**AMD Southeast Region Genomic Epidemiology Training**
**Retrospective Genomic Data Analysis**
4/15/24

# Retrospective Genomic Data Analysis
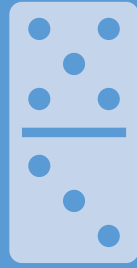
**Common Uses**
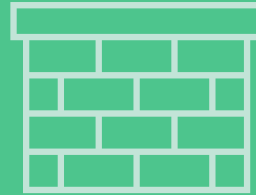
**Limitations**

**Forecasting**

# Literature Review on Genomic Pathogen Evolution

## SARS-CoV-2
- 6 articles
- Geographic Specificity
- Genetic Variation

## Bacteria Evolutionary Mechanisms
- 7 articles
- Evolutionary mechanisms of AR
- Genetic diversity and pathogenicity of endemic strains

## Eukaryotic Pathogens
- 7 articles
- 4 articles on Malaria
- 1 on zoonotic transmission of Crypto

## Host-Pathogen Interactions
- 1 article
- Development of panel for prediction of cervical cancer severity
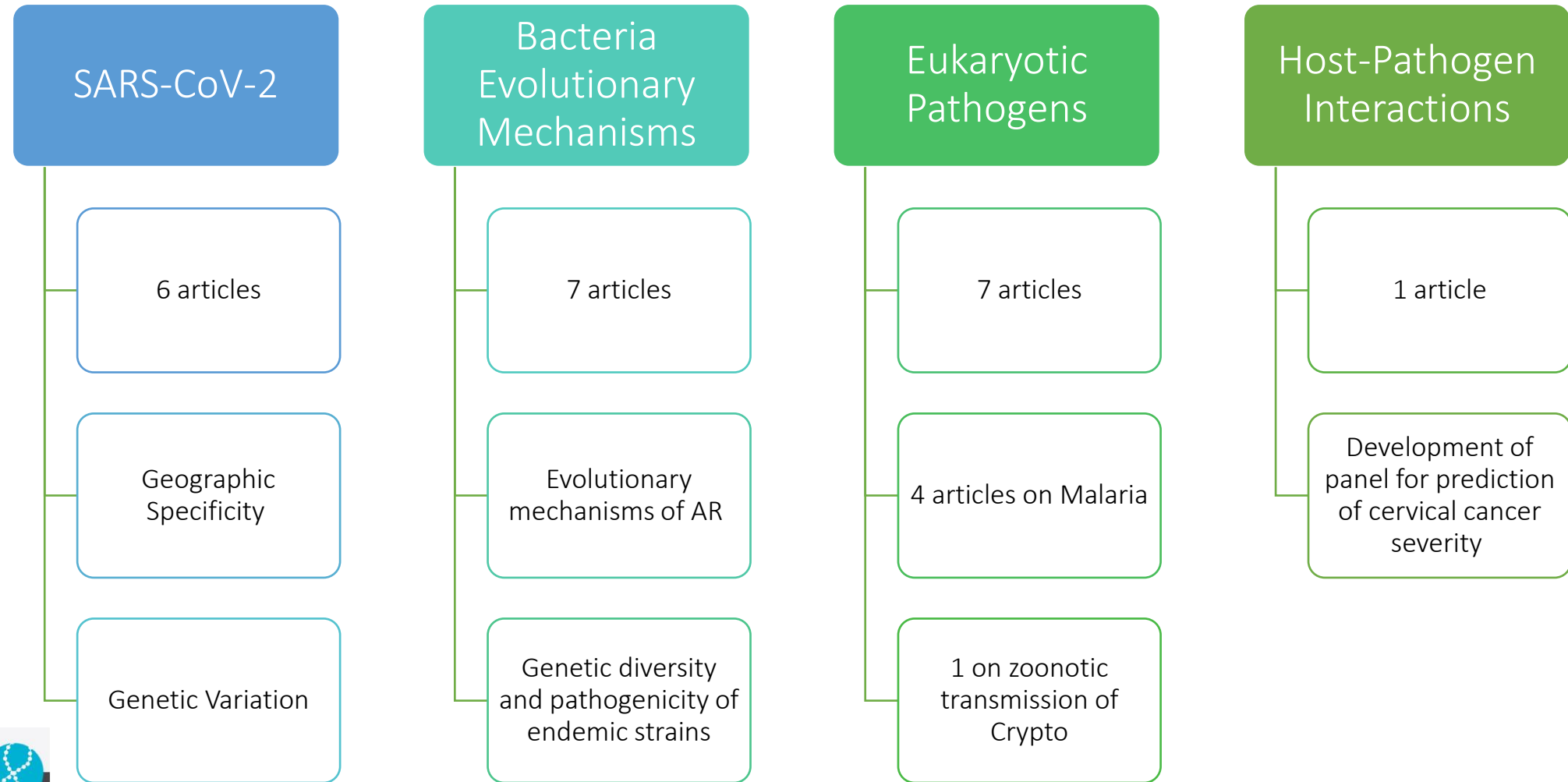
Advanced Molecular Detection
Southeast Region Bioinformatics

Jianying Gu et al. Editorial: Evolutionary Mechanisms of Infectious Diseases. Frontiers of Microbiology 2021.

