Outbreak Analysis

CSCI6802 - Finlay Maguire

Overview

- Define outbreaks and the range of outbreak types
- Explain outbreak investigations and the role of genomics
- Apply typing, clustering, and phylogenetics to iteratively find and refine a set of outbreak isolates
- Construct and interpret outbreak phylogenies using reference-free and reference-based approaches
- Evaluate the impact of recombination on outbreak analysis
- Interpret phylogenies in the context of transmission
- Apply the general principles of transmission inference

What is an outbreak?

Outbreaks are defined relative to baseline surveillance

“An outbreak is an increase in the occurrence of a disease above the baseline or expected endemic level” - *Canada IPAC Program Standards 2024*

Outbreaks are defined relative to baseline surveillance

"An outbreak is an increase in the occurrence of a disease above the baseline or expected endemic level" - *Canada IPAC Program Standards 2024*

- Within a period of **time** - *within 1 month*

- Connected to a **location** - *specific ward, hospital, or region*

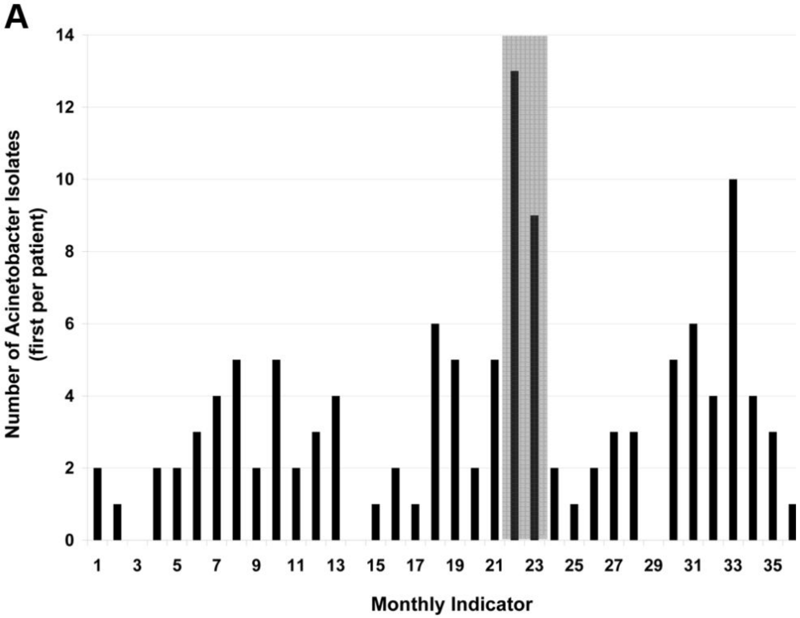
- In a **population** - *hospitalised patients, travellers*

- Specific **pathogens** - *rabies, viral hemorrhagic fevers, polio*

Outbreaks are defined relative to baseline surveillance

"An outbreak is an increase in the occurrence of a disease above the baseline or expected endemic level" - *Canada IPAC Program Standards 2024*

- Within a period of **time** - *within 1 month*
- Connected to a **location** - *specific ward, hospital, or region*
- In a **population** - *hospitalised patients, travellers*
- Specific **pathogens** - *rabies, viral hemorrhagic fevers, polio*

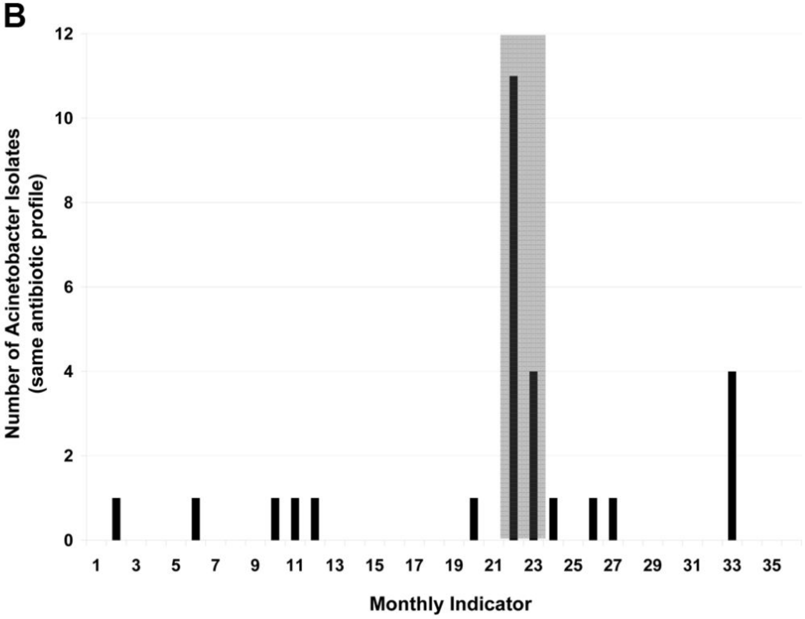


Huang, Susan S., et al. "Automated detection of infectious disease outbreaks in hospitals: a retrospective cohort study." PLoS medicine 7.2 (2010): e1000238.

Outbreaks are defined relative to baseline surveillance

"An outbreak is an increase in the occurrence of a disease above the baseline or expected endemic level" - *Canada IPAC Program Standards 2024*

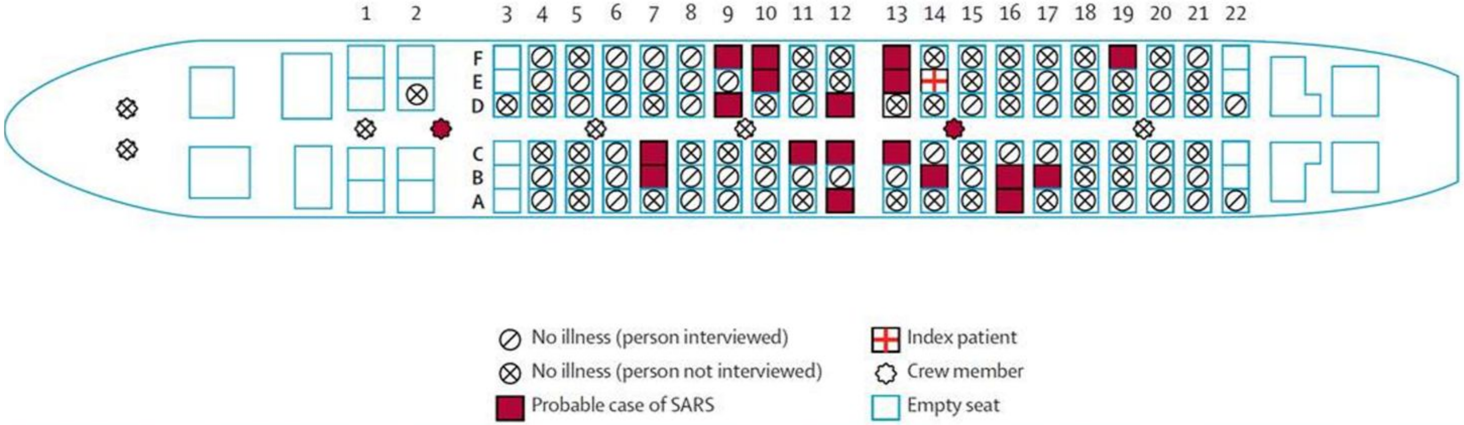
- Within a period of **time** - *within 1 month*
- Connected to a **location** - *specific ward, hospital, or region*
- In a **population** - *hospitalised patients, travellers*
- Specific **pathogens** - *rabies, viral hemorrhagic fevers, polio*
- Specific **phenotypes** - *carbapenemase*



Huang, Susan S., et al. "Automated detection of infectious disease outbreaks in hospitals: a retrospective cohort study." PLoS medicine 7.2 (2010): e1000238.

Many types of outbreak (scales, pathogen, mechanisms)

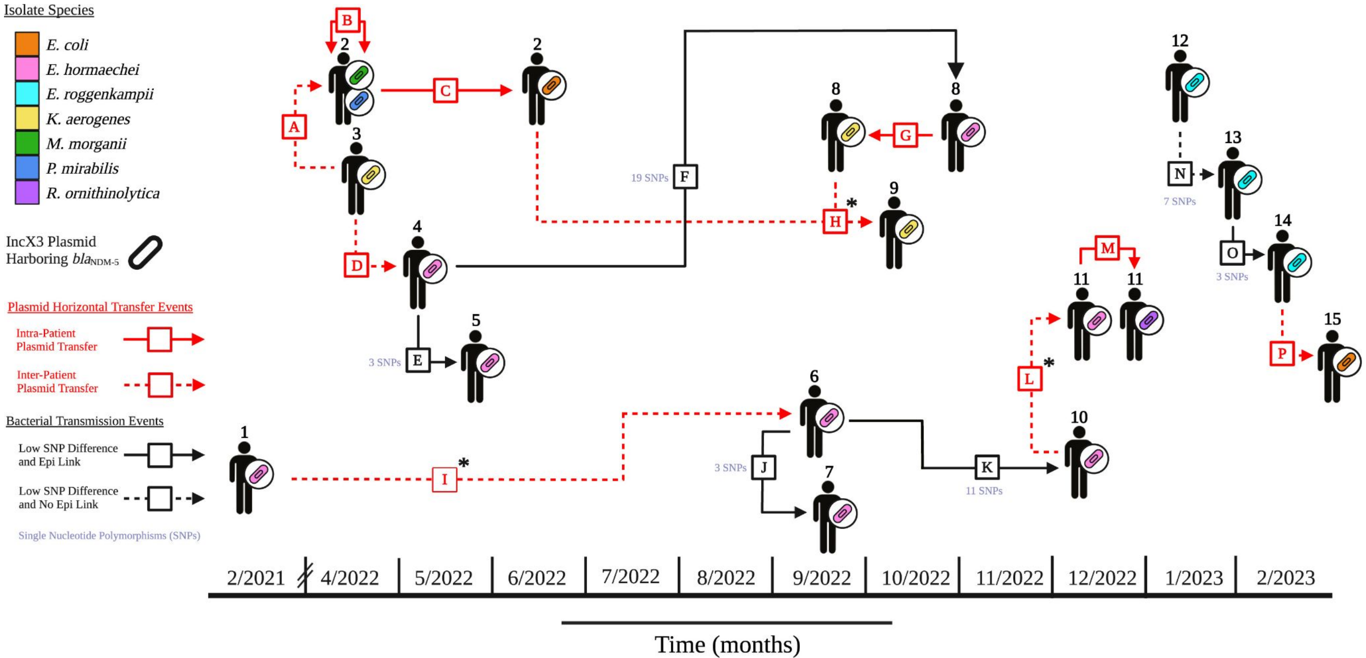
- Within an individual location:
    - March 15 2003: Boeing 737-300 with 1 SARS+ leading to a 22 person outbreak



Olsen, Sonja J., et al. "Transmission of the severe acute respiratory syndrome on aircraft." New England Journal of Medicine 349.25 (2003): 2416-2422.

Many types of outbreak (scales, pathogen, mechanisms)

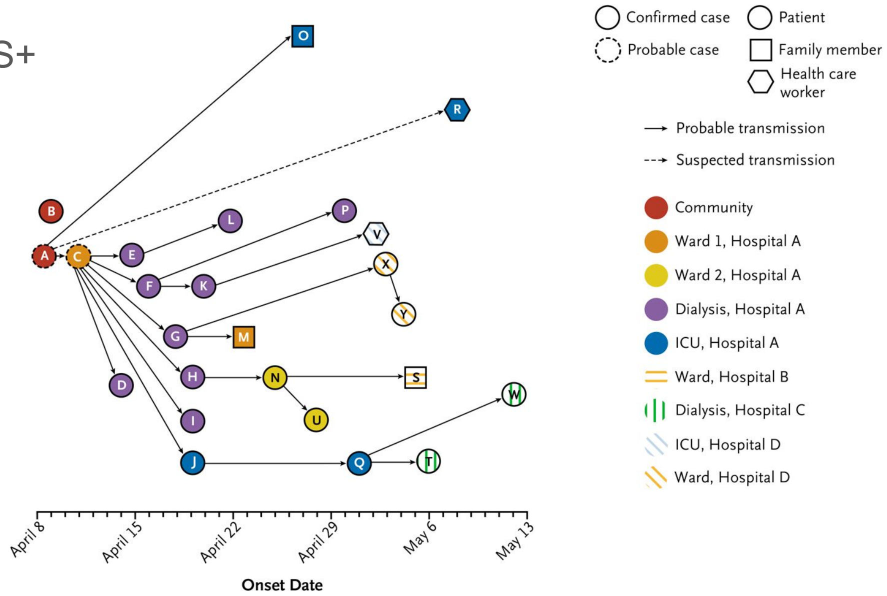
- Within an individual location:
  - March 15 2003: Boeing 737-300 with 1 SARS+ leading to a 22 person outbreak
  - 2021-2023: UPMC Presbyterian Hospital 15 patient & 7 species NDM-5 plasmid outbreak



Raabe, Nathan J., et al. "Real-time genomic epidemiologic investigation of a multispecies plasmid-associated hospital outbreak of NDM-5-producing Enterobacterales infections." *International Journal of Infectious Diseases* 142 (2024): 106971.

