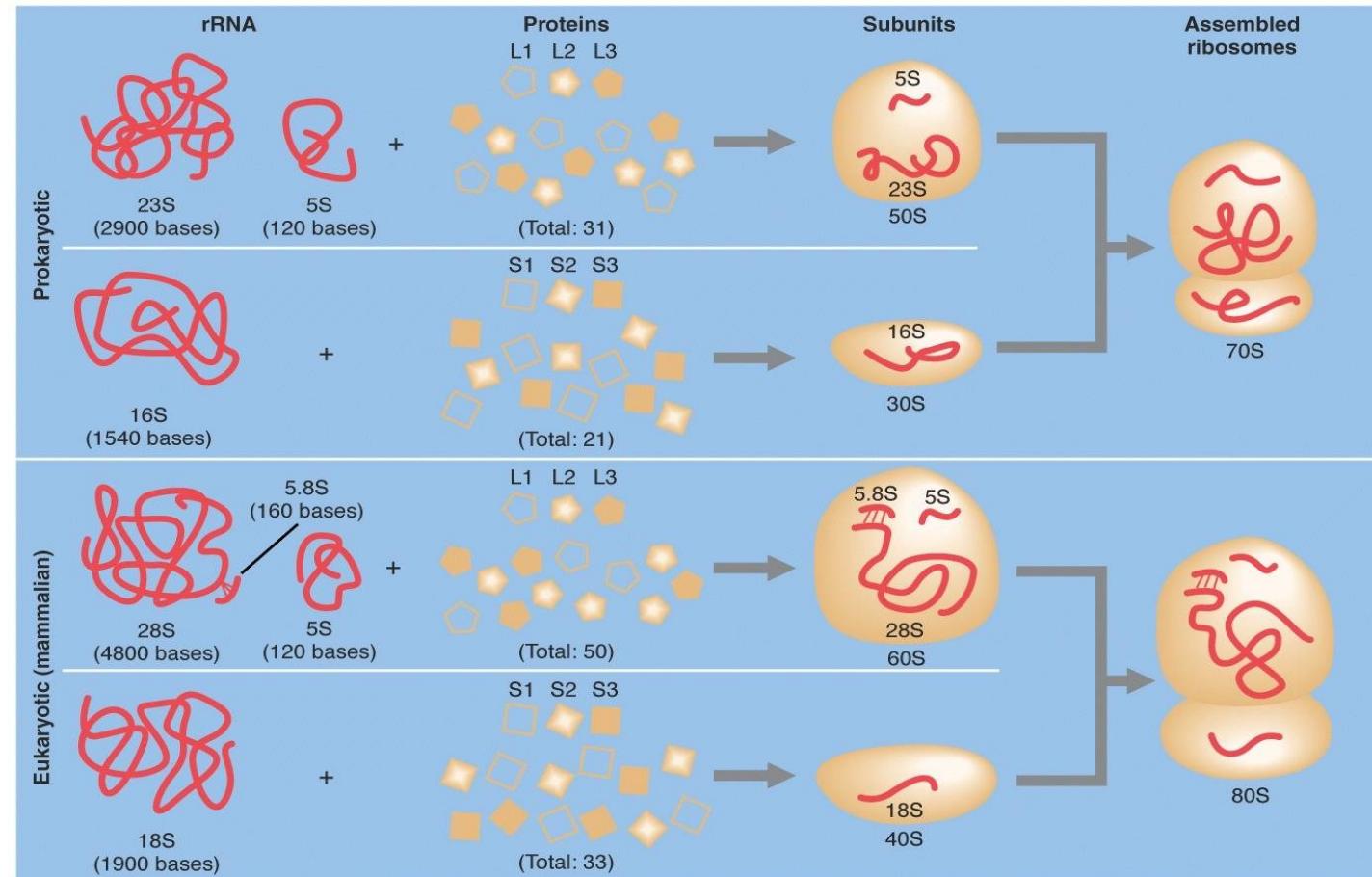
- “Archaea” (1977)
 - 3 domains
 - “King Philip Came Over For Good Soup”

16S rRNA is Part of the Ribosome

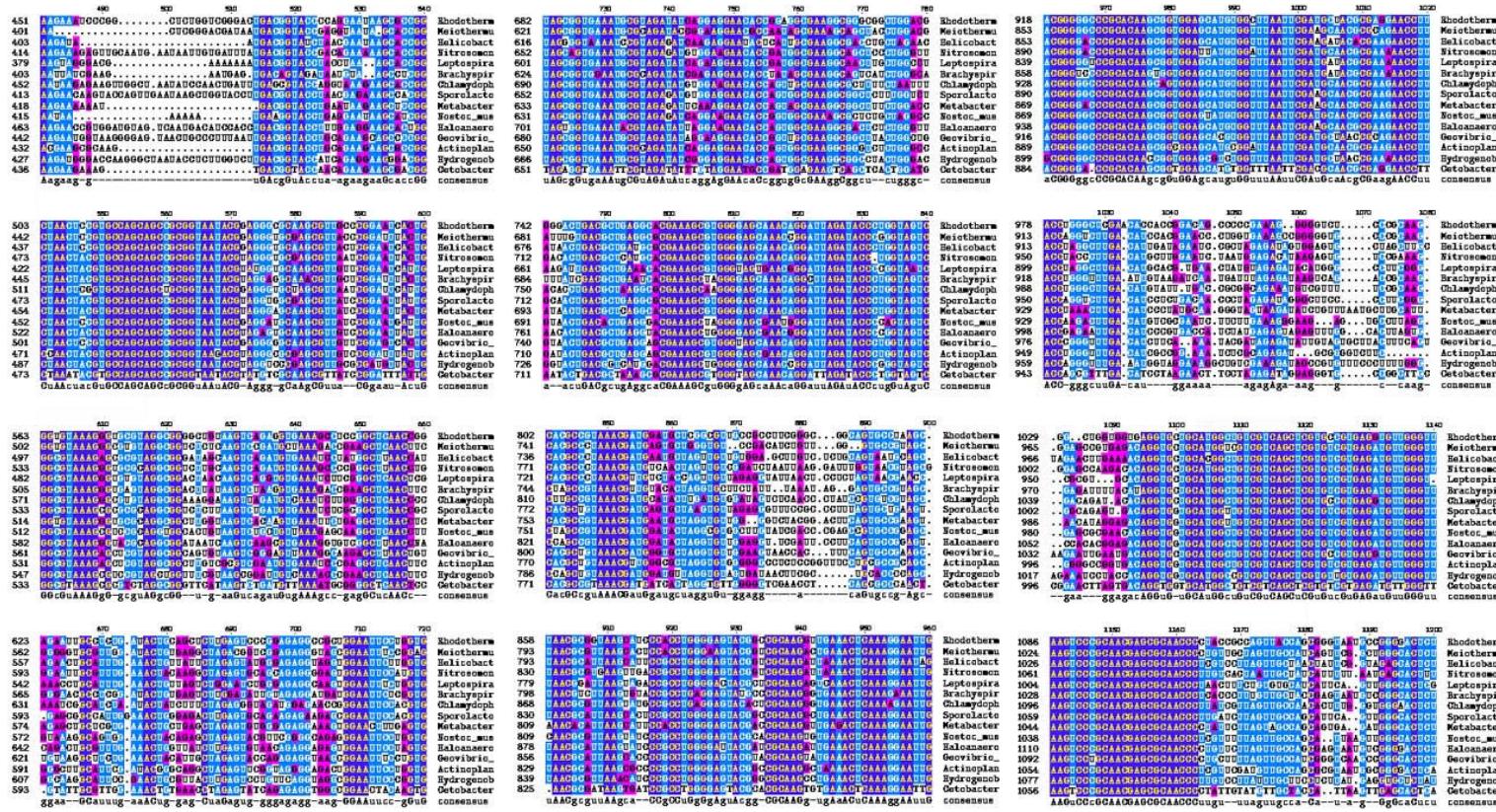


What makes 16S a good marker for taxonomy?

