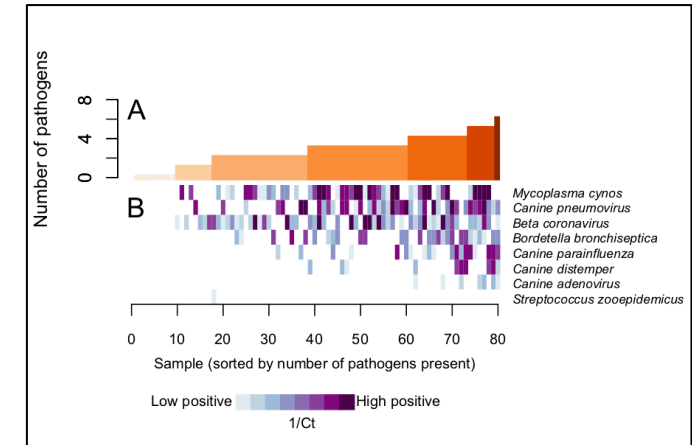
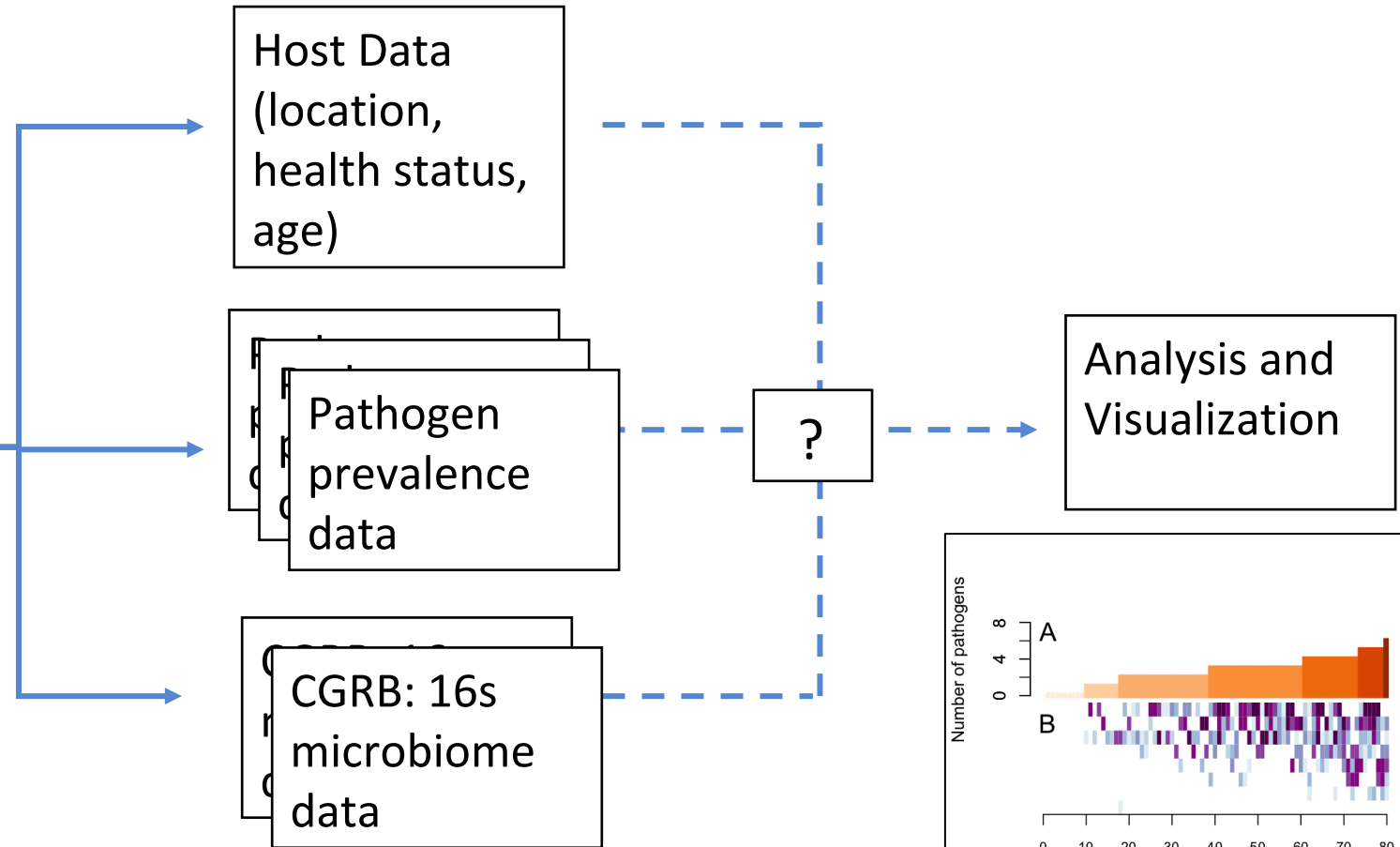