Many types of outbreak (scales, pathogen, mechanisms)

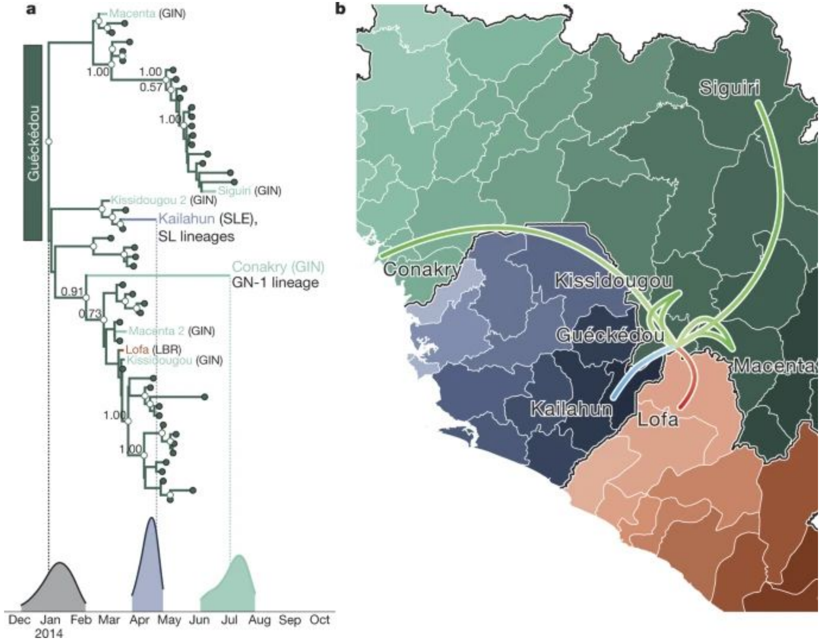
- Within an individual location:
  - March 15 2003: Boeing 737-300 with 1 SARS+ leading to a 22 person outbreak
  - 2021-2023: UPMC Presbyterian Hospital 15 patient & 7 species NDM-5 plasmid outbreak
- Across facilities nationally:
  - 2013: Multi-hospital MERS outbreak



Assiri, Abdullah, et al. "Hospital outbreak of Middle East respiratory syndrome coronavirus." *New England Journal of Medicine* 369.5 (2013): 407-416.

Many types of outbreak (scales, pathogen, mechanisms)

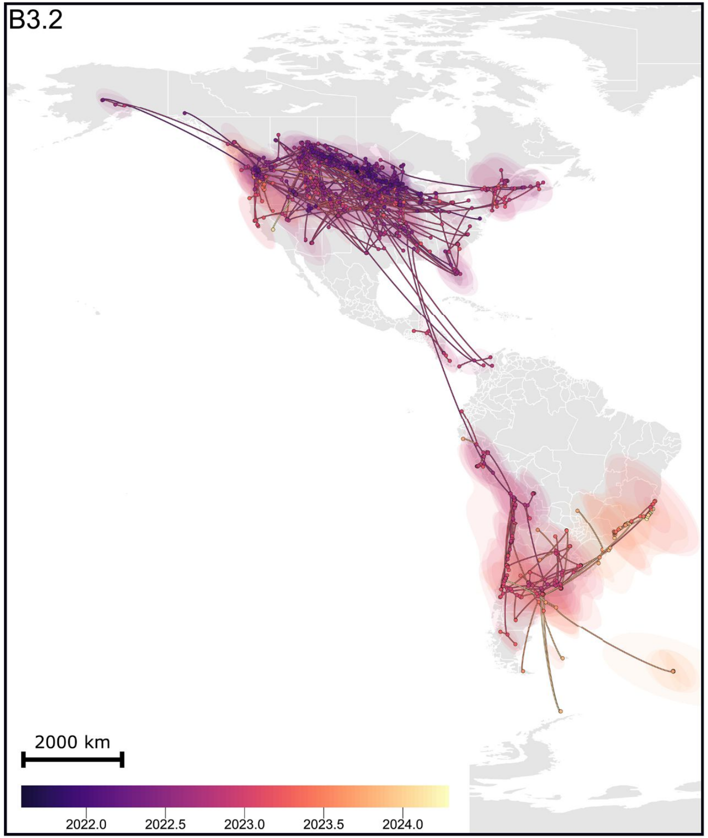
- Within an individual location:
  - March 15 2003: Boeing 737-300 with 1 SARS+ leading to a 22 person outbreak
  - 2021-2023: UPMC Presbyterian Hospital 15 patient & 7 species NDM-5 plasmid outbreak
- Across facilities nationally:
  - 2013: Multi-hospital MERS outbreak
- Internationally:
  - 2013-2016: West African Ebola epidemic



Dudas, Gytis, et al. "Virus genomes reveal factors that spread and sustained the Ebola epidemic." *Nature* 544.7650 (2017): 309-315.

Many types of outbreak (scales, pathogen, mechanisms)

- Within an individual location:
  - March 15 2003: Boeing 737-300 with 1 SARS+ leading to a 22 person outbreak
  - 2021-2023: UPMC Presbyterian Hospital 15 patient & 7 species NDM-5 plasmid outbreak
- Across facilities nationally:
  - 2013: Multi-hospital MERS outbreak
- Internationally:
  - 2013-2016: West African Ebola epidemic
  - 2022-Present: 2.3.4.4b North America HPAI

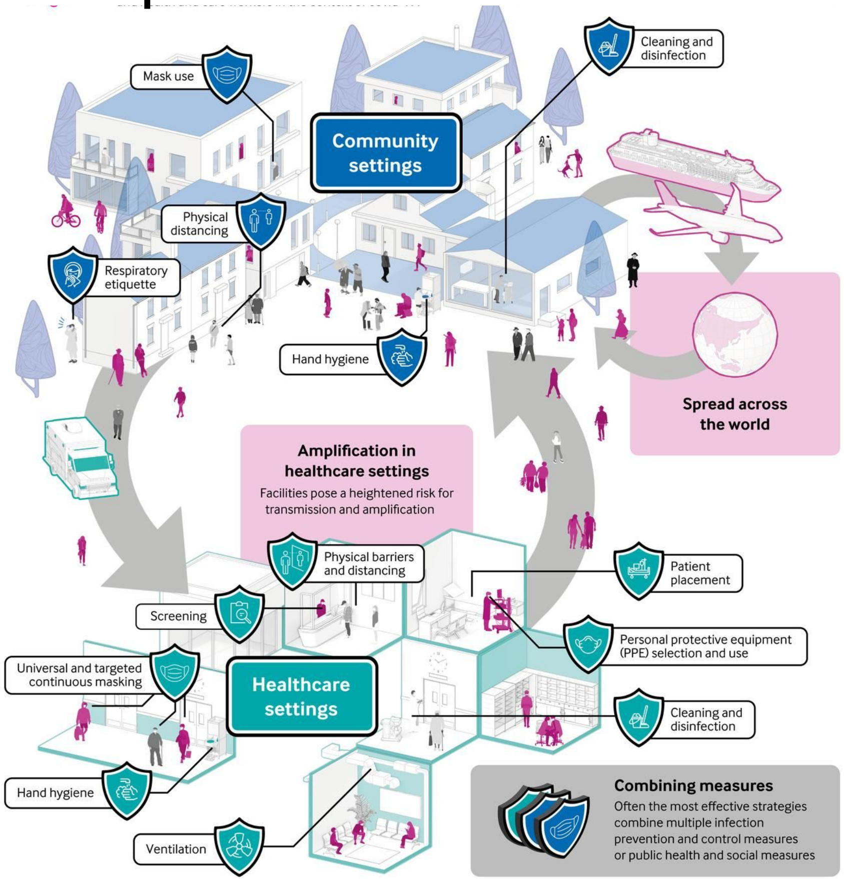


Signore, Anthony V., et al. "Spatiotemporal reconstruction of the North American A (H5N1) outbreak reveals successive lineage replacements by descendant reassortants." *Science Advances* 11.28 (2025): eadu4909.

Why investigate outbreaks?

Goal of investigation is to control & prevent further cases

- Control measures
    - Where are new cases coming from?
    - How do we remove that source?
    - Should we close a facility?
- Risk assessment
    - Who is at risk?
    - Why are they at risk?
- Cleaning & Disinfecting
    - What needs cleaning?
    - Which cleaning agent?
- Prophylaxis, Treatment & Vaccination
    - Which drugs and which vaccine?
    - Who should be prioritised?



Dunn, Kathleen, et al. "Summary of WHO infection prevention and control guideline for covid-19: striving for evidence based practice in infection prevention and control." bmj 385 (2024).

So, how do we actually investigate an outbreak and how is genomics involved?

Sequencing useful throughout outbreak investigations

- **Identification**: confirm outbreak & verify diagnosis
- **Generic Control**: Immediate/generic control measures

Iteratively:

- **Define**: develop a case definition
- **Find**: systematically identify potential linked cases
- **Describe**: epidemiological characters (time/place/person)
- **Hypothesise**: exposure, source, control, transmission
- **Evaluate**: test these hypotheses
- **Specific Control**: agent-specific control measures
- **Communicate**: develop public health messaging

Sequencing useful throughout outbreak investigations

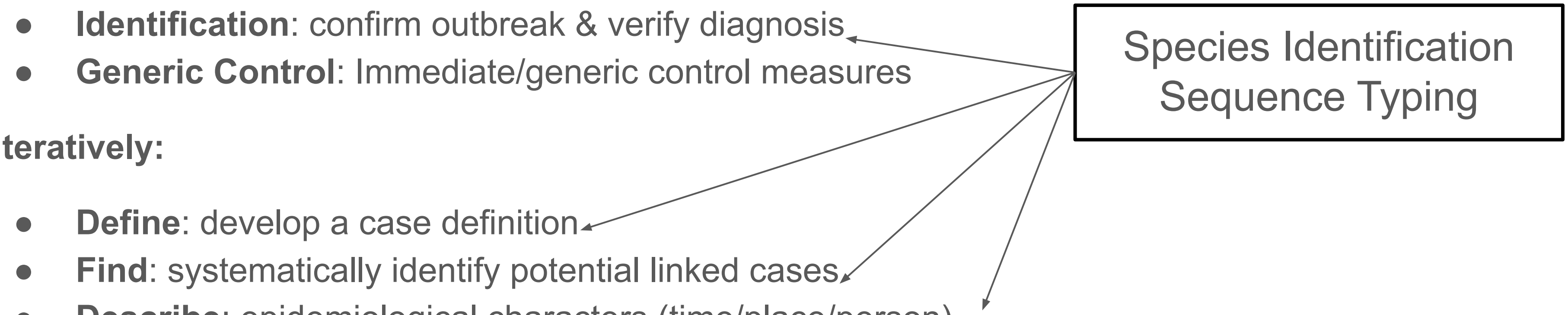
- **Identification**: confirm outbreak & verify diagnosis
- **Generic Control**: Immediate/generic control measures

Iteratively:

- **Define**: develop a case definition
- **Find**: systematically identify potential linked cases
- **Describe**: epidemiological characters (time/place/person)
- **Hypothesise**: exposure, source, control, transmission
- **Evaluate**: test these hypotheses
- **Specific Control**: agent-specific control measures
- **Communicate**: develop public health messaging

Data Curation/Sharing

Sequencing useful throughout outbreak investigations



- **Identification**: confirm outbreak & verify diagnosis
- **Generic Control**: Immediate/generic control measures

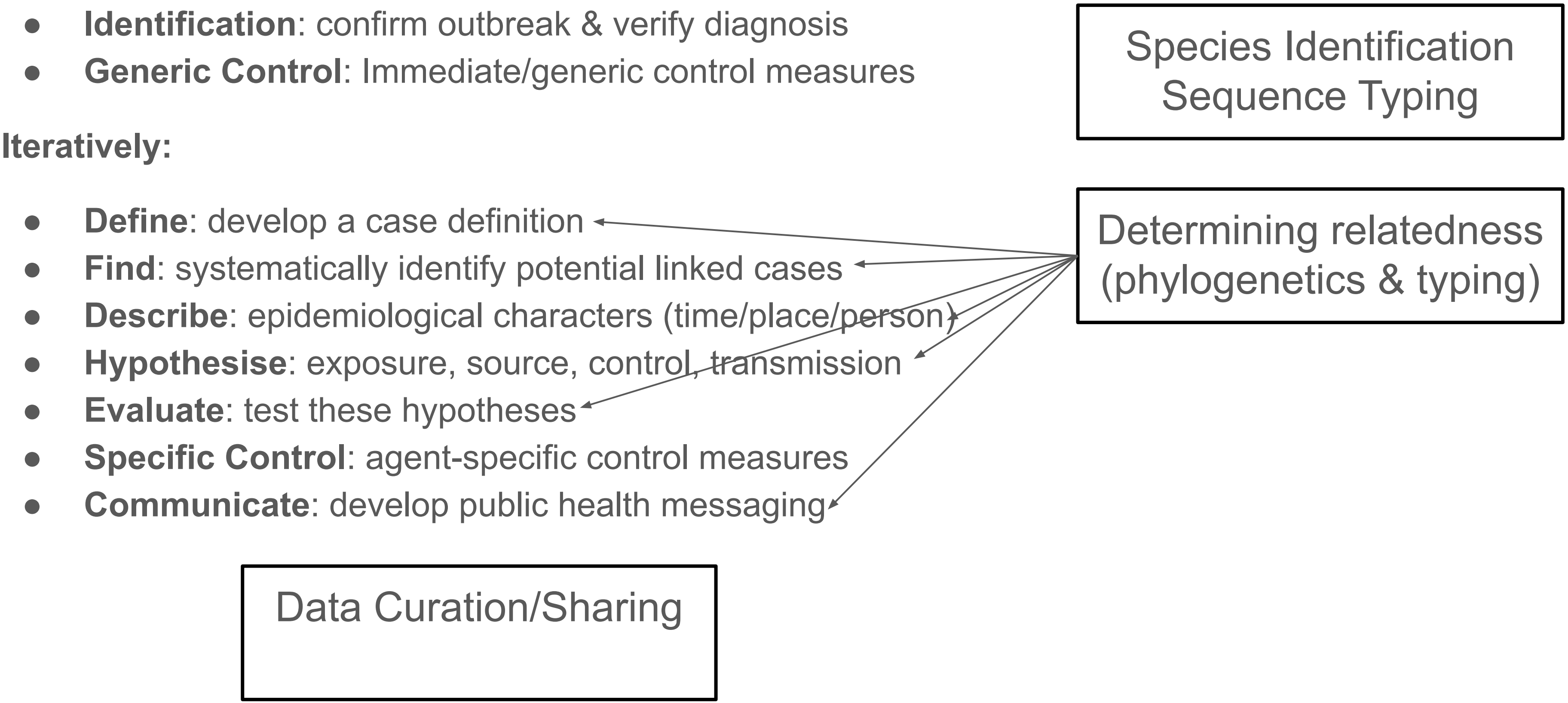
Iteratively:

- **Define**: develop a case definition
- **Find**: systematically identify potential linked cases
- **Describe**: epidemiological characters (time/place/person)
- **Hypothesise**: exposure, source, control, transmission
- **Evaluate**: test these hypotheses
- **Specific Control**: agent-specific control measures
- **Communicate**: develop public health messaging

Species Identification
Sequence Typing

Data Curation/Sharing

Sequencing useful throughout outbreak investigations



- **Identification**: confirm outbreak & verify diagnosis
- **Generic Control**: Immediate/generic control measures

Iteratively:

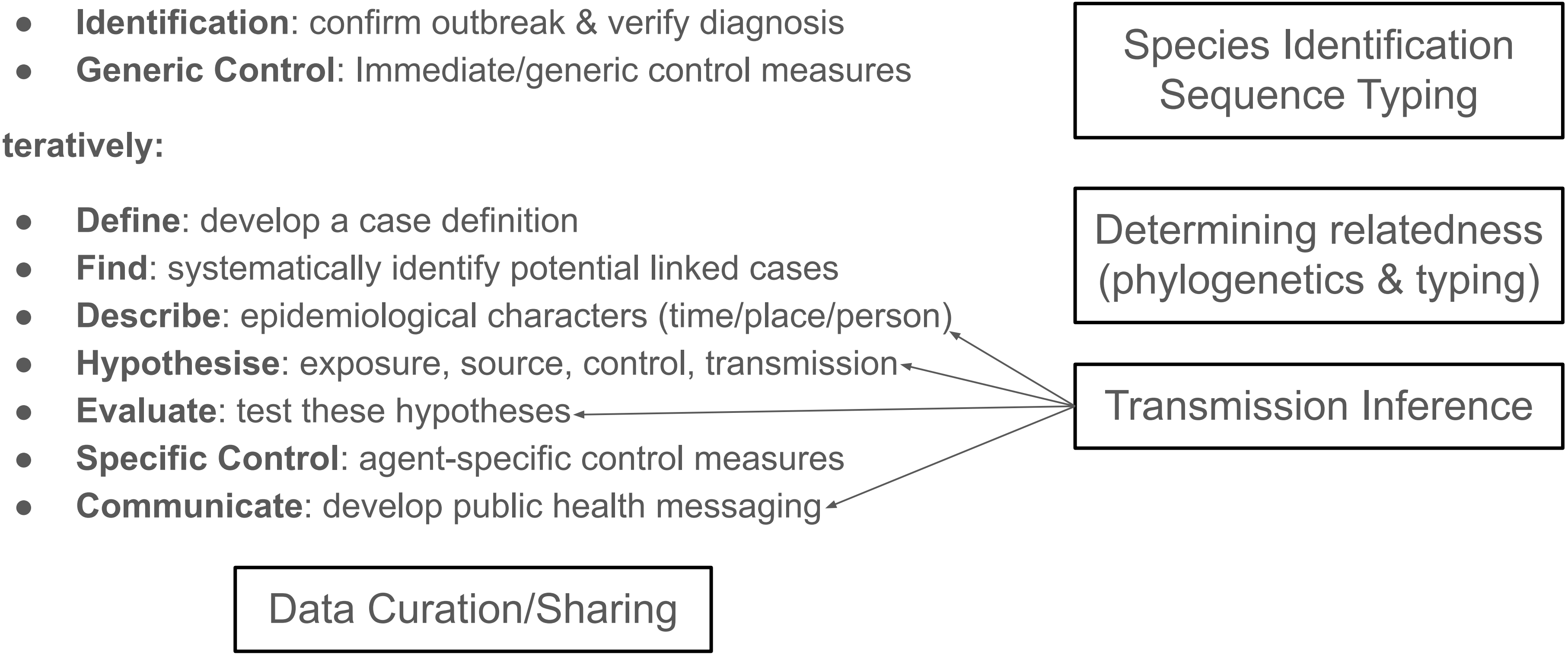
- **Define**: develop a case definition
- **Find**: systematically identify potential linked cases
- **Describe**: epidemiological characters (time/place/person)
- **Hypothesise**: exposure, source, control, transmission
- **Evaluate**: test these hypotheses
- **Specific Control**: agent-specific control measures
- **Communicate**: develop public health messaging

Species Identification
Sequence Typing

Determining relatedness
(phylogenetics & typing)

Data Curation/Sharing

Sequencing useful throughout outbreak investigations



- **Identification**: confirm outbreak & verify diagnosis
- **Generic Control**: Immediate/generic control measures

Iteratively:

- **Define**: develop a case definition
- **Find**: systematically identify potential linked cases
- **Describe**: epidemiological characters (time/place/person)
- **Hypothesise**: exposure, source, control, transmission
- **Evaluate**: test these hypotheses
- **Specific Control**: agent-specific control measures
- **Communicate**: develop public health messaging

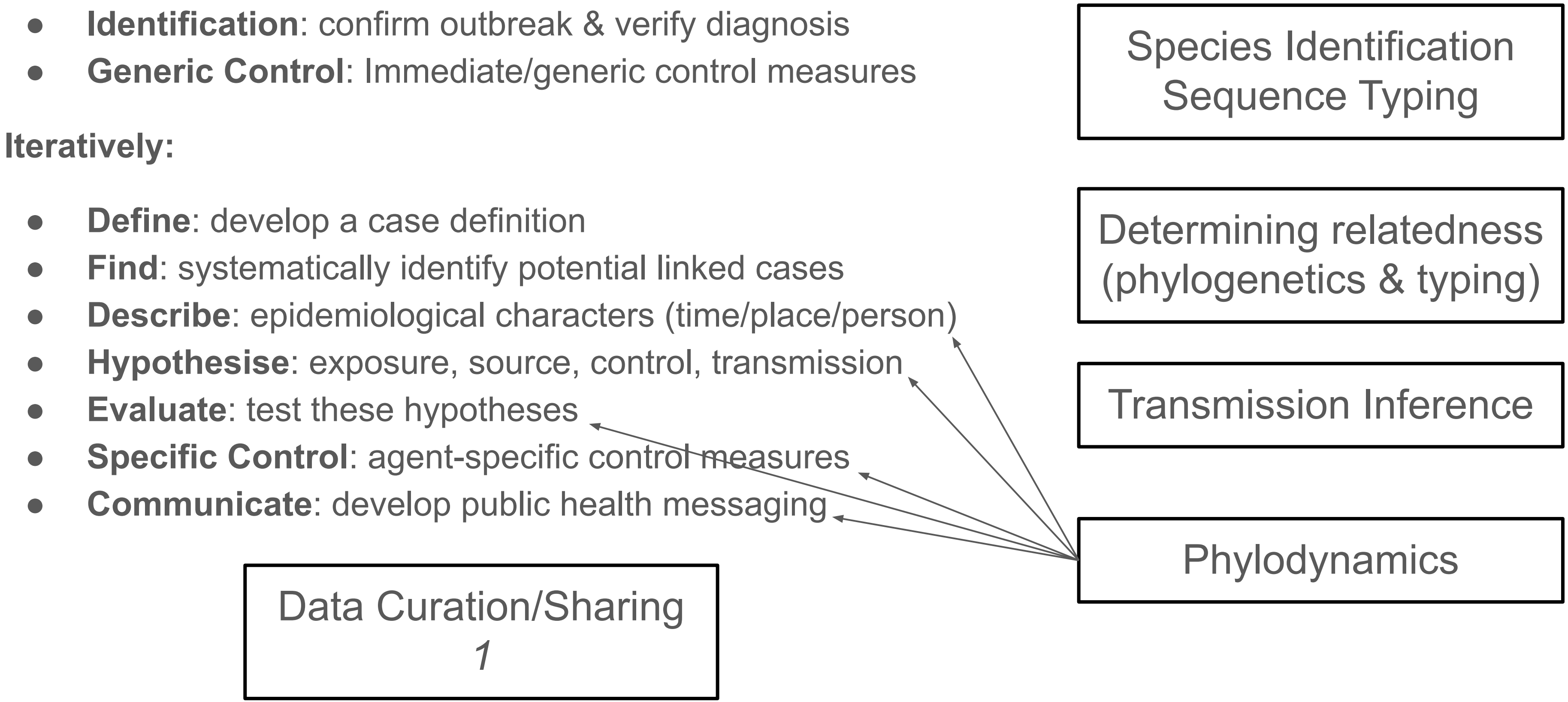
Species Identification
Sequence Typing

Determining relatedness
(phylogenetics & typing)

Transmission Inference

Data Curation/Sharing

Sequencing useful throughout outbreak investigations



- **Identification**: confirm outbreak & verify diagnosis
- **Generic Control**: Immediate/generic control measures

Iteratively:

- **Define**: develop a case definition
- **Find**: systematically identify potential linked cases
- **Describe**: epidemiological characters (time/place/person)
- **Hypothesise**: exposure, source, control, transmission
- **Evaluate**: test these hypotheses
- **Specific Control**: agent-specific control measures
- **Communicate**: develop public health messaging

Species Identification
Sequence Typing

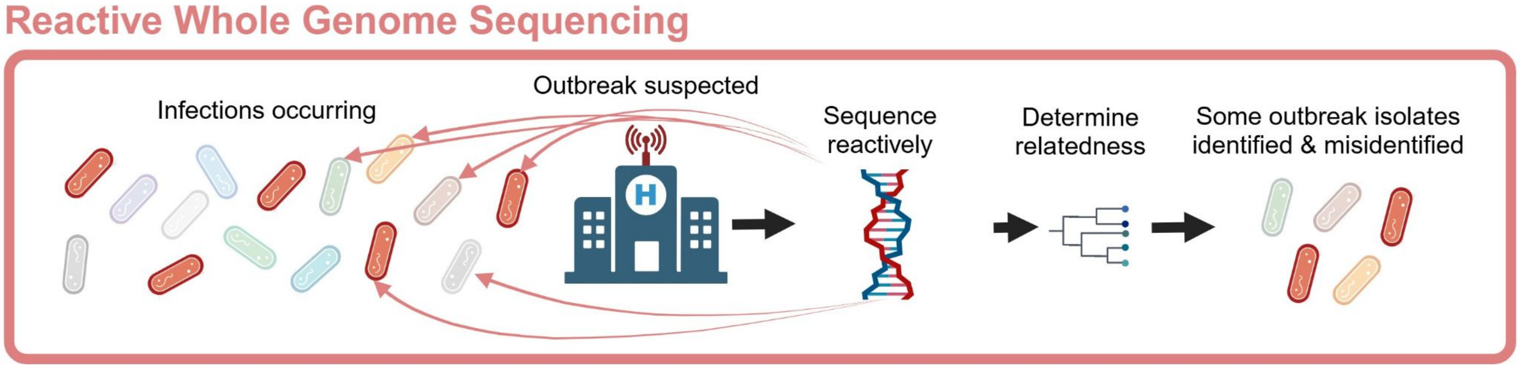
Determining relatedness
(phylogenetics & typing)