# Common Reasons for Retrospective Data Analysis

Evolutionary Mechanisms

Spillover event identification

Investigate past outbreaks

Identify past trends at various levels

Inform policy making

Advanced Molecular Detection
Southeast Region Bioinformatics

# Pathogen Evolution

- Factors that influence pathogen virulence and host susceptibility
  - Trade-off between transmission rate and mortality rate of host
- Genetic variation and population dynamics on pathogenesis
  - Acquisition of genetic material from different pathogens
- Mechanisms of Antibiotic Resistance
  - Role of medical interventions on drug resistance
- Role of disease control interventions on pathogens



**The evolution of antagonistic coevolutionary relationships**

Coevolution in emerging or relatively new infectious disease systems is likely to be characterized by quantitative resistance, whereas infectious disease systems with a long coevolutionary history are likely to be characterized by the evolution of qualitative resistance.

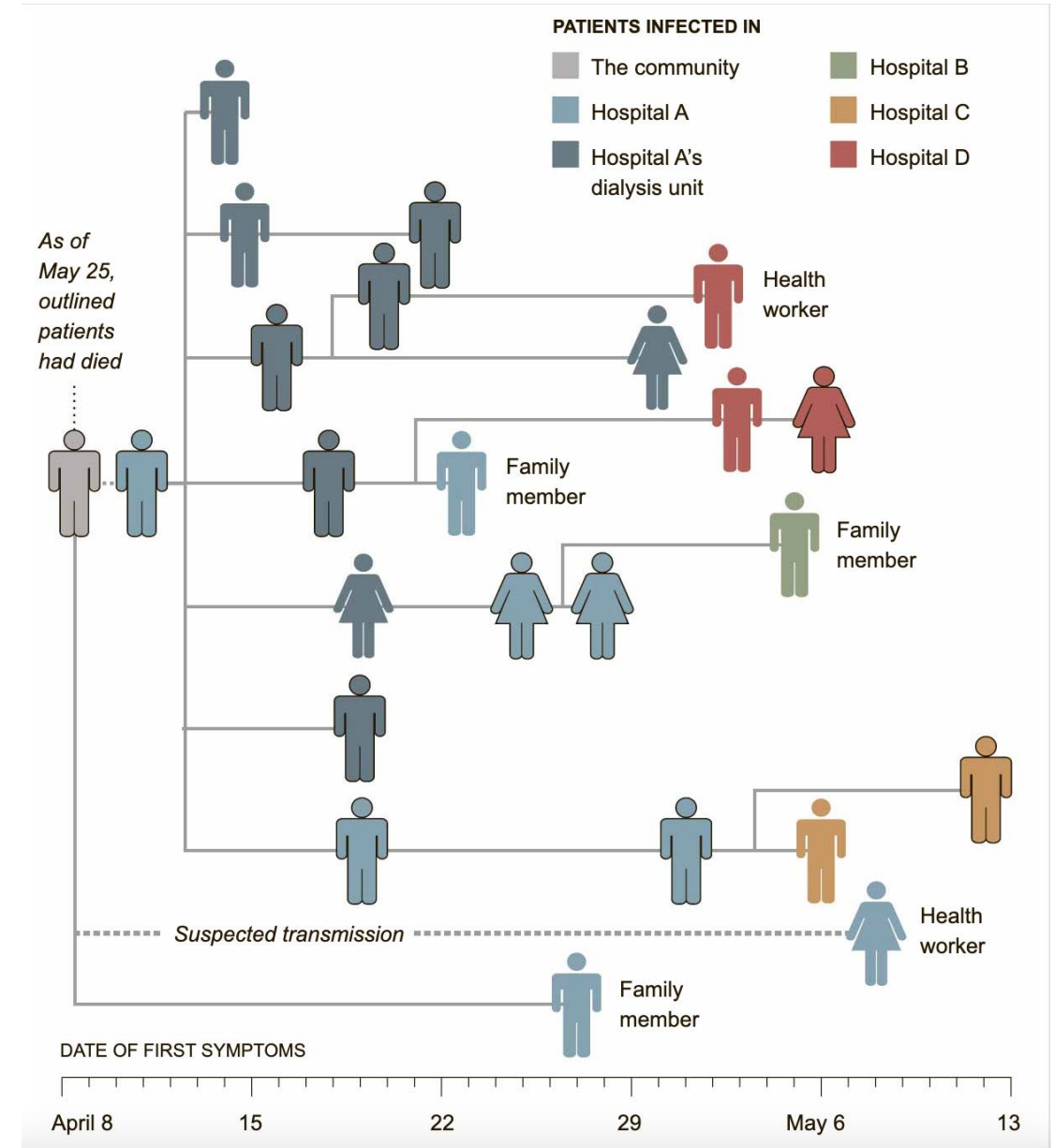Gu et al. 2023    5
Miller and Metcalf. 2019.
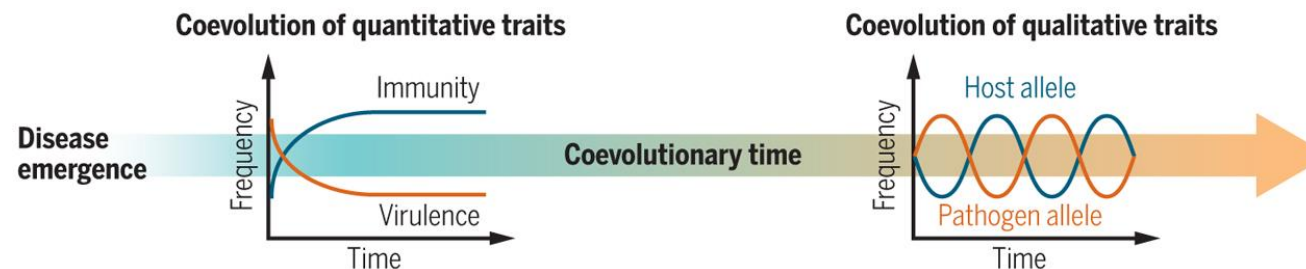
# Spillover Event Identification
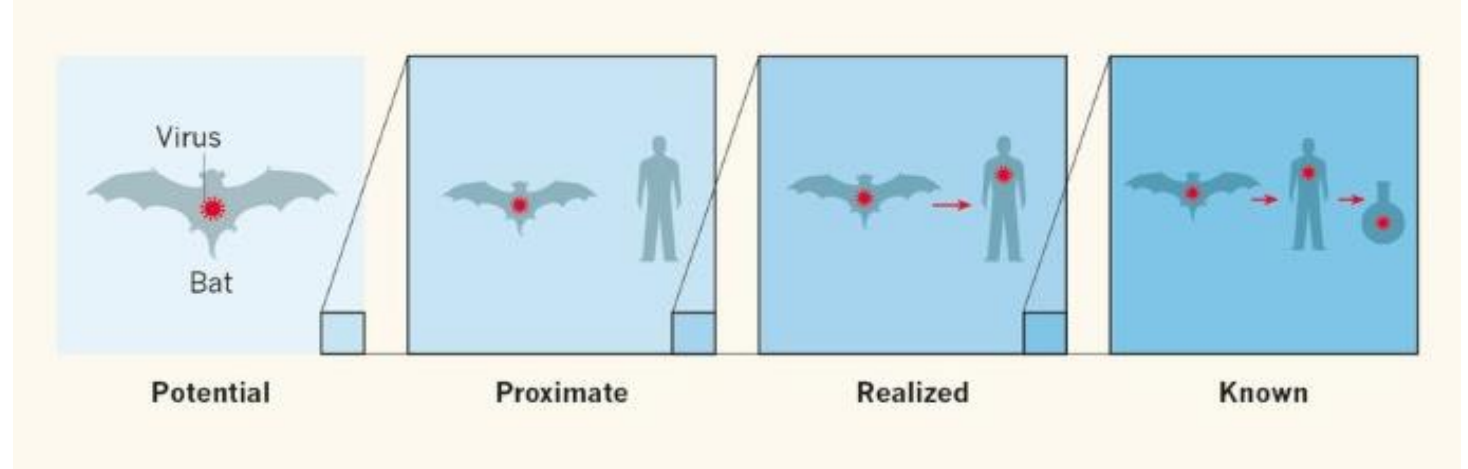


- Transmission from animals to humans
  - Can cause new variants to be introduced into circulation
- Genomic Epi successfully used to identify Ebola spillover in DRC along with GIS mapping

Lloyd-Smith. 2017.
Lee-Cruz et al. 2023.