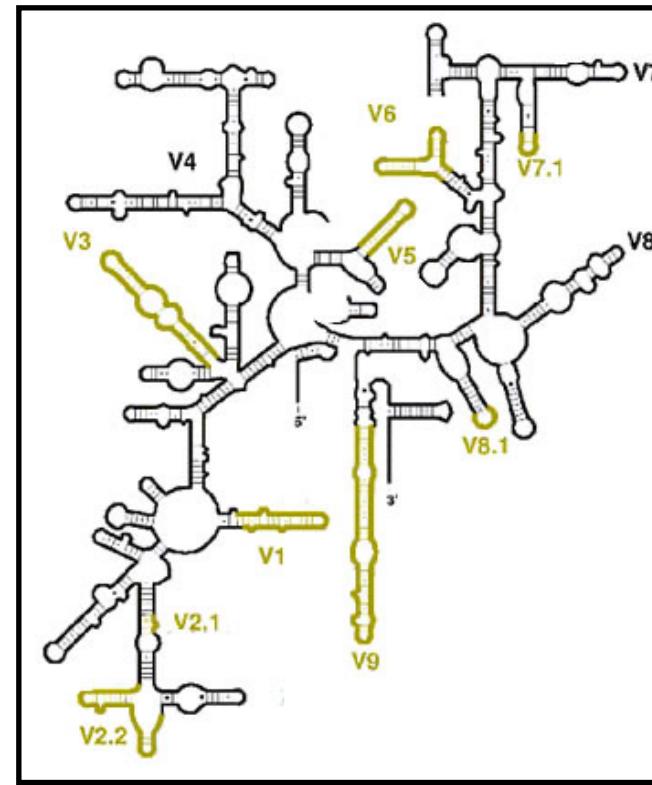
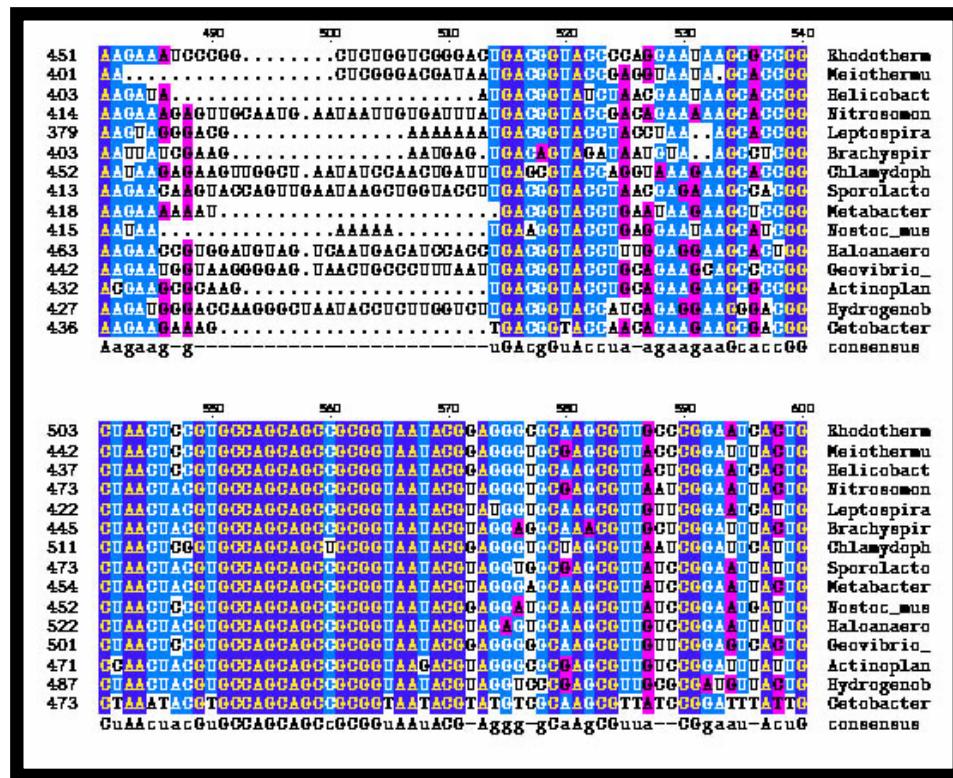
I. Ribosomes are Universal



II. 16S rRNA Gene is Highly Conserved



III. But, There are Variable Regions



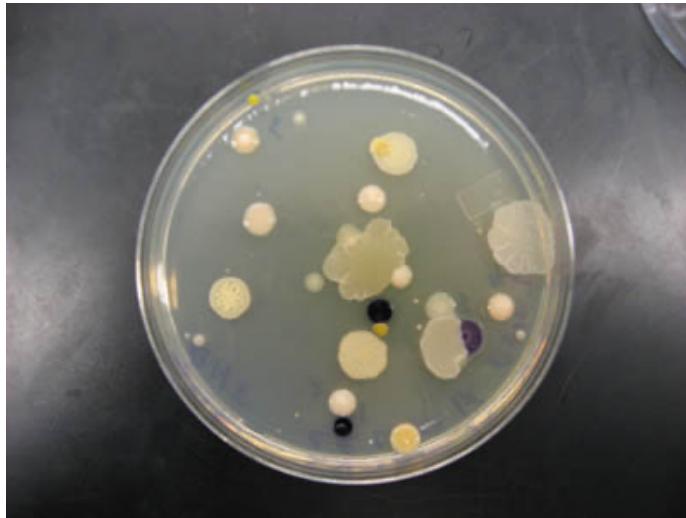
- Conserved residues important for structure and function

What Makes 16S rRNA a Good Marker for Phylogeny?

- 16S rRNA is present in all microorganisms (universal)
- 16S rRNA gene is highly conserved
- Mutations in variable regions can be used to trace evolutionary changes

Do you think there are other genes that are good markers for taxonomical classification?

The Microbial(Bacterial) Sequence Approach



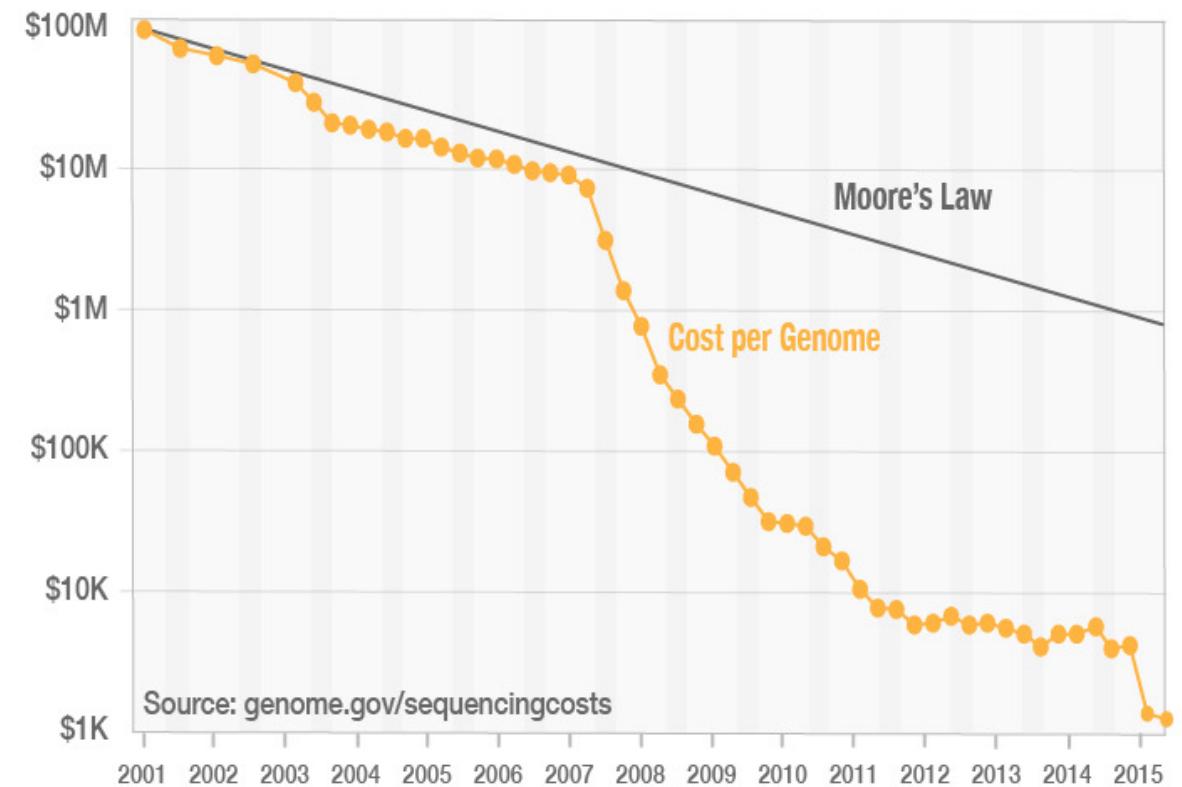
Only ~1% of microbes can be cultured using current techniques



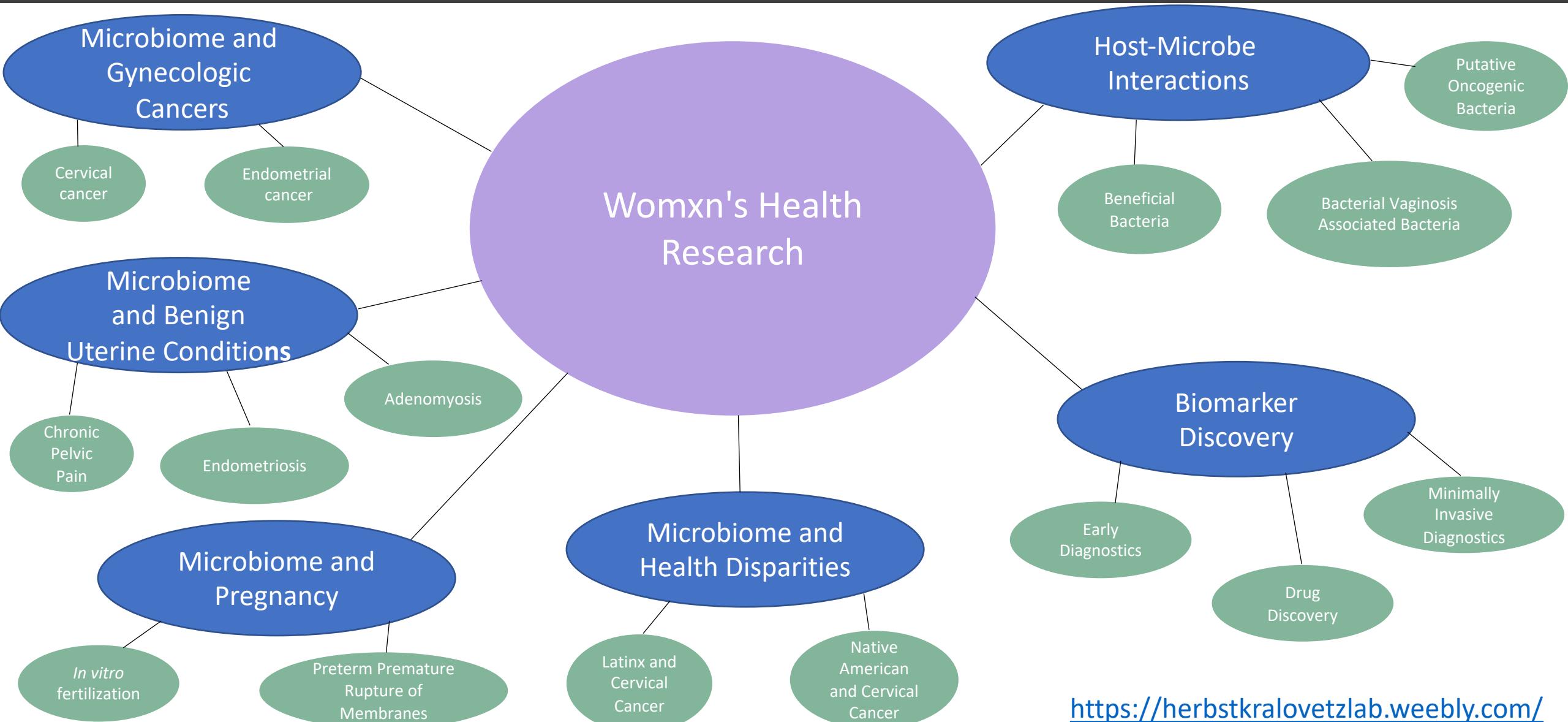
Use sequencing approach to identify DNA from microbes without culturing