Transmission Inference

Phylodynamics

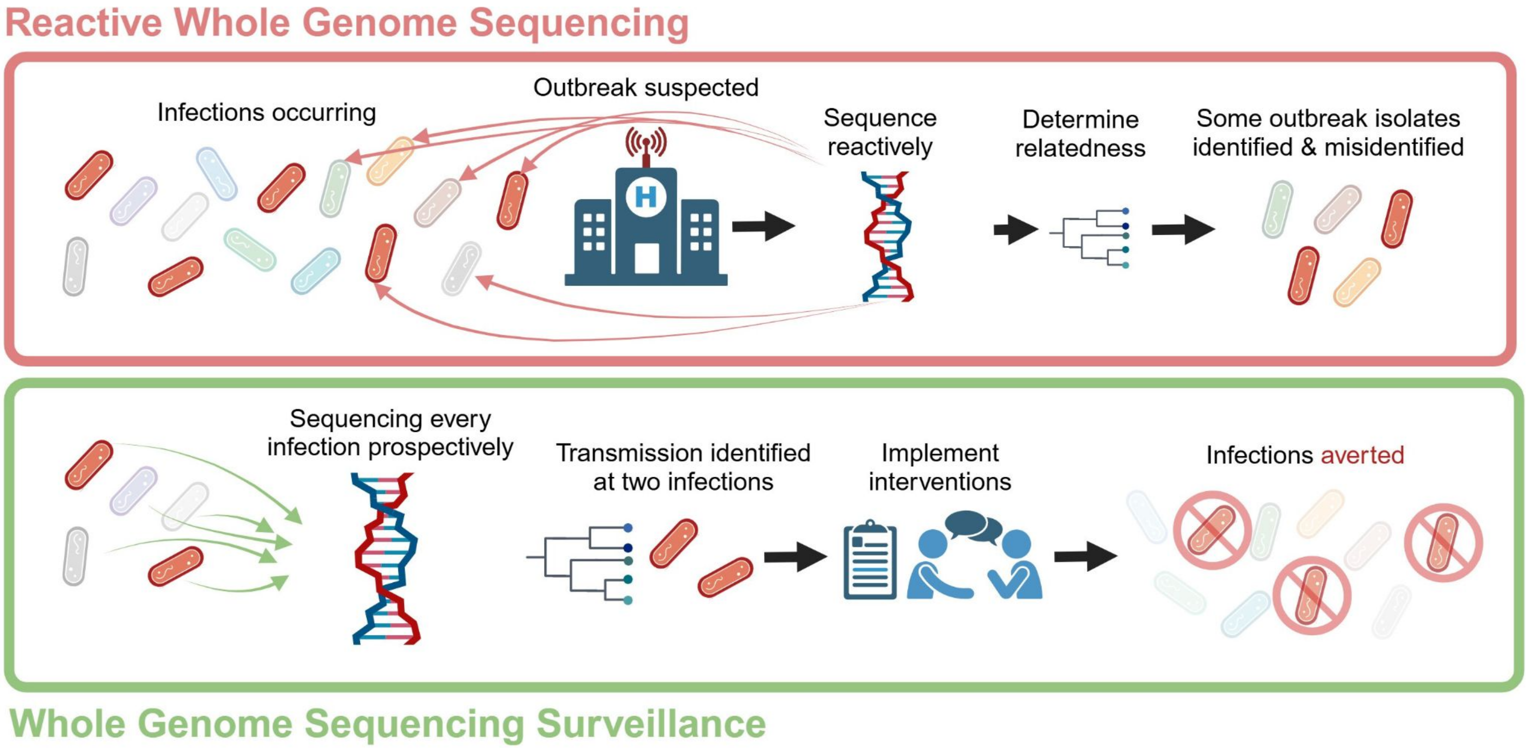
Data Curation/Sharing
1

Genomics can be incorporated reactively or prospectively



Sundermann, Alexander J., et al. "Pathogen genomics in healthcare: overcoming barriers to proactive surveillance." Antimicrobial agents and chemotherapy 69.1 (2025): e01479-24.

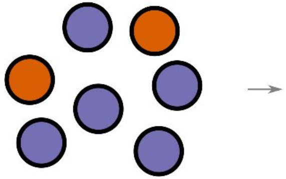
Genomics can be incorporated reactively or prospectively



Sundermann, Alexander J., et al. "Pathogen genomics in healthcare: overcoming barriers to proactive surveillance." Antimicrobial agents and chemotherapy 69.1 (2025): e01479-24.

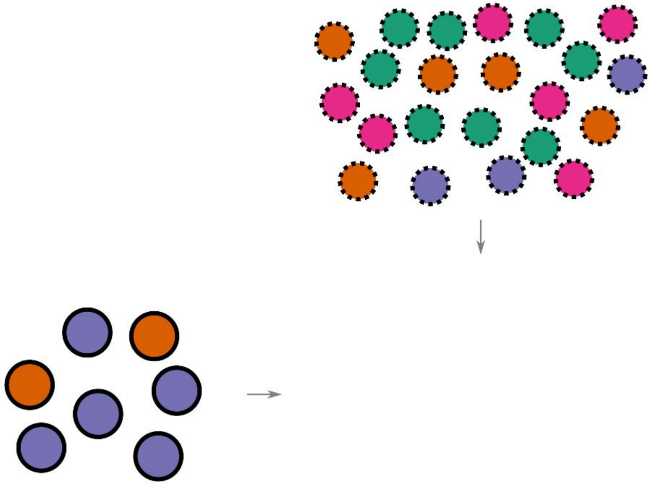
So, how do we identify outbreak-linked isolates using genomics?

Iteratively refine isolates with typing, clusters & phylogeny



INITIAL SUSPECTED
OUTBREAK CLUSTER

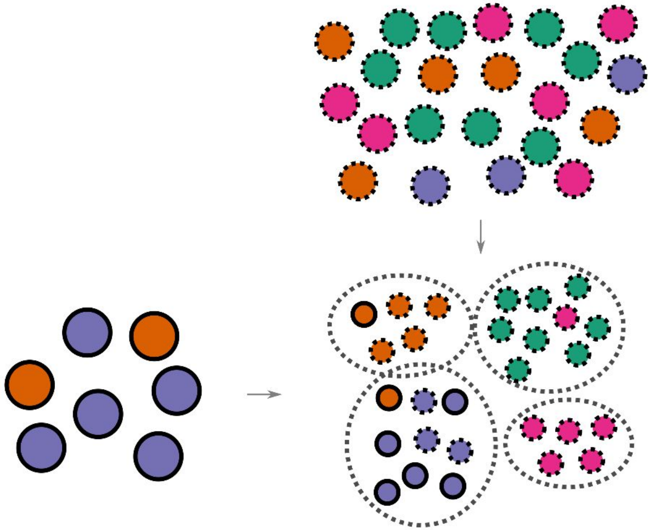
Iteratively refine isolates with typing, clusters & phylogeny



ISOLATE COLLECTION (database, sampling, lab freezers)

INITIAL SUSPECTED OUTBREAK CLUSTER

Iteratively refine isolates with typing, clusters & phylogeny

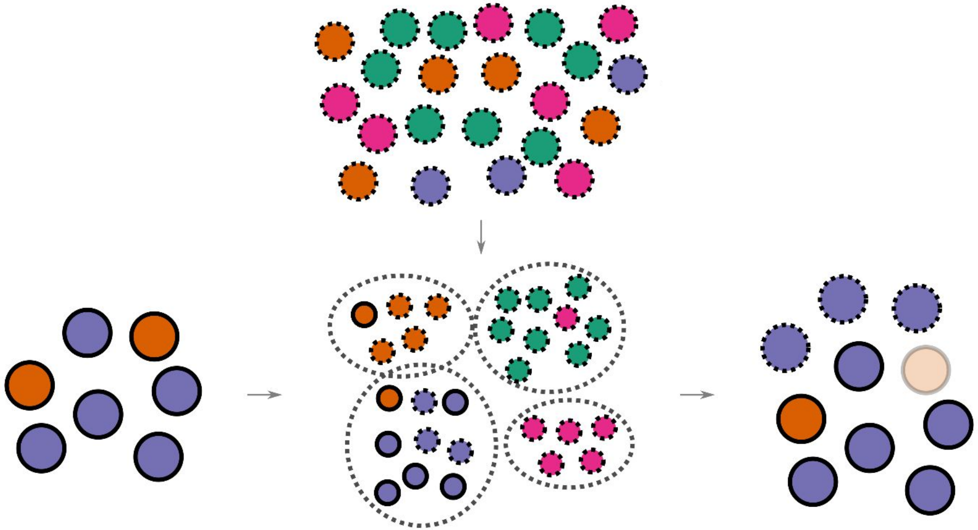


ISOLATE COLLECTION
(database, sampling,
lab freezers)

INITIAL SUSPECTED
OUTBREAK CLUSTER

RAPID SCREENING
TYPING & CLUSTERING
(MLST, MASH, AMR,
cg/wgMLST, PANGO)

Iteratively refine isolates with typing, clusters & phylogeny



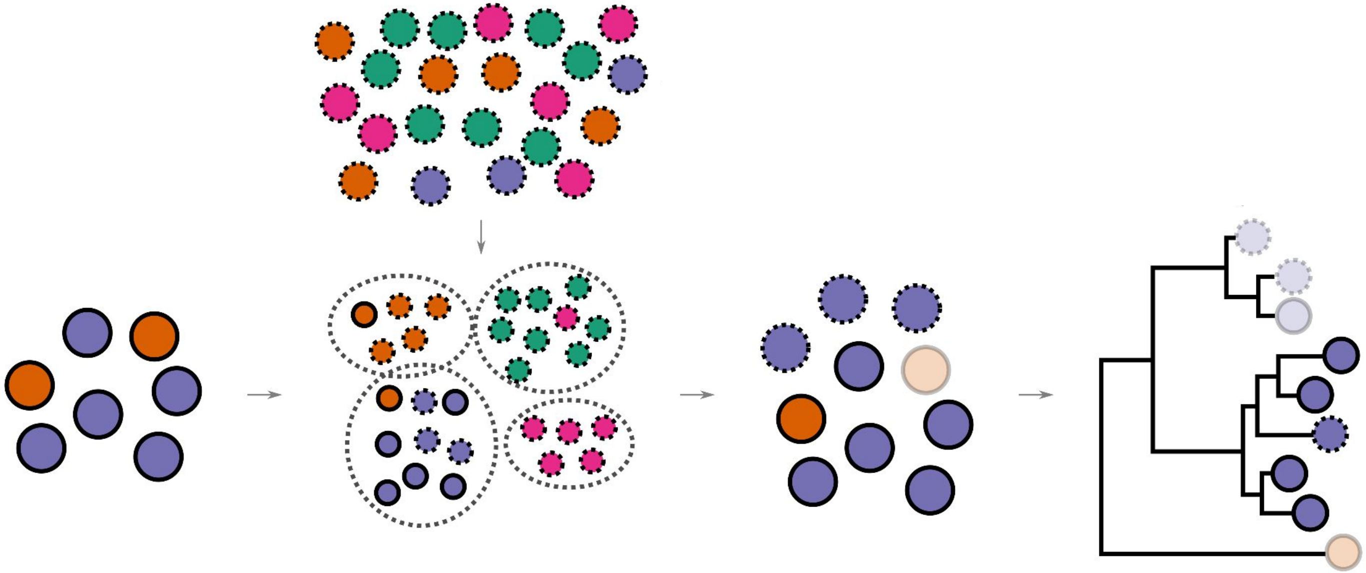
ISOLATE COLLECTION
(database, sampling,
lab freezers)

INITIAL SUSPECTED
OUTBREAK CLUSTER

RAPID SCREENING
TYPING & CLUSTERING
(MLST, MASH, AMR,
cg/wgMLST, PANGO)

REFINED OUTBREAK
CLUSTER

Iteratively refine isolates with typing, clusters & phylogeny



ISOLATE COLLECTION
(database, sampling,
lab freezers)