6

# Past Outbreak Investigation

- Retrospective analysis can inform dynamics and timing of past outbreaks
  - Can inform understanding of current and future outbreaks
  - Reconstruct outbreak transmission to isolate points of introduction
  - Useful for planning new interventions and disease prevention



PATIENTS INFECTED IN

- The community
- Hospital A
- Hospital A's dialysis unit
- Hospital B
- Hospital C
- Hospital D

As of May 25, outlined patients had died

Health worker

Family member

Family member

Suspected transmission

Health worker

Family member

DATE OF FIRST SYMPTOMS

April 8    15    22    29    May 6    13

Advanced Molecular Detection
Southeast Region Bioinformatics

New York Times. 2013.    7

# Analysis of Past Trends

- Identify endemic strains across areas and time

- Discover and characterize undetected outbreaks

- Expansion of AMR strains over time

- Can overlay Epi information such as policies in place at the time or major population events

# Data to Inform Policymaking

- A review of 6,000 SARS-CoV-2 samples revealed locally specific haplotypes
  - In tandem with epi investigations supported stay-at-home orders and travel restrictions
- Addressing endemic circulation of disease
  - Vaccination campaigns
- Limit points of new introduction of disease
  - Travel restrictions
  - Zoonotic disease monitoring
- Hospital hygiene procedures, routine screening for CROs, isolation protocols

# Limitations of Retrospective Data Analysis

- Time lag between clinical usefulness and availability of genomic sequences

- Genomic data not available for most clinical cases

- Most sampling is convenience based
  - Lack of representation amongst people groups

- Limited time and availability of bioinformaticians
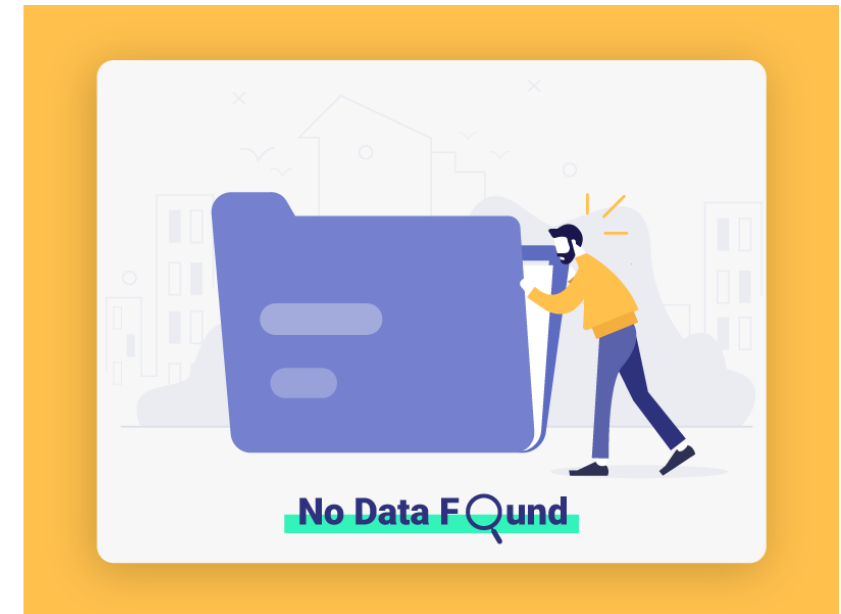
# Outbreak Forecasting

- Requires more longitudinal genomic sampling of a geographical region

  - Takes A LOT of data

  - Can utilize pathogen sequence databases such as GISAID and NCBI to increase the number of samples in the study

    - Let us know if you need help finding/creating a database for your chosen pathogen

- Goal is not to predict exact sequences or mutations but to predict prevalence of types or emergence of new strains

  - Design intervention strategies

  - Assess risk of AR

**Advanced Molecular Detection**
Southeast Region Bioinformatics

# Forecasting Models

- Models combine sequence data and estimated underlying parameters gathered from phylogenetic trees
  - Epi data and investigation can be added to clarify or verify estimated parameters
  - Usually, cannot include pathogen diversity as a model element
- Can be incorporated into existing models for enhanced clarity
  - Malaria uses population structures and immune selection
  - Incorporate genomics to identify vaccine strain candidates
- Adding genomic data to currently used forecasting models adds specificity and can be used to better prepare for outbreaks

# Conclusion

- There are many uses for retrospective genomic epidemiology studies
  - Pathogen Evolution
  - Spillover Event Identification
  - Past Outbreak Investigation
  - Past Trends
- Limitations include lack of samples and lack of staff availability
- Genomic data can augment current forecasting models to better inform decision and policymaking



Advanced Molecular Detection
Southeast Region Bioinformatics

Advanced Molecular Detection
Southeast Region Bioinformatics

# Questions?

Bphl-sebioinformatics@flhealth.gov

**TBD**
Lead Bioinformatician & Supervisor
TBD@flhealth.gov

**Molly Mitchell, PhD**
Bioinformatician
Molly.Mitchell@flhealth.gov

**Lakshmi Thsaliki, MS**
Bioinformatician
Lakshmi.Thsaliki@flhealth.gov

**Sam Marcellus, MPH**
Bioinformatician
Samantha.marcellus@flhealth.gov