Different Approaches to Investigating the Microbiome

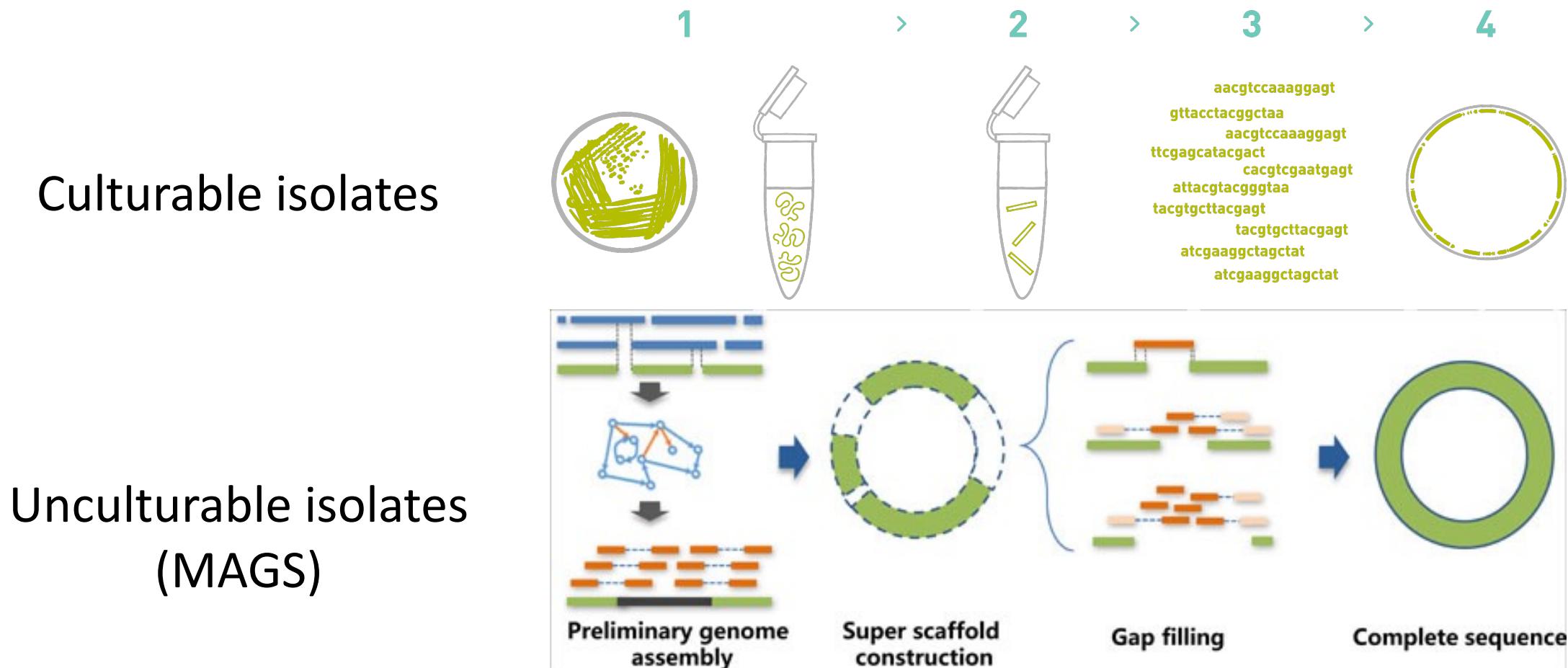
- Whole Genome Sequencing
 - Later developed Metagenomic assembled of Genomes
- Marker Gene Sequencing
- Whole Metagenomic Sequencing
- Your scientific question will determine your method
- Also cost!! -> Risk assessment



Herbst-Kralovetz lab and the vaginal microbiome



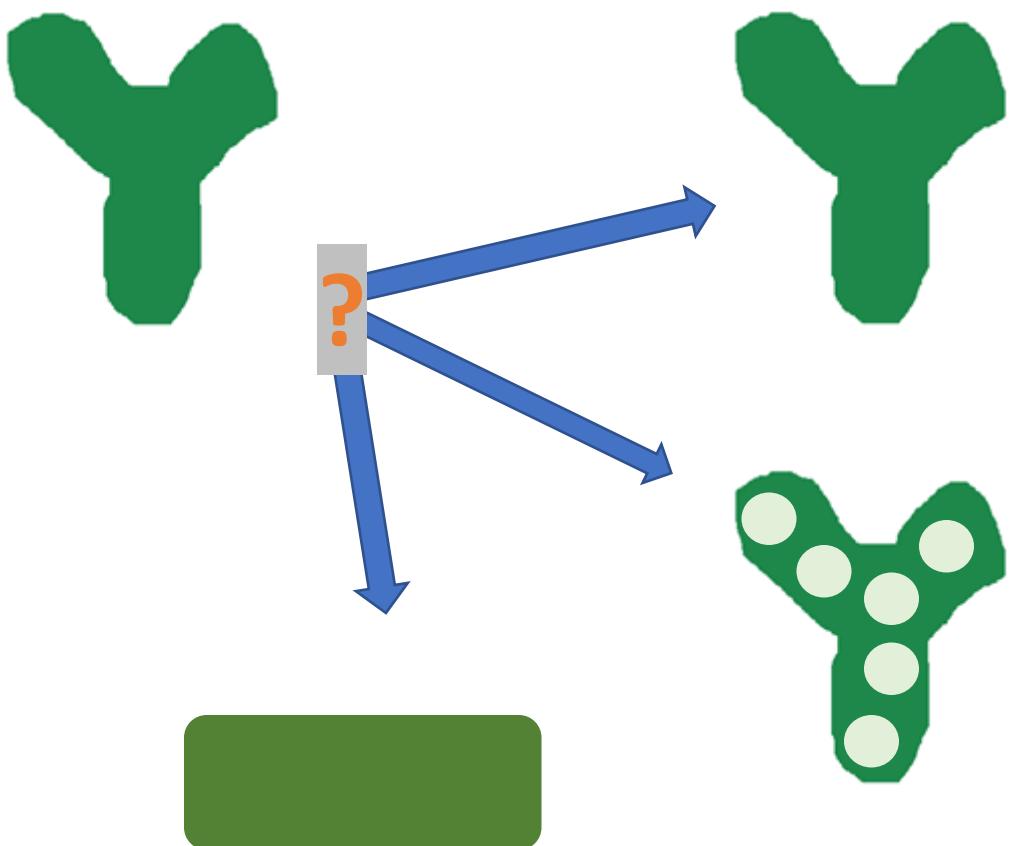
Whole Genome Sequencing



Comparative Genomics and Taxonomic classification

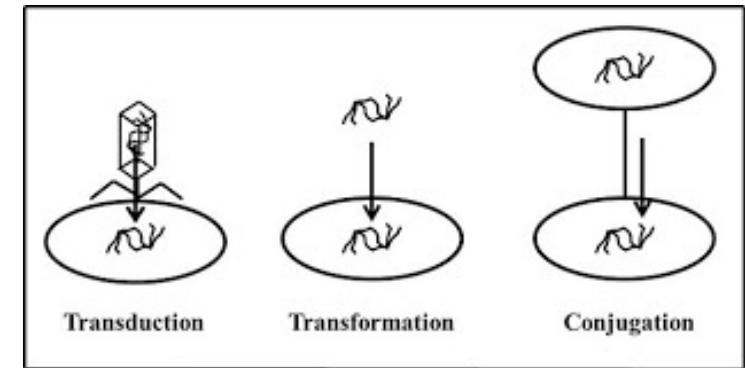
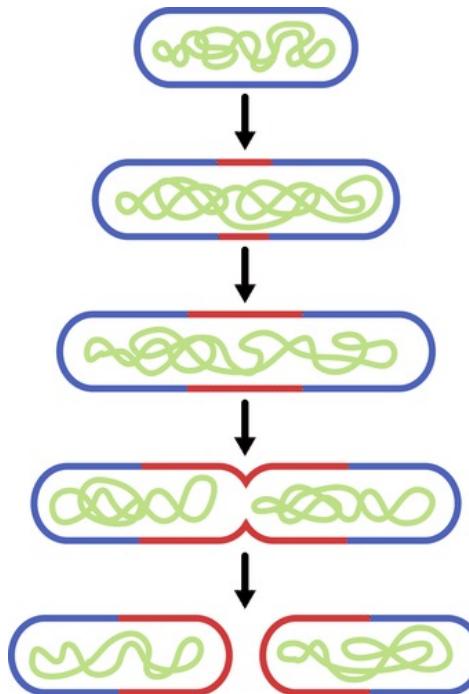