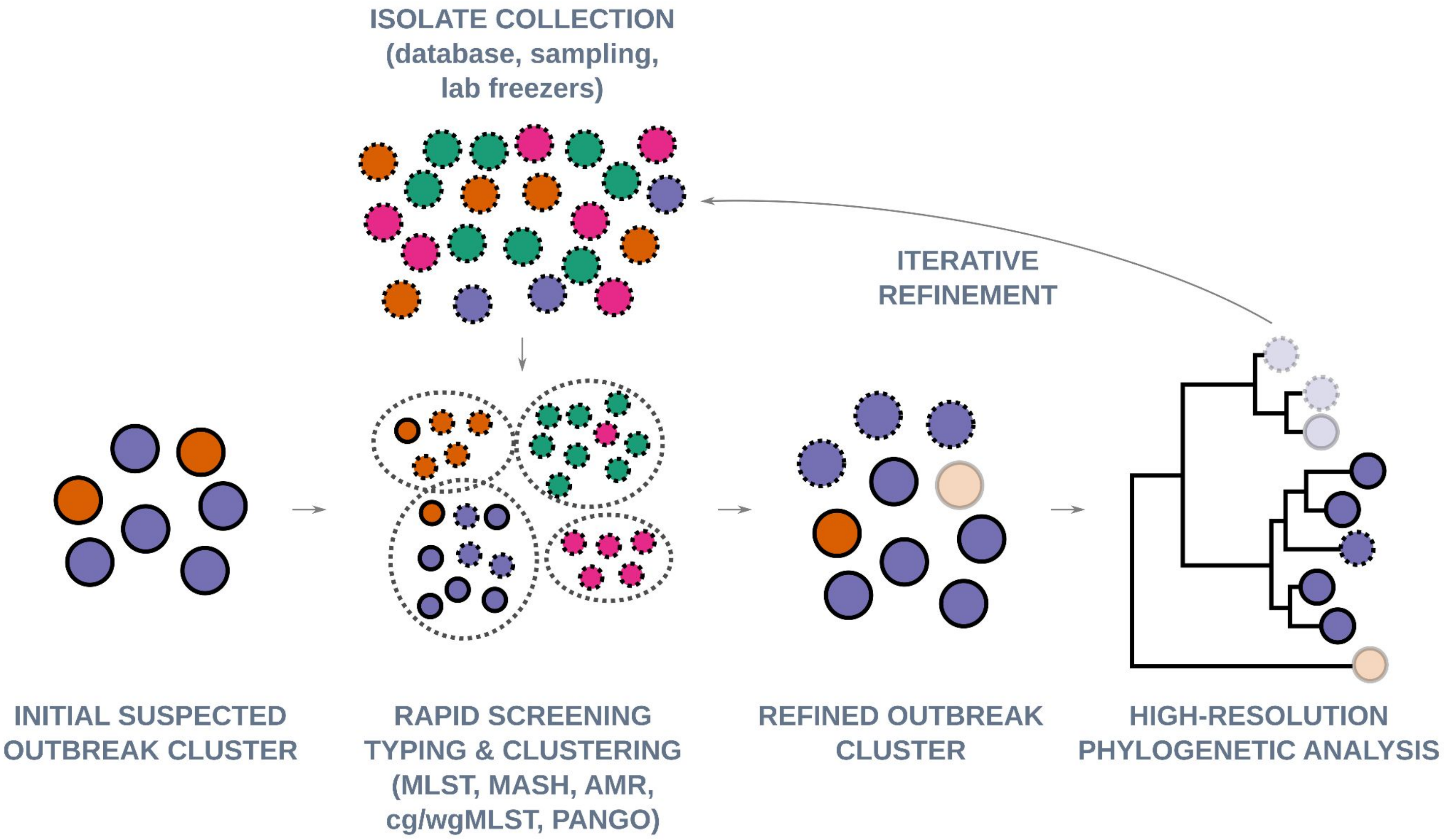
INITIAL SUSPECTED
OUTBREAK CLUSTER

RAPID SCREENING
TYPING & CLUSTERING
(MLST, MASH, AMR,
cg/wgMLST, PANGO)

REFINED OUTBREAK
CLUSTER

HIGH-RESOLUTION
PHYLOGENETIC ANALYSIS

Iteratively refine isolates with typing, clusters & phylogeny



ISOLATE COLLECTION
(database, sampling,
lab freezers)

ITERATIVE
REFINEMENT

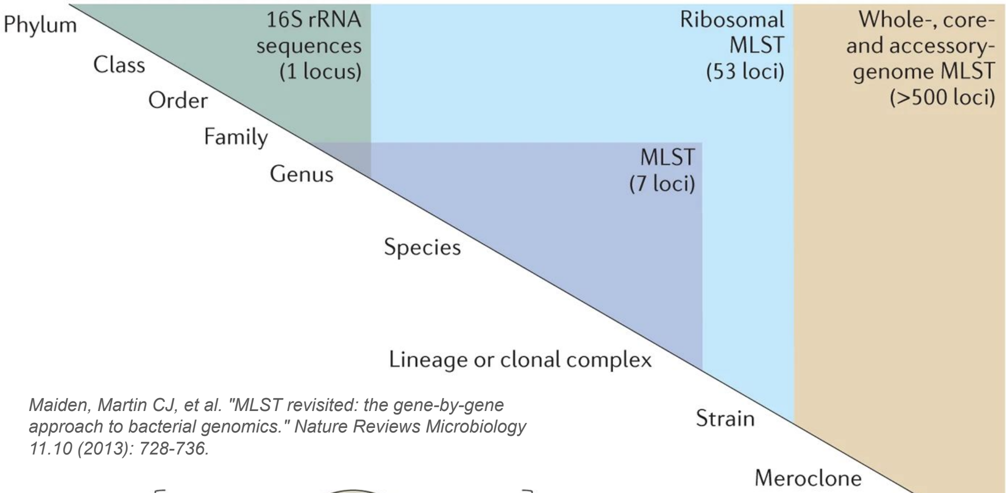
INITIAL SUSPECTED
OUTBREAK CLUSTER

RAPID SCREENING
TYPING & CLUSTERING
(MLST, MASH, AMR,
cg/wgMLST, PANGO)

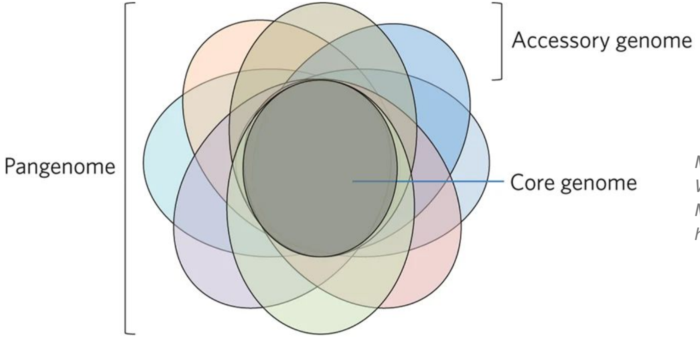
REFINED OUTBREAK
CLUSTER

HIGH-RESOLUTION
PHYLOGENETIC ANALYSIS

Appropriate typing tool is pathogen and outbreak specific



Maiden, Martin CJ, et al. "MLST revisited: the gene-by-gene approach to bacterial genomics." Nature Reviews Microbiology 11.10 (2013): 728-736.



McInerney, J., McNally, A. & O'Connell, M. Why prokaryotes have pangenomes. Nat Microbiol ***2****, 17040 (2017). https://doi.org/10.1038/nmicrobiol.2017.40*

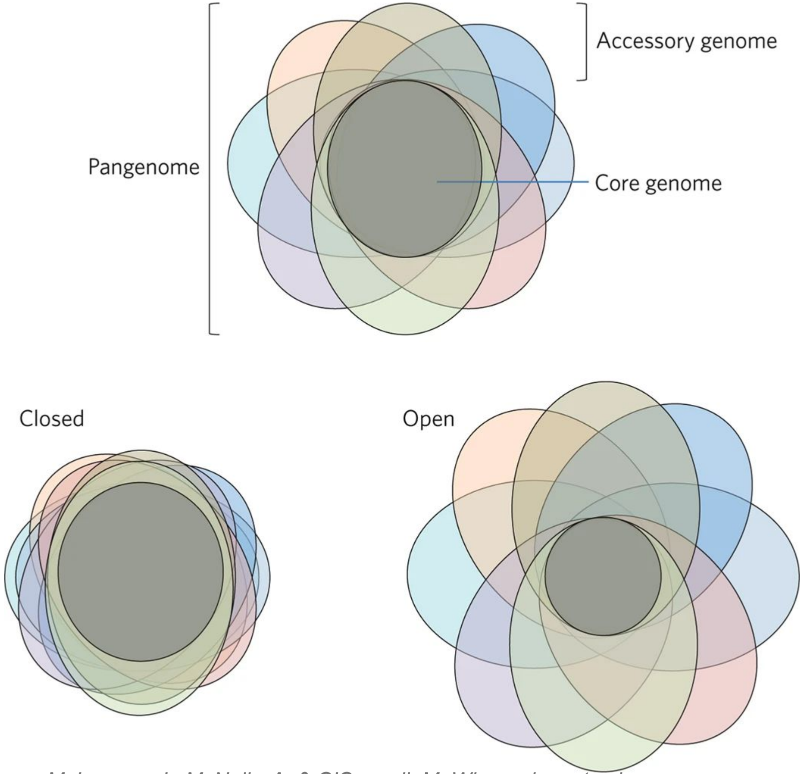
Ideally we to use a typing system:

- Appropriate for outbreak type
- Fast enough for outbreak size
- High enough resolution
- Granular but stable clusters
- Concordant with epidemiology

Pangenomics key to comparing non-viral genomes

- **Core genome**: genes in ~all isolates
- **Accessory genome**: genes with patchy distributions

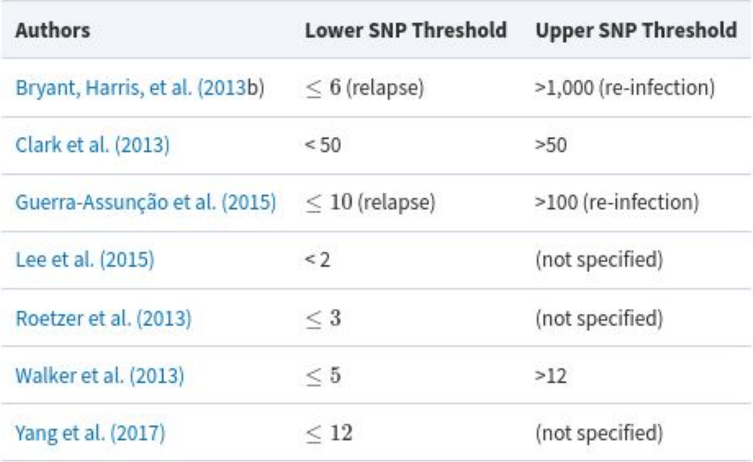
- Defined relative to isolate set
- Completeness/QC is important
- LGT in subset of isolates (e.g., plasmid acquisition) impacts accessory genome
- Less diverse set of isolates (e.g., outbreak) means larger core genome
- Large core genome means more resolution to compare isolates



McInerney, J., McNally, A. & O'Connell, M. Why prokaryotes have pangenomes. Nat Microbiol ***2****, 17040 (2017). https://doi.org/10.1038/nmicrobiol.2017.40*

Fixed threshold (SNPs, alleles) clusters are not ideal

- Fixed thresholds impose a **discrete boundary** on a **continuous**, **context-dependent**, and **probabilistic relationship**.
- Prioritize simplicity over accuracy.
- Substitution rate variation (mutation rate, infection type, drift/selection), within-host diversity, recombination, technical variation
- **Alternatives**: rate-calibrated cut-offs (lineage-specific), probabilistic clustering, transmission modelling, dynamics/adaptive thresholds

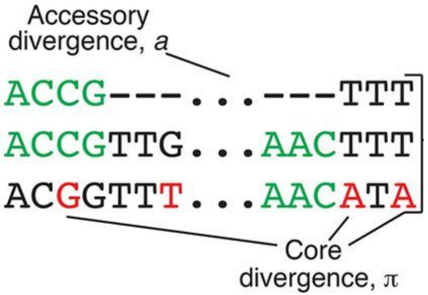


Authors	Lower SNP Threshold	Upper SNP Threshold
Bryant, Harris, et al. (2013b)	≤ 6 (relapse)	>1,000 (re-infection)
Clark et al. (2013)	< 50	>50
Guerra-Assunção et al. (2015)	≤ 10 (relapse)	>100 (re-infection)
Lee et al. (2015)	< 2	(not specified)
Roetzer et al. (2013)	≤ 3	(not specified)
Walker et al. (2013)	≤ 5	>12
Yang et al. (2017)	≤ 12	(not specified)

Stimson, James, et al. "Beyond the SNP threshold: identifying outbreak clusters using inferred transmissions." Molecular biology and evolution 36.3 (2019): 587-603.

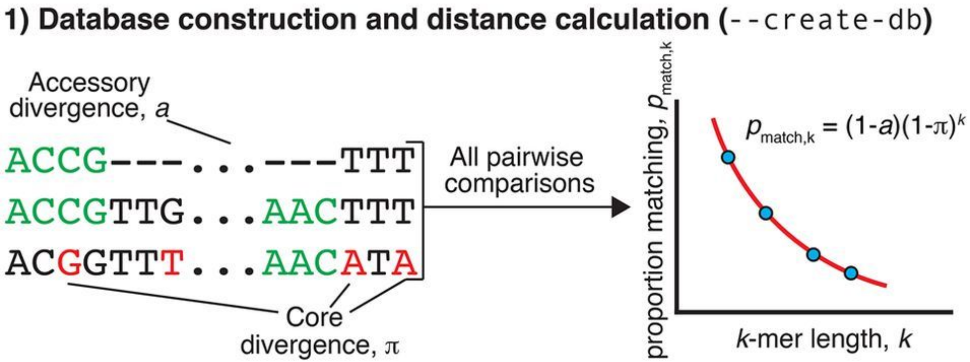
PopPUNK: rapid whole genome comparison & clustering

1) Database construction and distance calculation (`--create-db`)



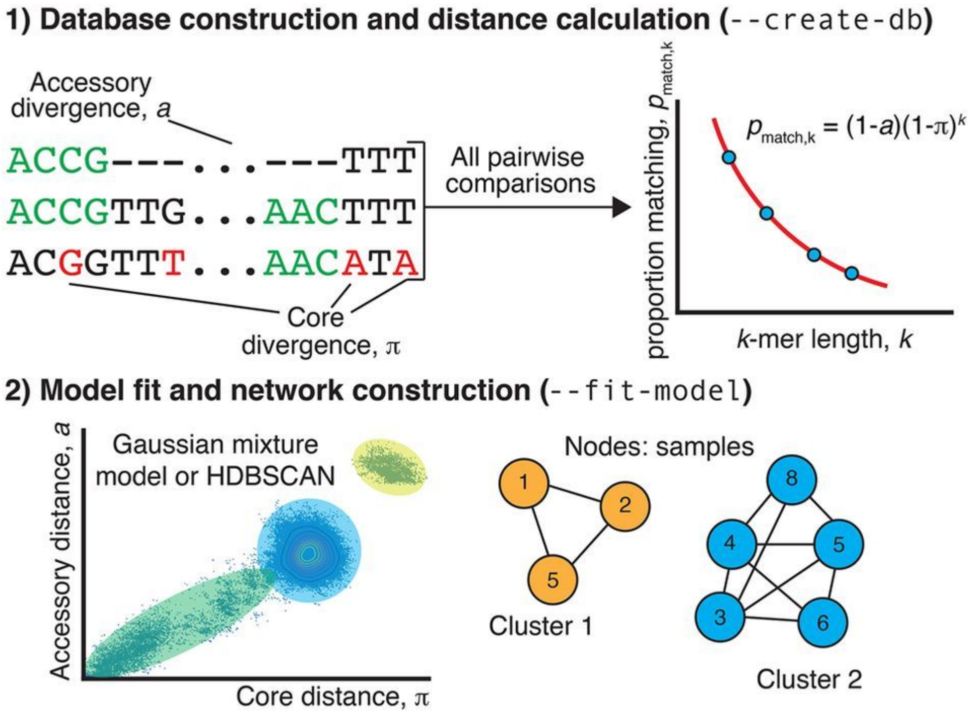
Lees, John A., et al. "Fast and flexible bacterial genomic epidemiology with PopPUNK." Genome research 29.2 (2019): 304-316.