A detailed scanning electron micrograph (SEM) of a microbial community, likely a biofilm. The image shows a dense network of grey, fibrous extracellular polymeric substances (EPS) that form a complex, interconnected mesh. Scattered throughout this matrix are numerous individual bacterial cells. These cells are represented by various colors: bright green, vibrant yellow, deep purple, and soft pink. They exhibit different shapes, including rod-like (bacilli) and oval (cocci) forms, suggesting a diverse population of microorganisms. The overall texture is highly irregular and three-dimensional, typical of a natural microbial ecosystem.

# Managing and integrating multiple pathogen and microbiome datasets

Liz Knorr

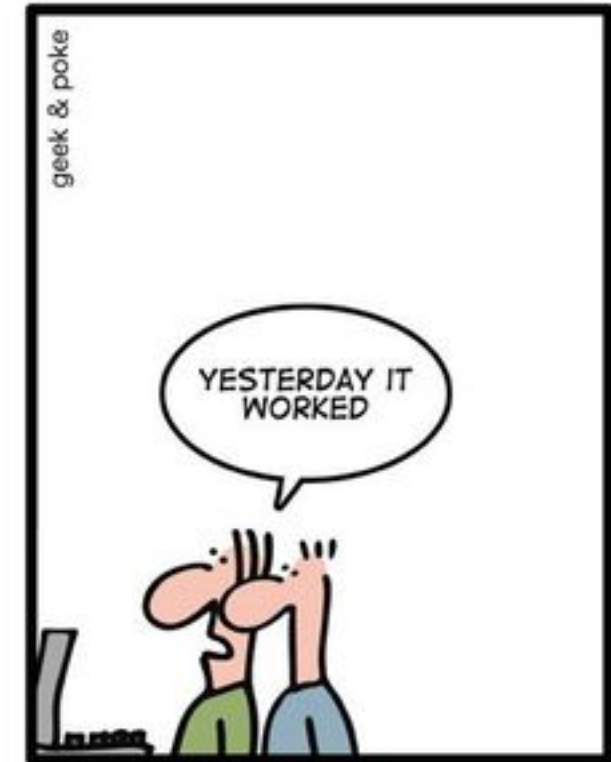
# Background



# Problem

- Multiple datasets I need to combine and be able to analyze in different ways
  1. Sample collection
  2. Respiratory pathogen diagnostic panels for each sample
  3. 16s microbiome data for each sample
  4. Demographic data for each sample
- Experienced trouble identifying and matching samples; no checkpoints, hardcoding
- Multiple copies of raw data

More data collection coming soon!



# Potential Approaches

- Best-practices for data management
  1. Data repository
  2. Combine data sets
  3. Clean data sets
  4. Reproducibility checkpoints

# End Result

- Reproducible workflow that allows me to integrate, manipulate, and analyze datasets
  - Pathogen data analysis
  - Microbiome data analysis
  - Host/demographic data
- Workflow that I can use for future pathogen and microbial abundance data