Predoctoral degree Nicole

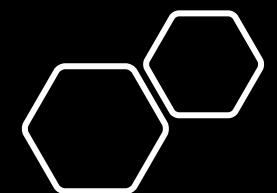


Main areas :

- Genome alignment
- Gene prediction
- Gene regulatory prediction
- Phylogeny
- Gene gain and loss events
- Evolution
 - Mutations
 - Selective pressures
 - Horizontal gene transfer

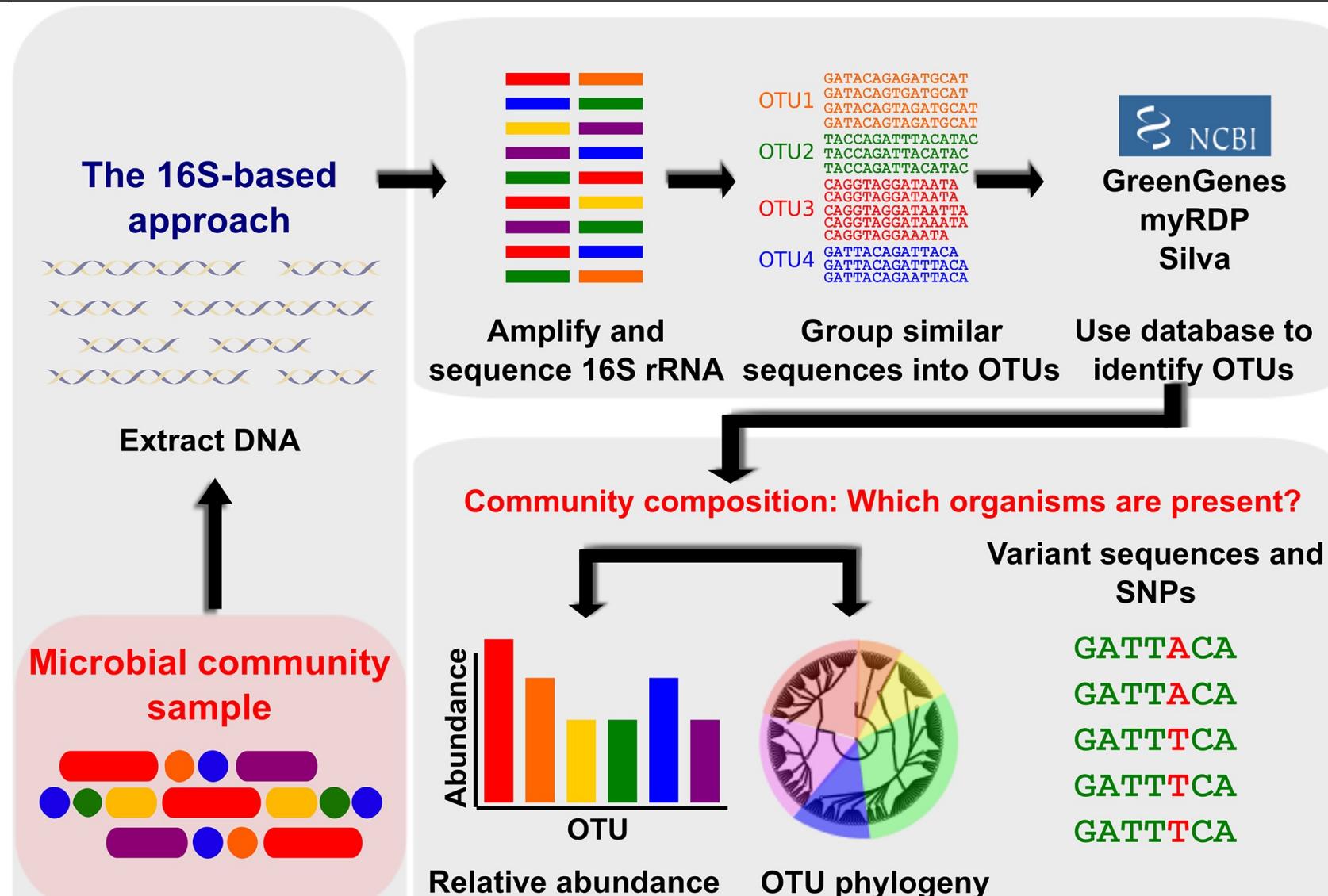


"Comparative genomics is a field of biological research in which researchers use a variety of tools to compare the complete genome sequences of different species. By carefully comparing characteristics that define various organisms, researchers can pinpoint regions of similarity and difference." - NIH