PopPUNK: rapid whole genome comparison & clustering



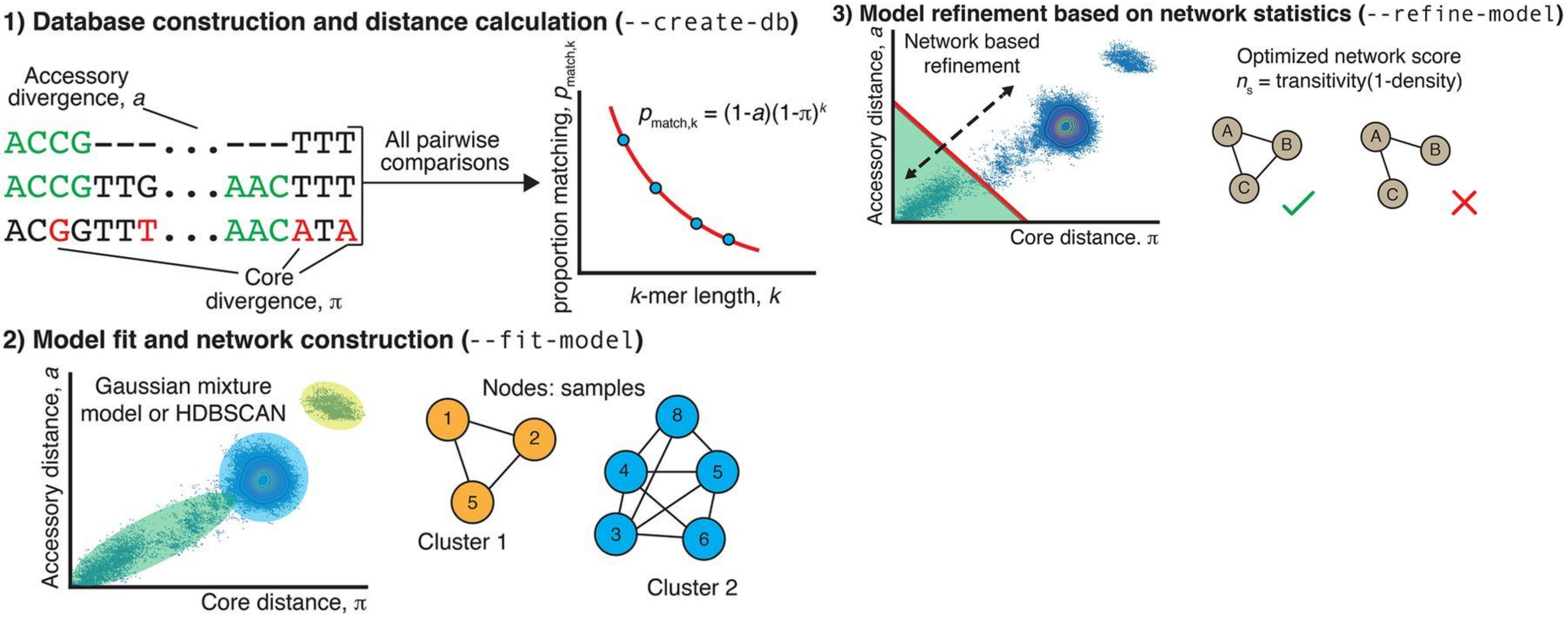
Lees, John A., et al. "Fast and flexible bacterial genomic epidemiology with PopPUNK." Genome research 29.2 (2019): 304-316.

PopPUNK: rapid whole genome comparison & clustering



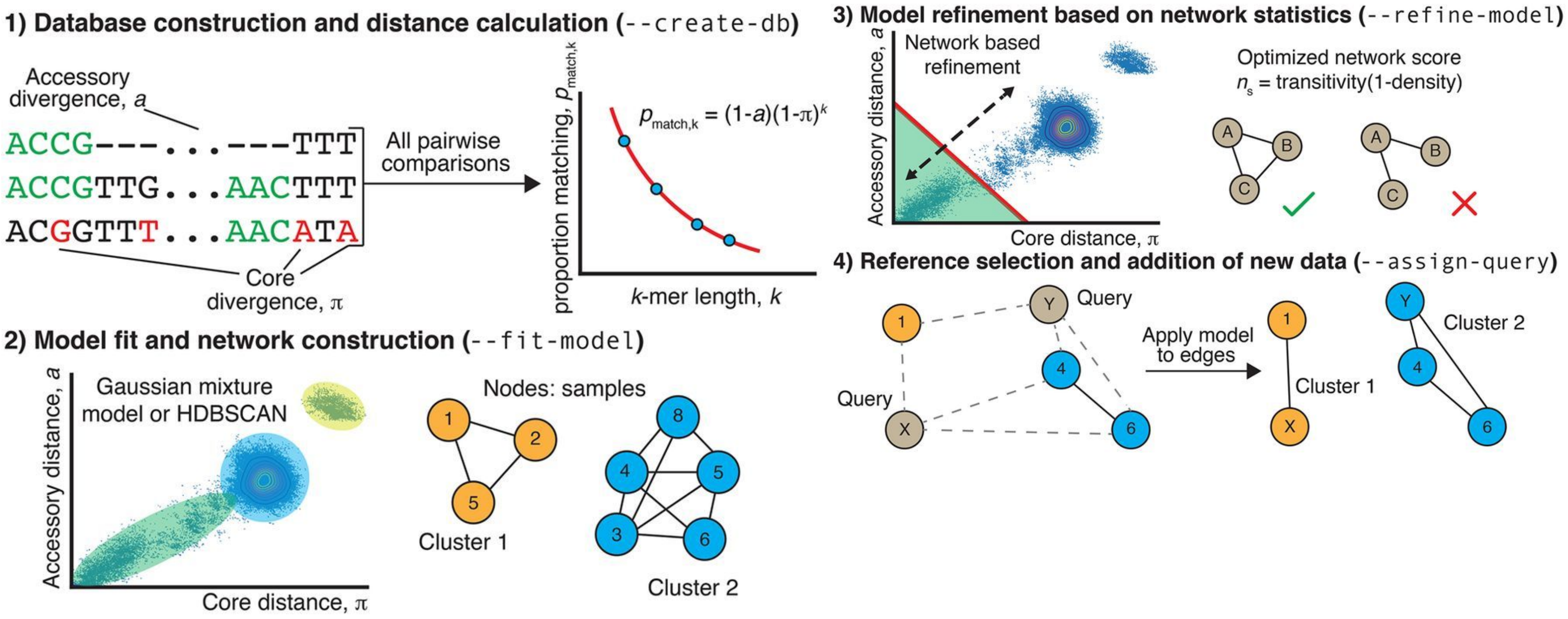
Lees, John A., et al. "Fast and flexible bacterial genomic epidemiology with PopPUNK." Genome research 29.2 (2019): 304-316.

PopPUNK: rapid whole genome comparison & clustering



Lees, John A., et al. "Fast and flexible bacterial genomic epidemiology with PopPUNK." Genome research 29.2 (2019): 304-316.

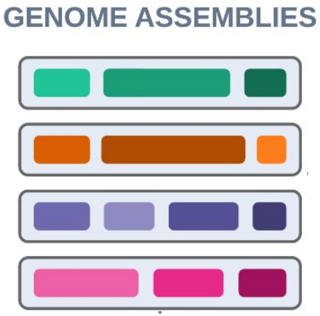
PopPUNK: rapid whole genome comparison & clustering



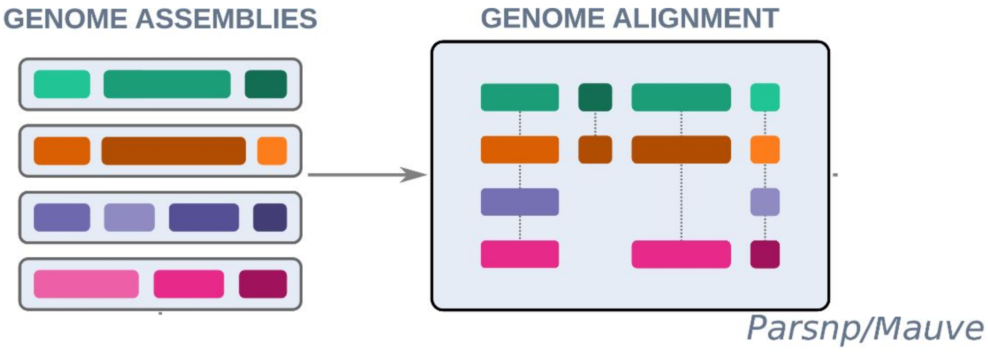
Lees, John A., et al. "Fast and flexible bacterial genomic epidemiology with PopPUNK." Genome research 29.2 (2019): 304-316.

We’ve refined our isolates using typing/clustering, how do we generate a high-resolution outbreak phylogeny?

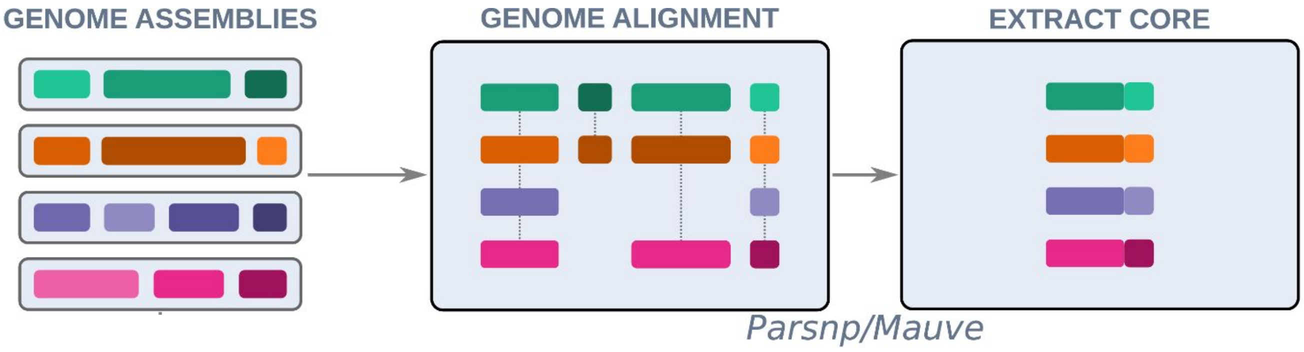
Core genomes can be aligned without a reference genome



Core genomes can be aligned without a reference genome



Core genomes can be aligned without a reference genome



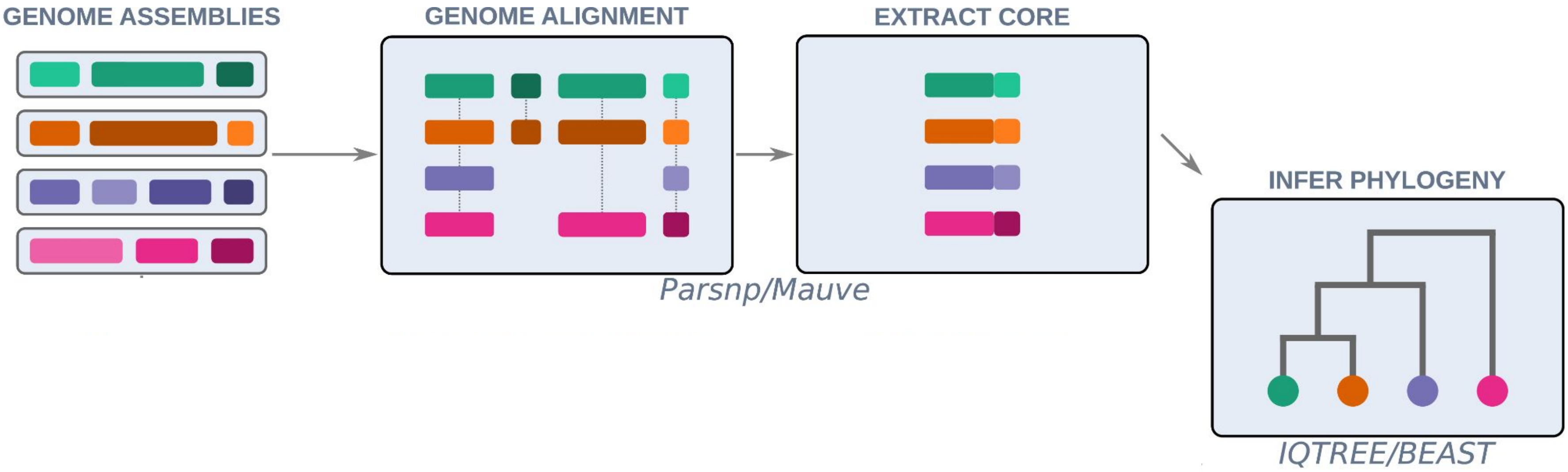
GENOME ASSEMBLIES

GENOME ALIGNMENT

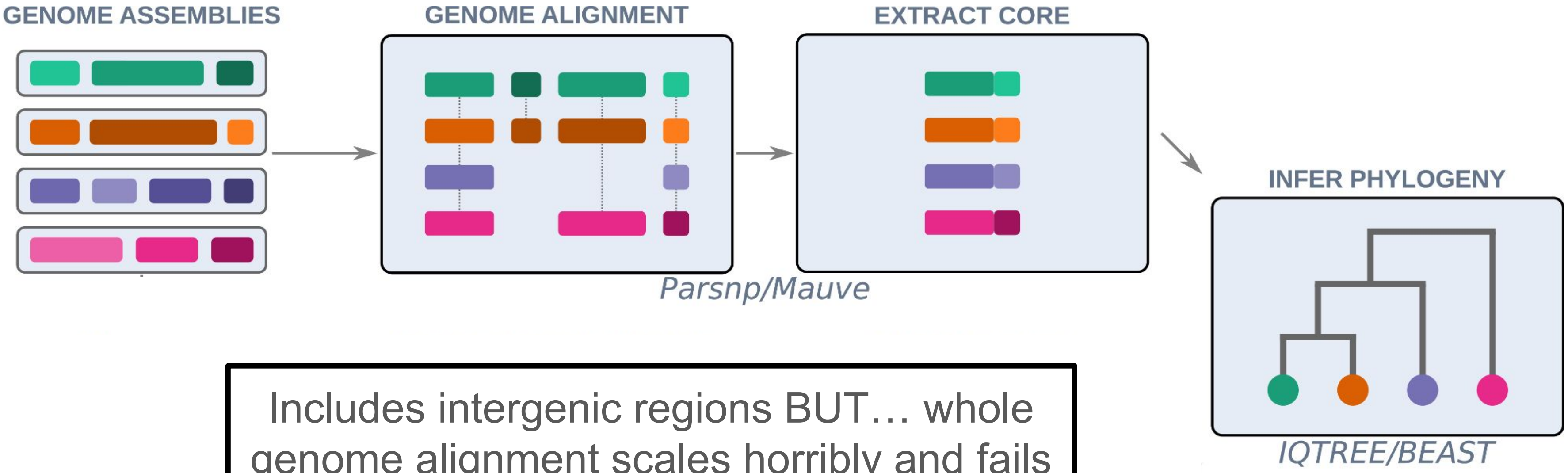
EXTRACT CORE

Parsnp/Mauve

Core genomes can be aligned without a reference genome



Core genomes can be aligned without a reference genome



GENOME ASSEMBLIES

GENOME ALIGNMENT

EXTRACT CORE

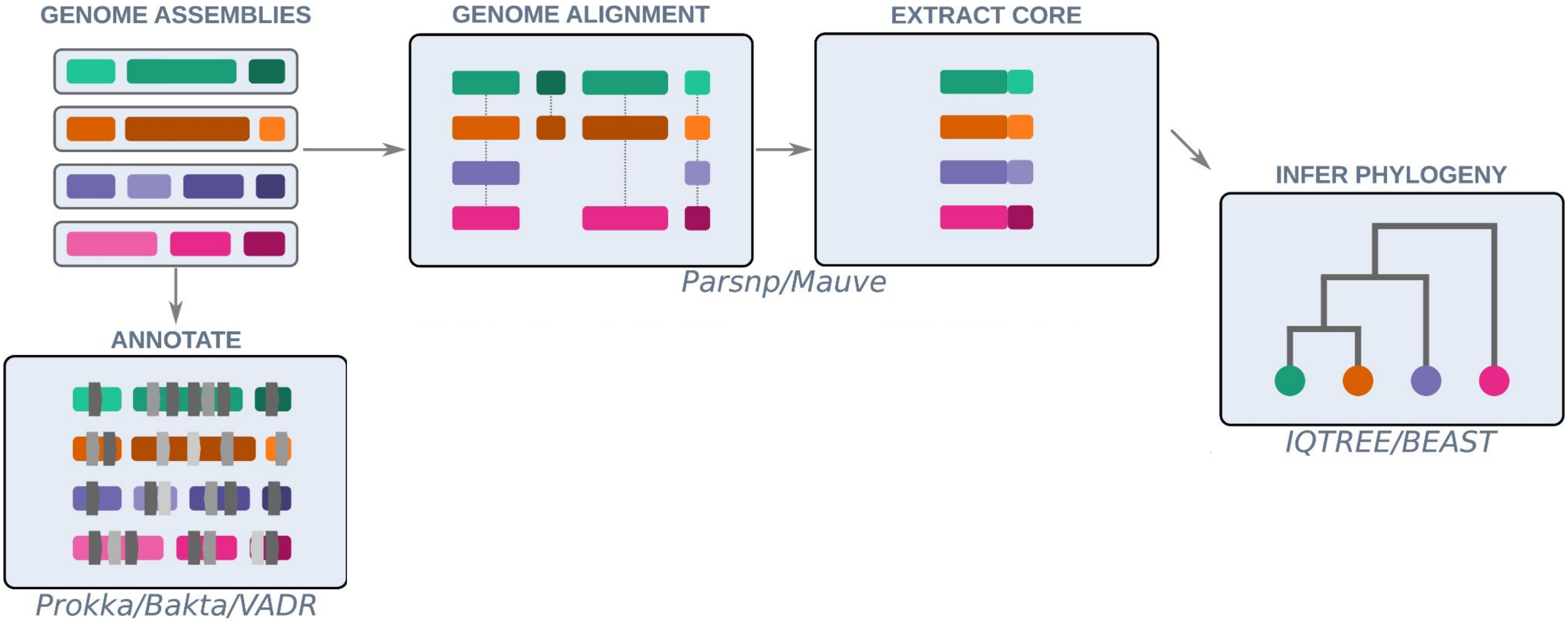
Parsnp/Mauve

INFER PHYLOGENY

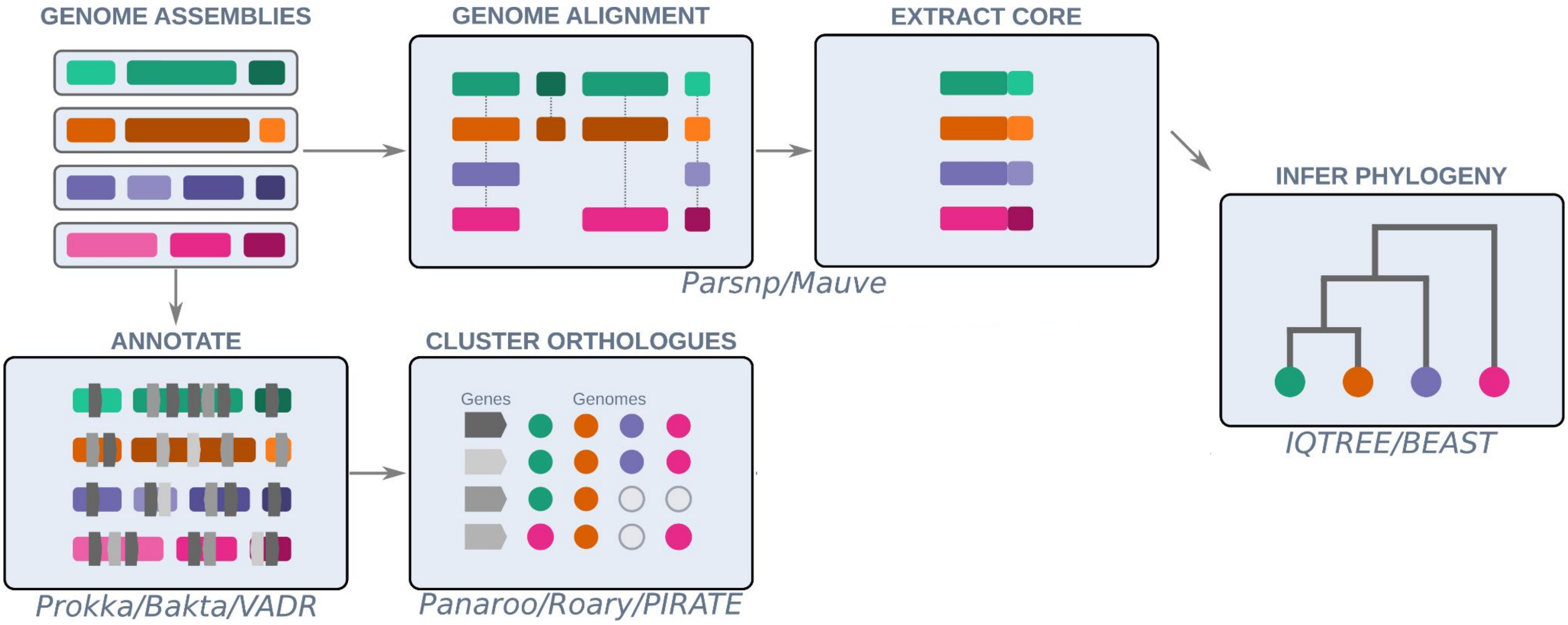
IQTREE/BEAST

Includes intergenic regions BUT… whole genome alignment scales horribly and fails with even moderate recombination and diversity

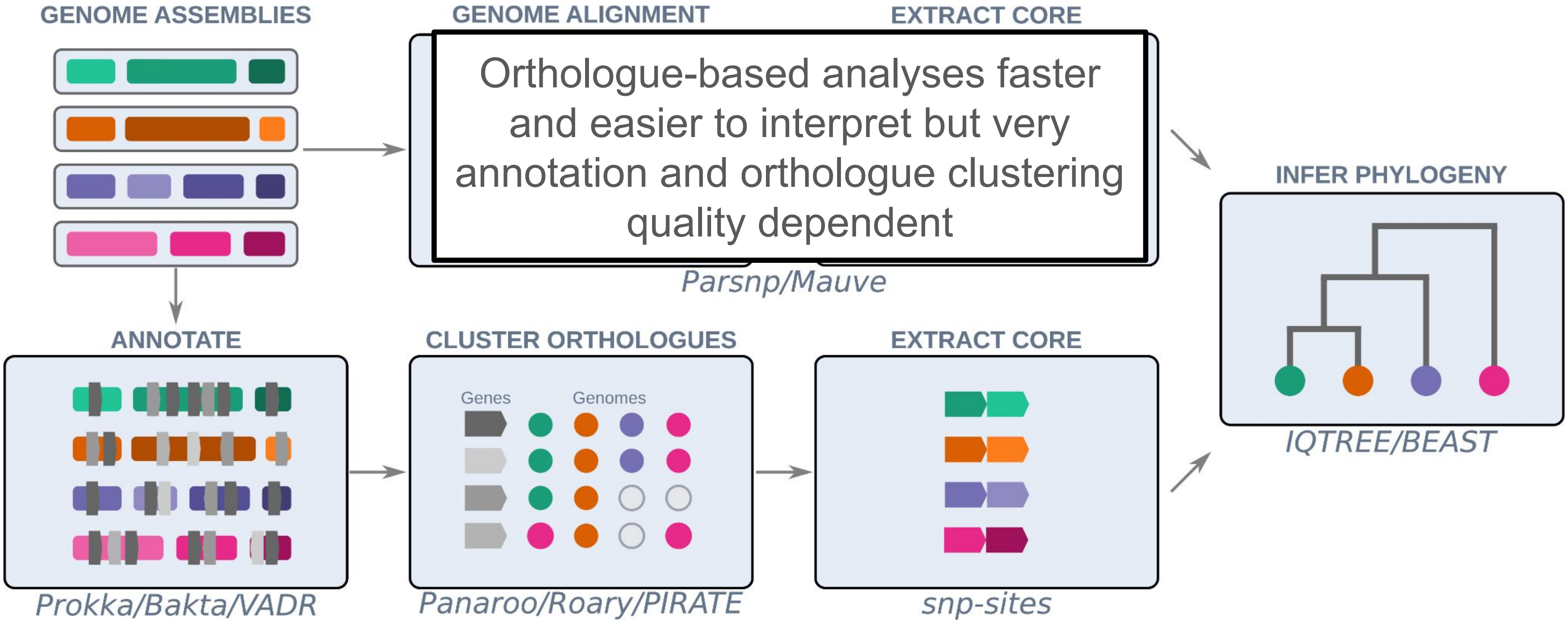
Core genomes can be aligned without a reference genome



Core genomes can be aligned without a reference genome



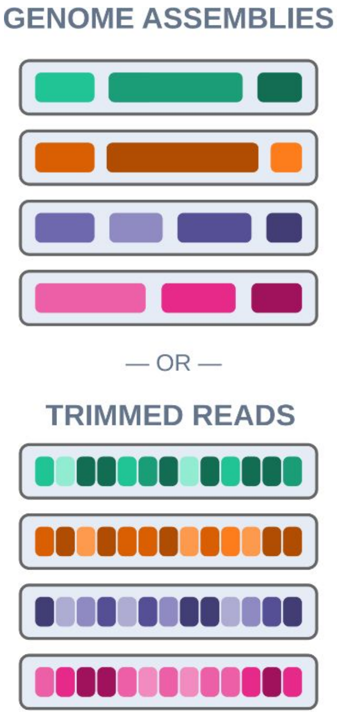
Core genomes can be aligned without a reference genome



Orthologue-based analyses faster and easier to interpret but very annotation and orthologue clustering quality dependent

Can we maximise sensitivity by avoiding *de novo* assembly?