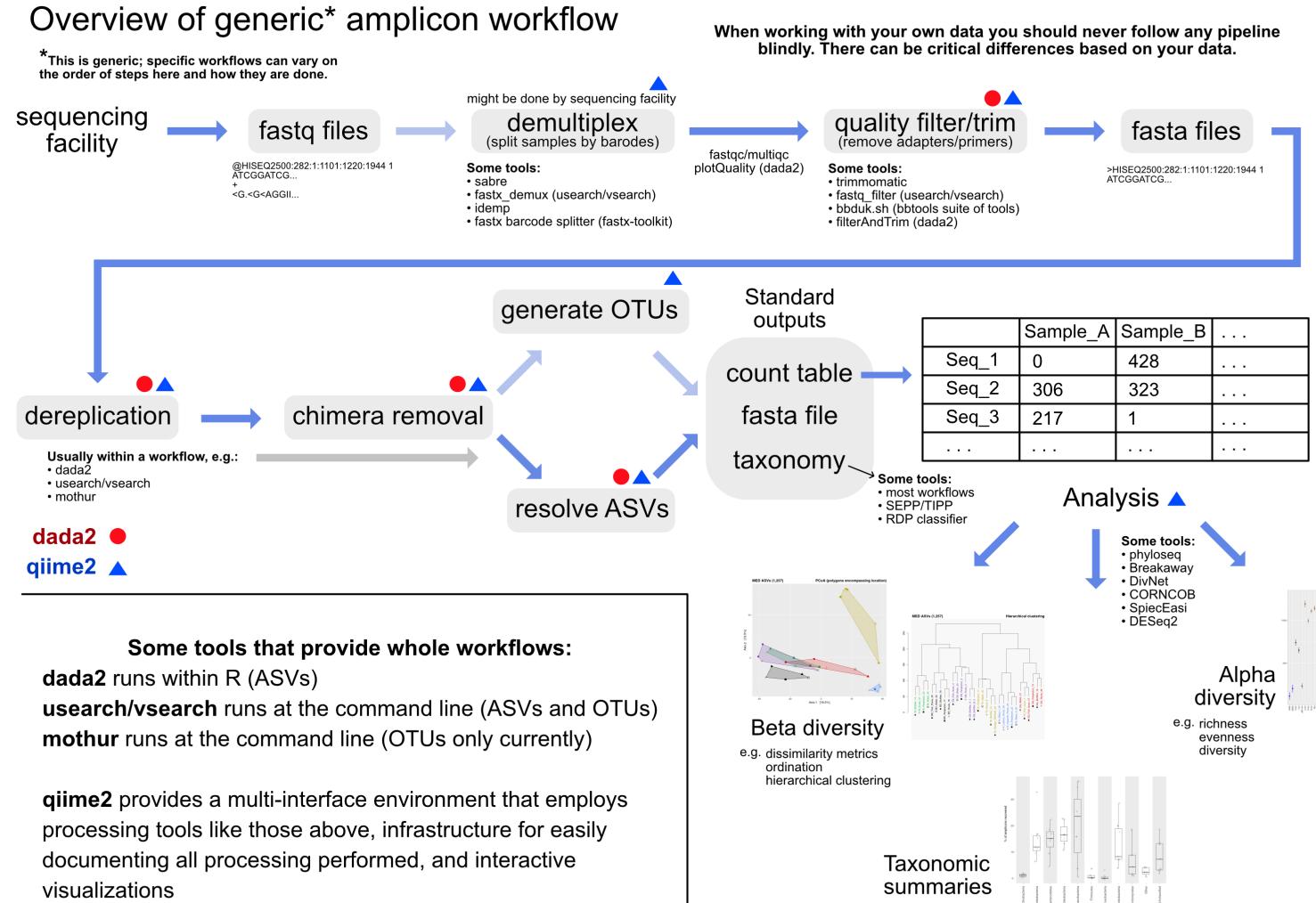
Marker Gene Microbiome Methods

Generally 16S rRNA, but others could be cpn60, GyrB, etc

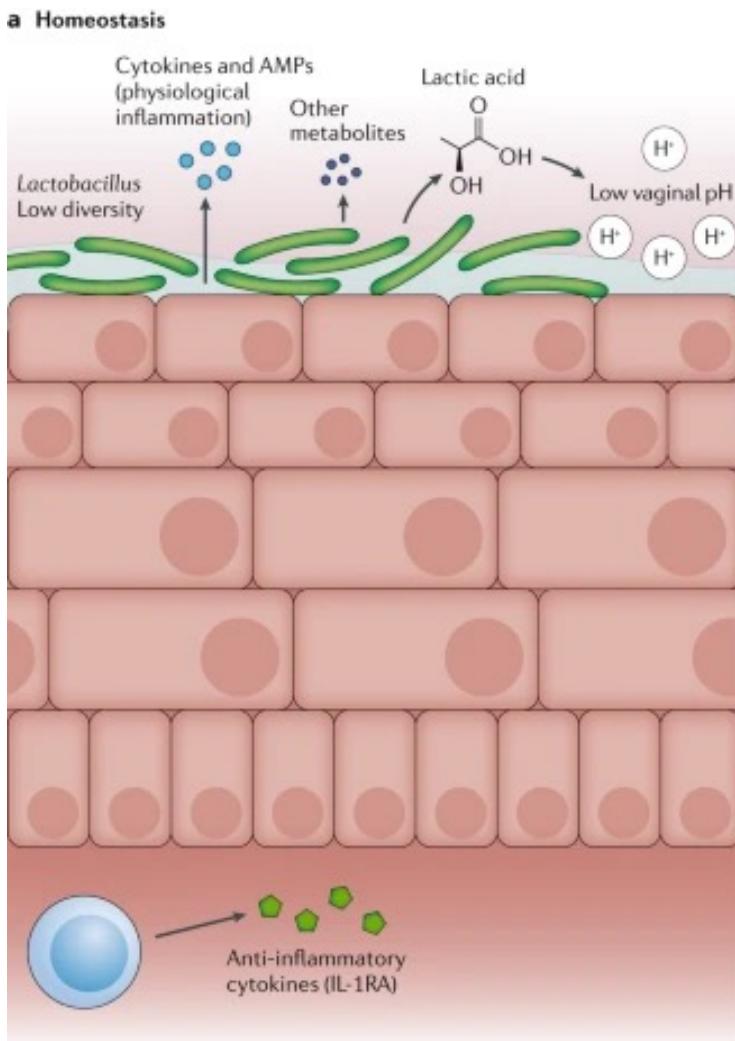


Marker Gene Microbiome Methods

Generally 16S
rRNA, but
others could be
cpn60, GyrB, etc



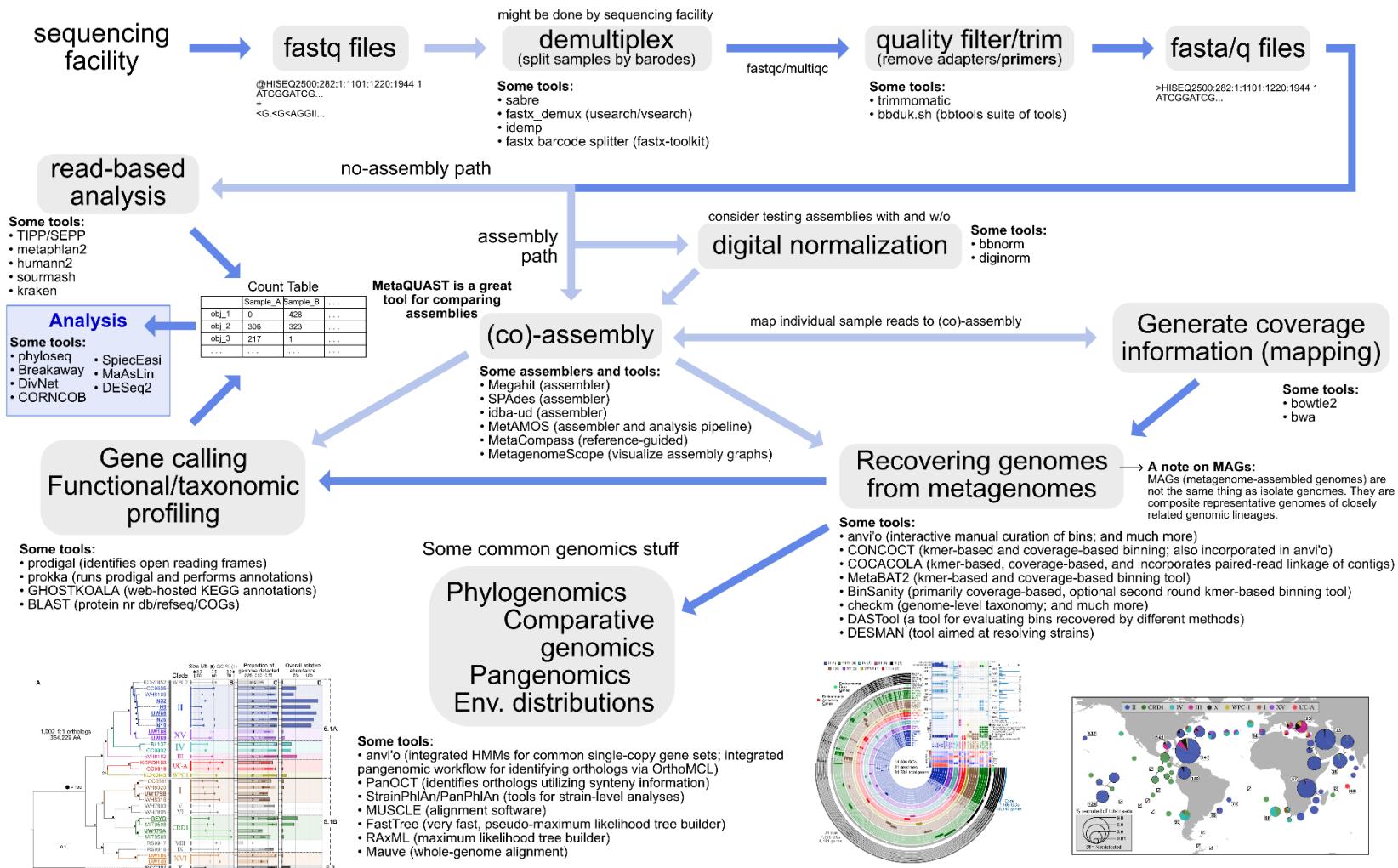
16S rRNA and omic data revealed Cervicovaginal microbiome paradigm



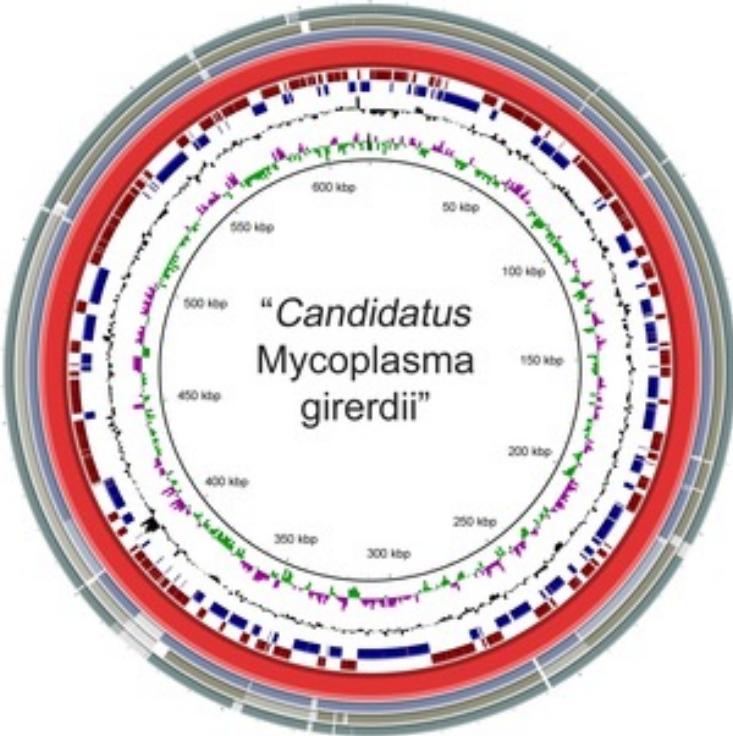
Whole Metagenomic Microbiome Methods

Overview of generic* metagenomics workflow

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



MAGs for Novel Microbe Discovery

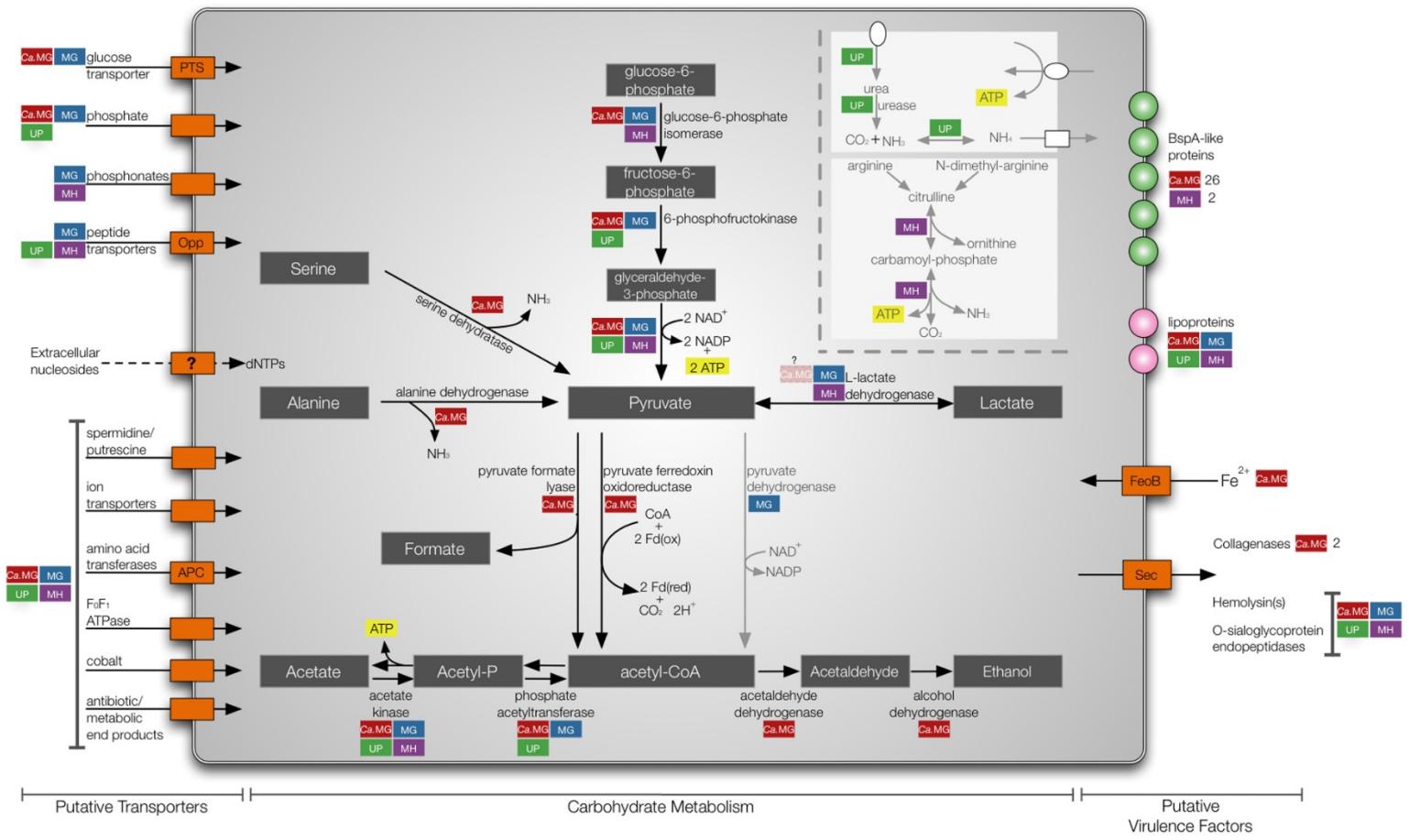


"Ca. M. girerdii" strain VCU_M1

	VCU_M1
Genome size (bp)	619 kb
DNA G+C content (%)	28.6%
rRNA operons	1
tRNA genes	31
Protein-coding genes	572
Genes assigned to COGs	415

Strains

- VCU_M1
- VCU_PA1
- VCU_JB1
- VCU_G1

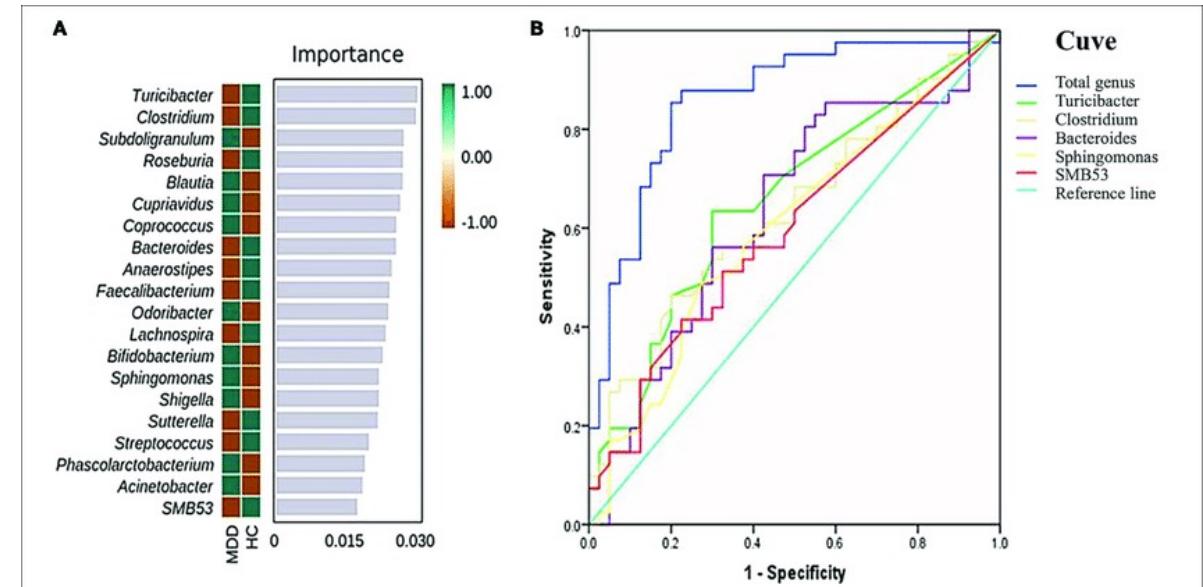


Fettweis et al, 2014

Types of Downstream Microbiome Analysis

- 16S/WMGS

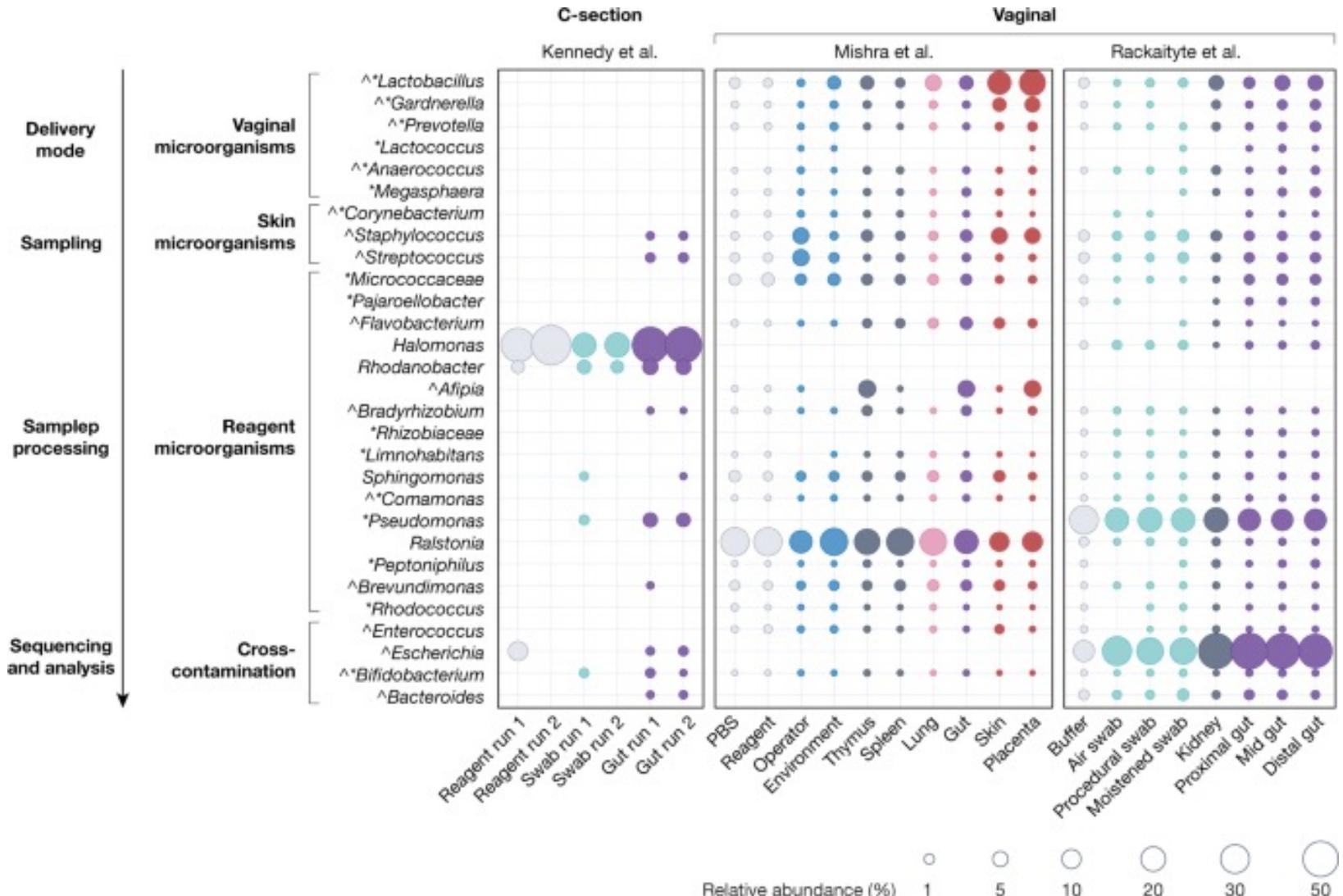
- Mosicki et al as an example:
<https://www.frontiersin.org/articles/10.3389/fcimb.2020.569022/full>
- Diversity
 - Alpha -> richness, evenness within samples
 - Beta -> diversity distance of group
- Taxonomic abundance
- Differential abundance
- Network analysis/co-occurrence
- Machine-learning for prediction on certain variables(Biomarker Discovery)
- Putative metabolic contribution
- Microbial evolution - WMGS
- Host-Microbe Interactions (Genes identity) - WMGS



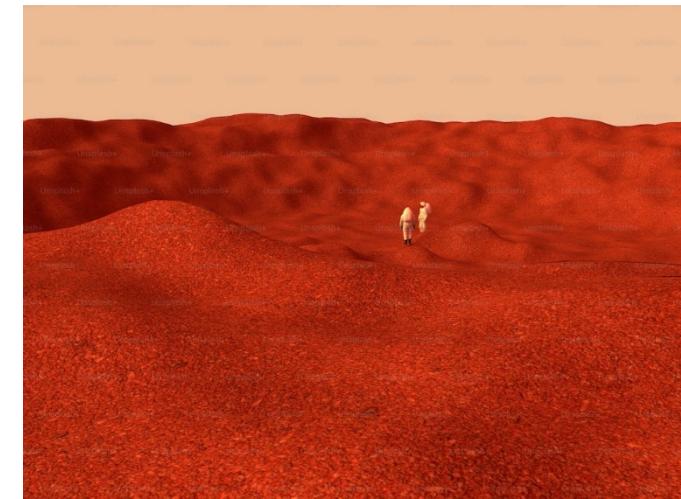
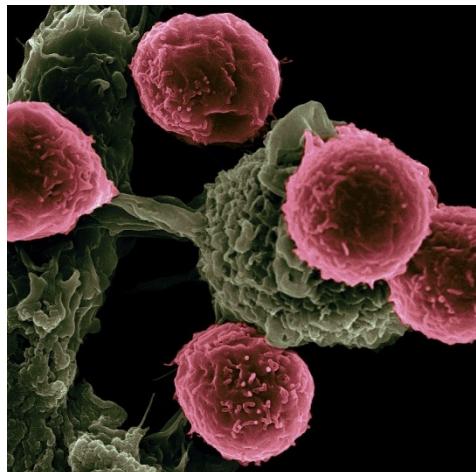
Liu et al, 2022

Low-Biomass Samples and Controls

Study Design and Upstream prep Matters!



How can Microbiome Fit into your Research



Possibilities Endless



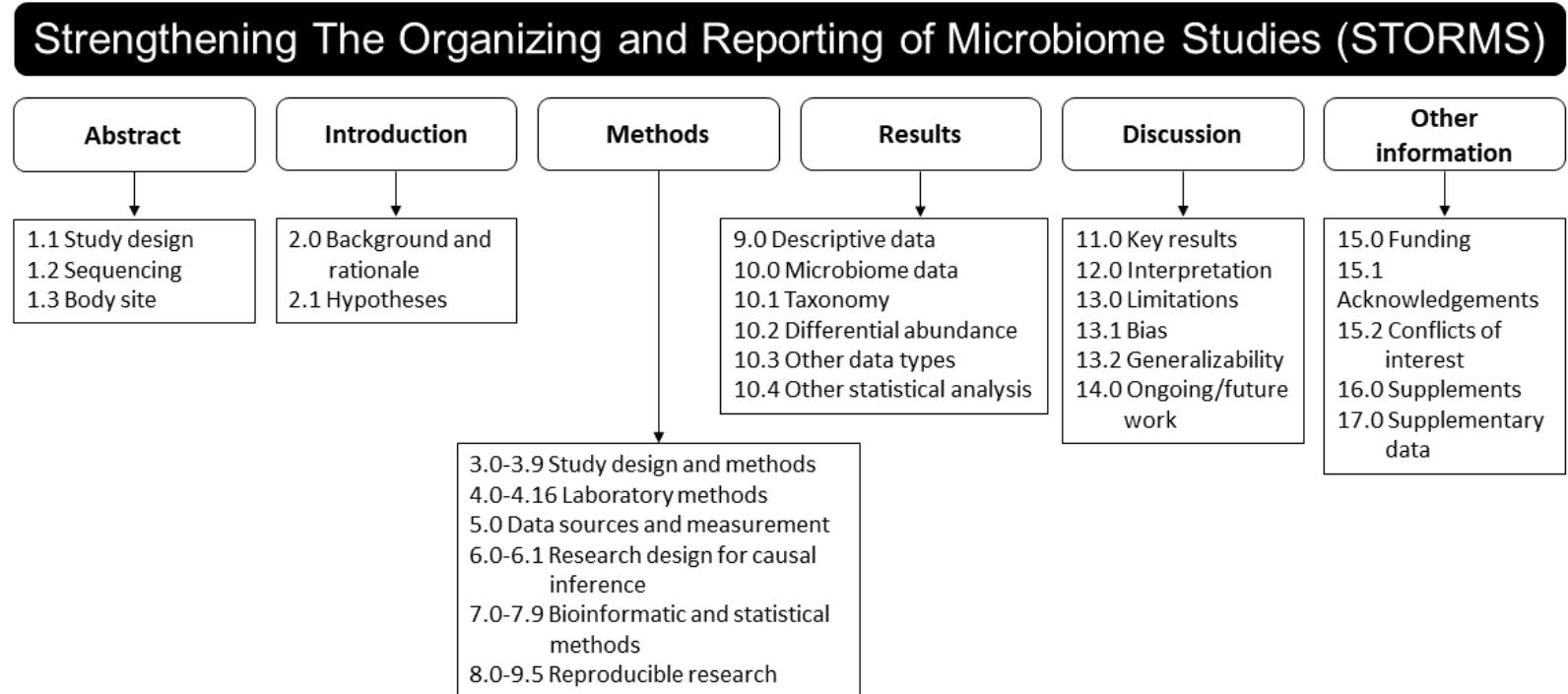
Also....

- Protozoans
- Fungi
- Archaea
- Viruses

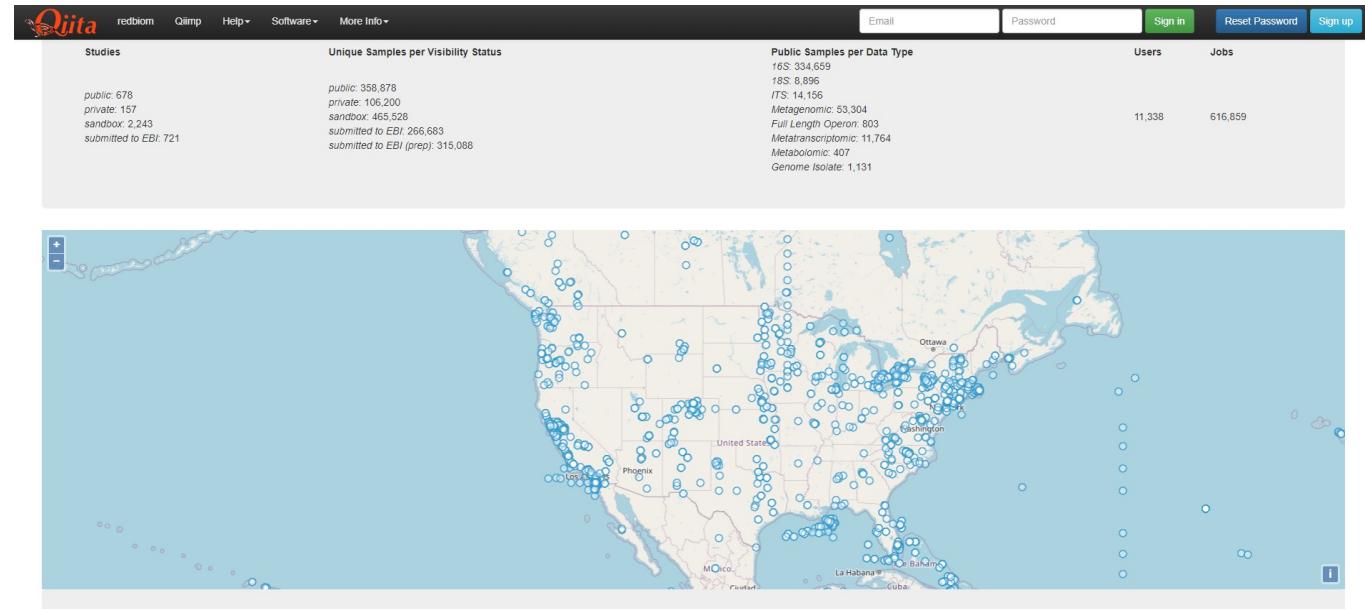
These require more research in the microbiome field and different considerations

Standardization

Strengthening The Organization and Reporting of Microbiome Studies (STORMS) is a checklist for reporting on human microbiome studies, organized into six sections covering all sections of a scientific publication, presented as a table with space for author-provided comments and intended for inclusion in supplementary materials.



Publicly Available Data



Qiita (canonically pronounced cheetah) is a software package intended for analysis and administration of multi-omics datasets.

Qiita provides a free and open platform for users to:

1. Easily share and reuse existing data-sets in the form of studies.
2. Perform analyses (by combining one or more studies/data-sets) using published and unpublished data that meet the standards as described by this documentation.
3. Easily interface with the EBI repository for automated deposition. Query and interact with Qiita data programmatically.

Deposition of Data

The screenshot shows the dbGaP homepage with a search bar at the top. Below it is a banner with a blue eye graphic and text about the database's purpose. The main content area has sections for 'Access dbGaP Data' (Advanced Search, Controlled Access Data, Public ETR Download, Collections, Summary Statistics), 'Resources' (dbGaP Data Browser, Phenotype-Genotype Integrator, dbGaP RSS Feed, Software), and 'Important Links' (How to Submit, FAQ, Code of Conduct, Security Procedures, Contact Us). A 'Latest Studies' table lists several studies with columns for Study ID, Embargo Release, Details, Participants, Type Of Study, Links, and Platform. The table includes rows for studies like 'Gabriella Miller Kids First Pediatric Research Project in Microtia in Hispanic Population', 'Kids First Genetics of Kidney and Ureter Tract Malformations', and 'Gabriella Miller Kids First Pediatric Research Program in Craniofacial Microsomia'.

Sequencing data

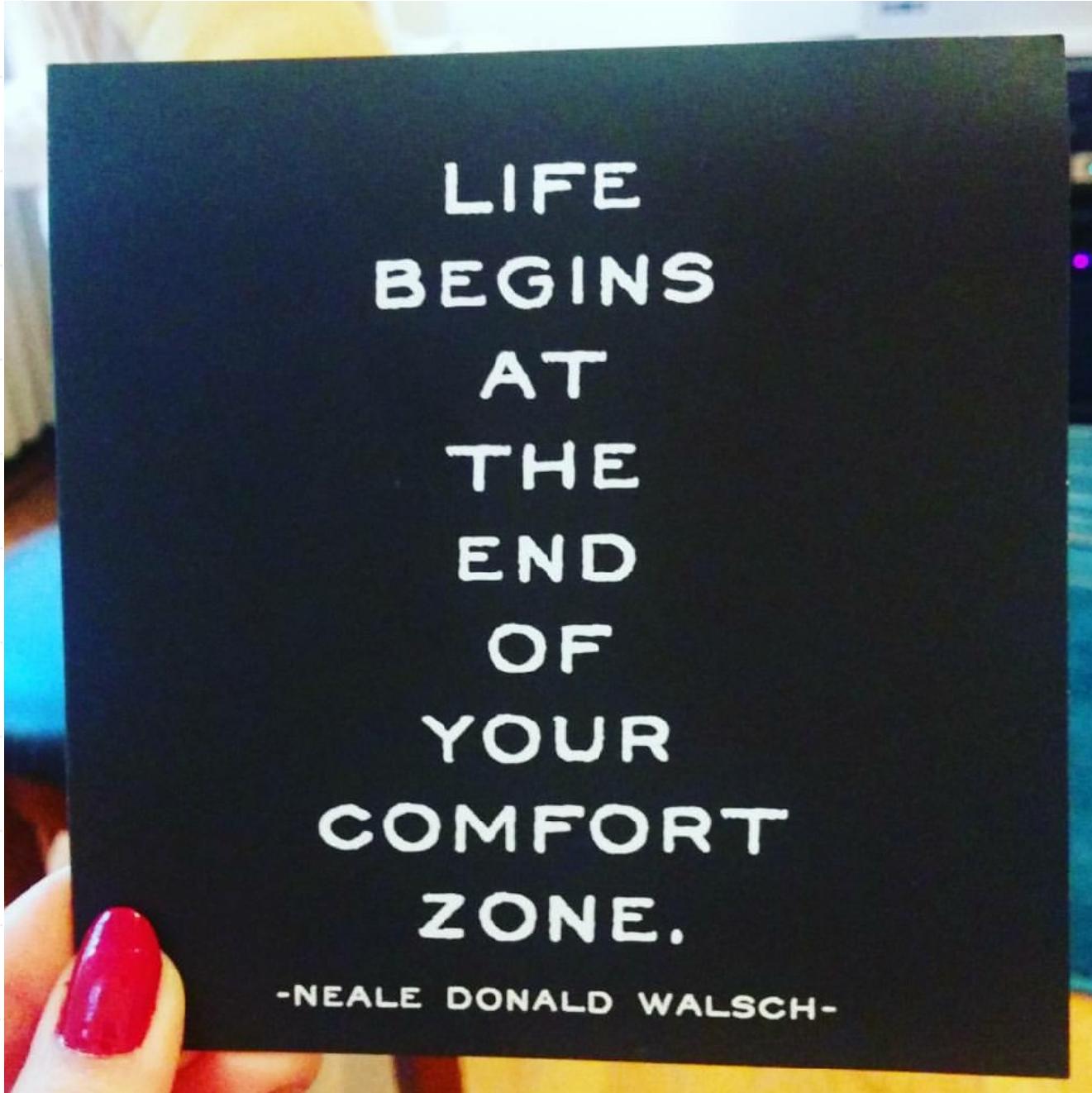
Deposition of data is important and considered in your data management plan whether working with sequence data or samples these are places to be considered.

Metadata is key in this step as well, go through a database and you may find yourself with a headache...

The screenshot shows the BEI Resources product page for HM-30 *Bifidobacterium* sp., Strain 12_1_47BFAA. The page includes sections for 'ATTACHMENTS' (Product Information Sheet, Certificate of Analysis by Lot), 'PERMITS' (Shipping to the U.S.: State of Hawaii), and 'RELATED PRODUCTS' (HM-300, Genomic DNA from *Bifidobacterium* sp., Strain 12_1_47BFAA). The main product details section includes fields for Organism, Designation, Biosafety Level, Availability Status, Store at, Contributor, Comments, and Citations. The 'Comments' section notes a quantity limit of 1 per year and shipping information for the U.S. State of Hawaii.

Bacterial Cultures

<https://www.impactt-microbiome.ca/platform-3-microbial-human-tissue-repositories/>



Card from my Tia when I got into Graduate School

Applicable to all stages of a (Data) Scientists career

Acknowledgments

- **Herbst-Kralovetz Current Lab Members**

- Dr. Melissa Herbst-Kralovetz
- Dr. Pawel Laniewski
- Phoebe Crossley
- Hajar Rahee
- Cole Sherwood
- Emerald Bell
- Vianney Mancilla
- Michelle Aminov
- Clinical Research Team:
 - Dr. Dana Chase, BUMC-P Clinicians, and Erika Flores

- **College of Medicine-Phoenix - Department of Obstetrics and Gynecology**

- **Participants from our Vaginal Cohort Studies**

- UA Cervical Cancer Study & UA Endometrial Cancer Study

- **Early Investigator Support**

- Postdoctoral Affairs Office, UA-CoM-P Research Office, and Native American Cancer Prevention – GUIDeS: U54CA143924, U54CA143925



- **Funding**

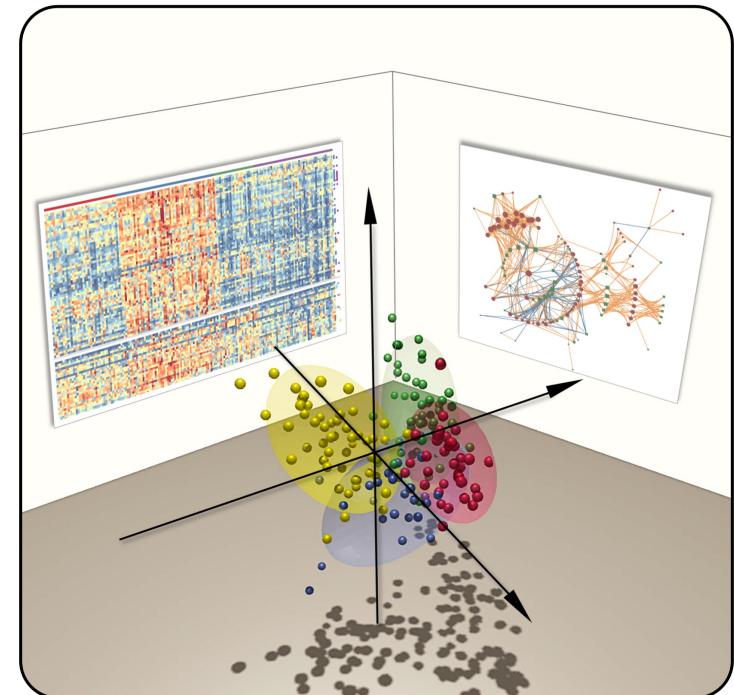
- Office of Research for Women's Health (3P30CA023074-39S3)
- Community Foundation for Southern Arizona
- Flinn Foundation
- Mary Kay Foundation
- Data Science Institute – Data Science Fellows
- BIO5 Institute – Postdoctoral Fellowship

Resources for future Microbiome learning

- Qiime2
 - Youtube: <https://www.youtube.com/@QIIME2>
- Qiime2 Galaxy
 - <https://forum.qiime2.org/t/qiime-2-available-on-public-galaxy-server/26220>
- STAMPS Workshop
 - <https://www.mbl.edu/education/advanced-research-training-courses/course-offerings/strategies-and-techniques-analyzing-microbial-population-structures>
- PANDA Core For Genomics and Microbiome Research is offering a new iCourse "Introduction to Microbiome Community Analysis Methods"
 - <https://international.uahs.arizona.edu/news/2022/11/summer-icourse-spotlight-introduction-microbiome-community-analysis-methods>

GUI(R packages) for Omic Analysis

- <https://www.microbiomeanalyst.ca/MicrobiomeAnalyst/home.xhtml>
 - <https://github.com/jimeneznr/MicrobiomeAnalystR>
- <https://www.metaboanalyst.ca/>
- <https://www.omicsanalyst.ca/>
- <https://metorigin.met-bioinformatics.cn/home/>
- <https://borenstein-lab.github.io/MIMOSA2shiny/>



OmicsAnalyst

Exercise

- Microbiome field is great as many folks usally have as thorough tutorials
 - Here are some I follow for resources:
 - <https://github.com/topics/microbiome-workflow>
- This one I use with folks in my lab to get their toes wet
 - Thanks to Nabiilah Ardini Fauziyyah for this tutorial
 - <https://rpubs.com/nabiilahardini/microbiome>
- If you come across an interesting package in Microbiome space there are many, I always keep my ear to the ground to see what is upcoming.
- Feel free to Reach out to developers via Github, I often encounter that the developer normally has some sort of forum if largely used package or will adjust the current tool to fit the users needs or give advice on work arounds
 - MIMOSA2 and VOLARE developers have been very friendly and excited that our team is utilzing their tools.