SNP phylogenies are sensitive but reference dependent

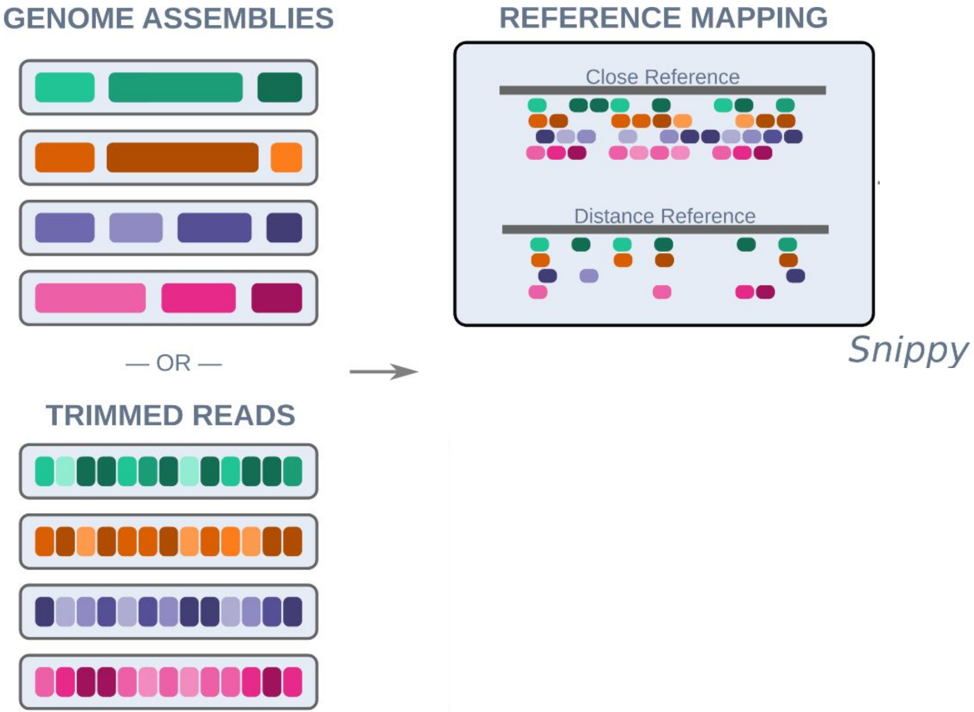


GENOME ASSEMBLIES

— OR —

TRIMMED READS

SNP phylogenies are sensitive but reference dependent



GENOME ASSEMBLIES

— OR —

TRIMMED READS

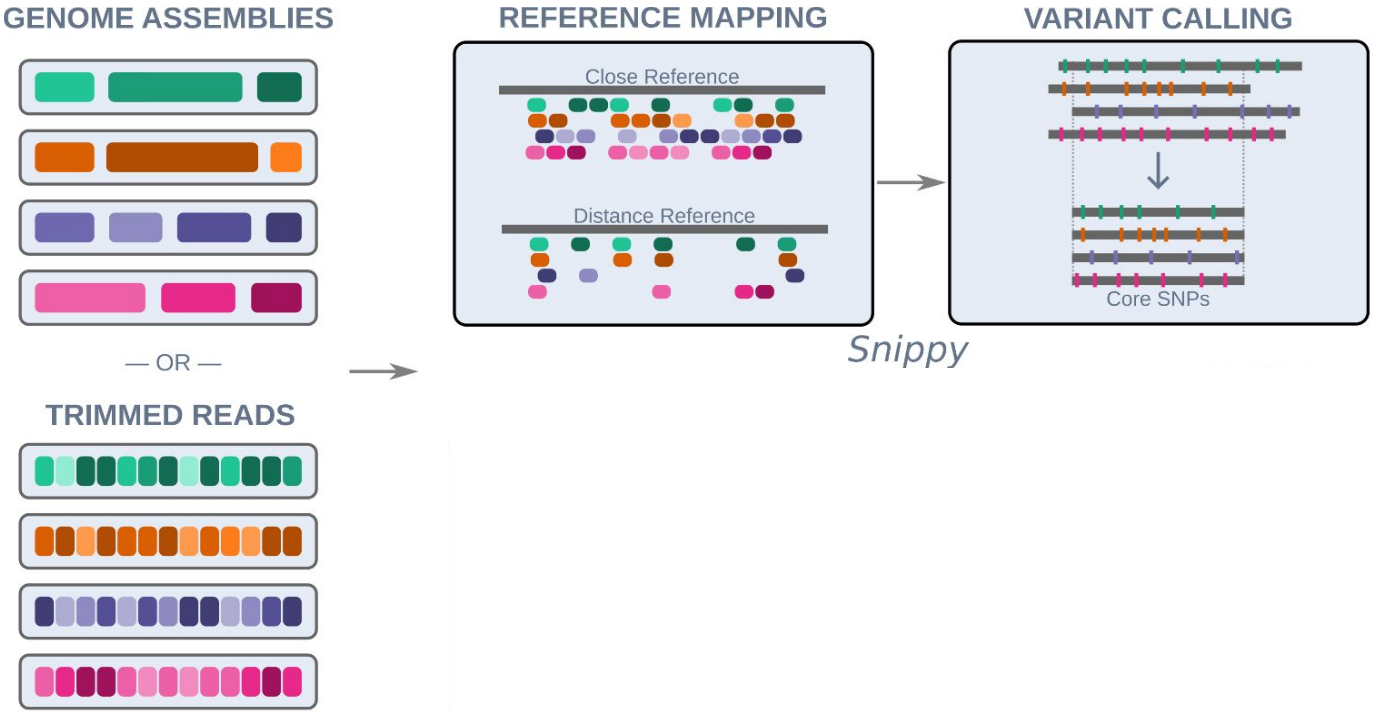
REFERENCE MAPPING

Close Reference

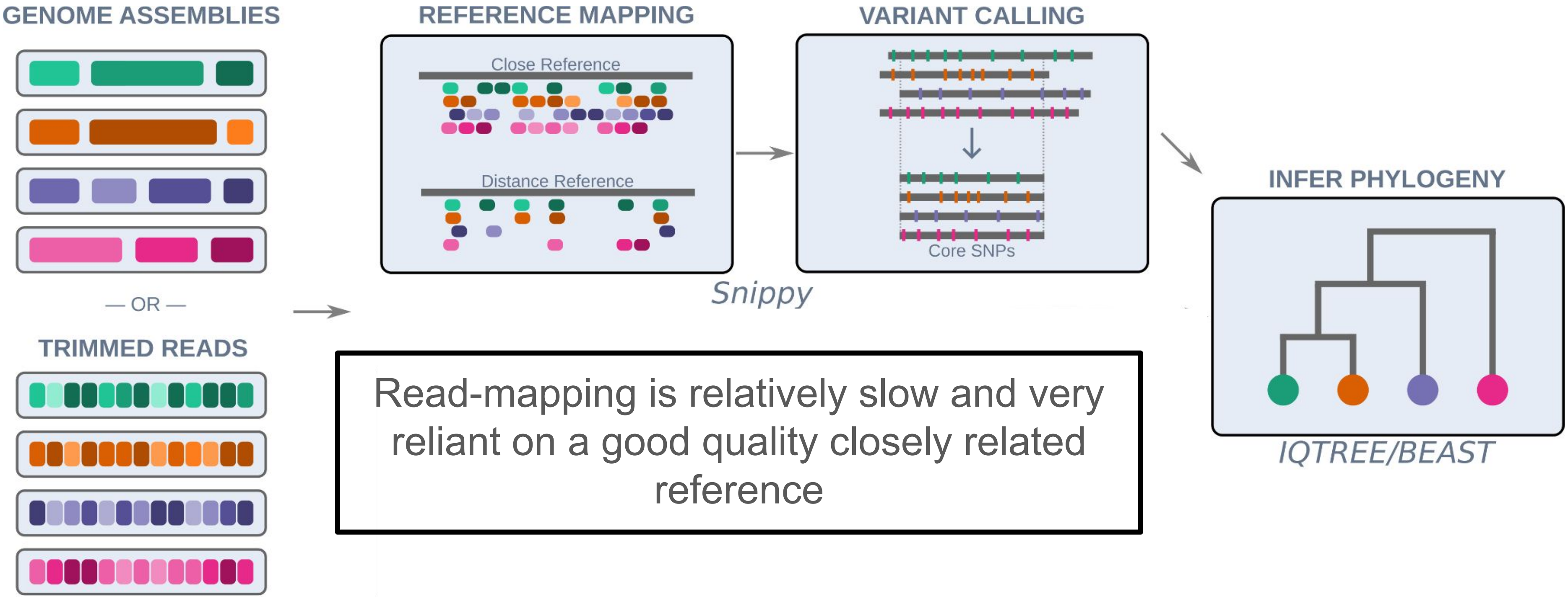
Distance Reference

Snippy

SNP phylogenies are sensitive but reference dependent

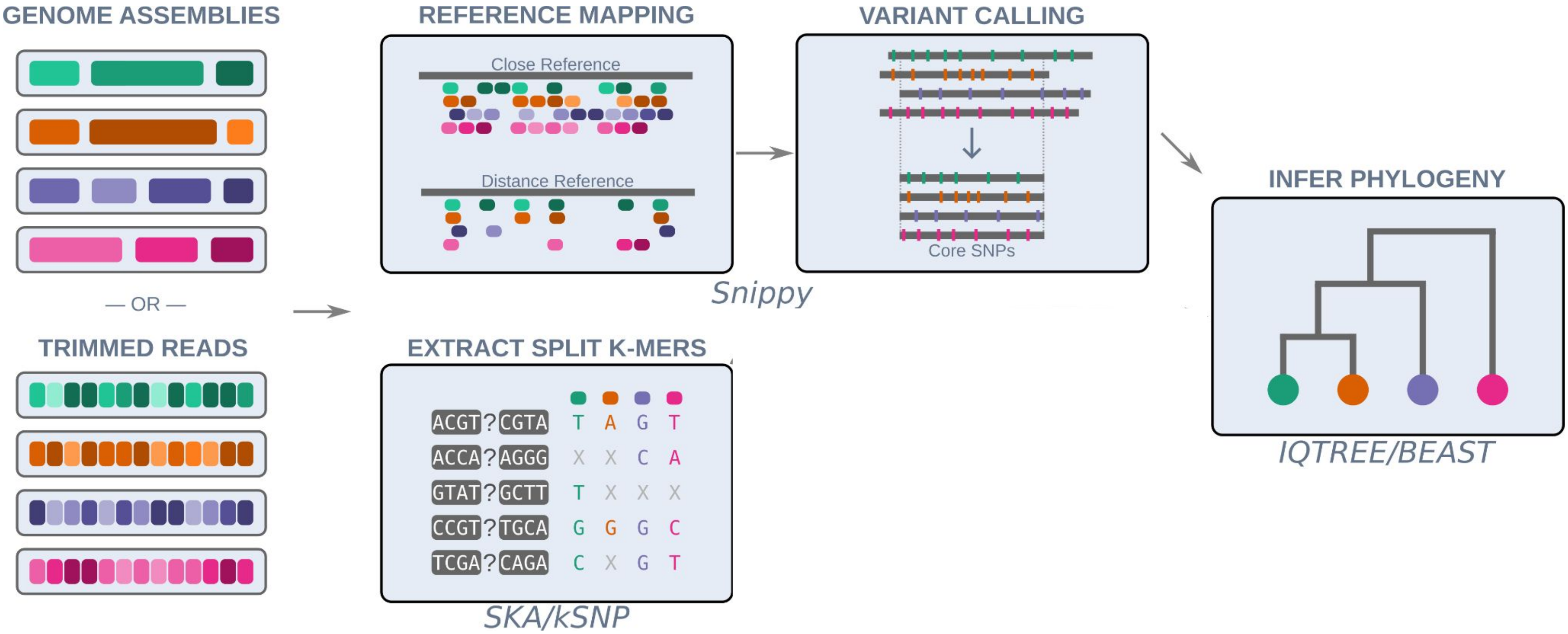


SNP phylogenies are sensitive but reference dependent

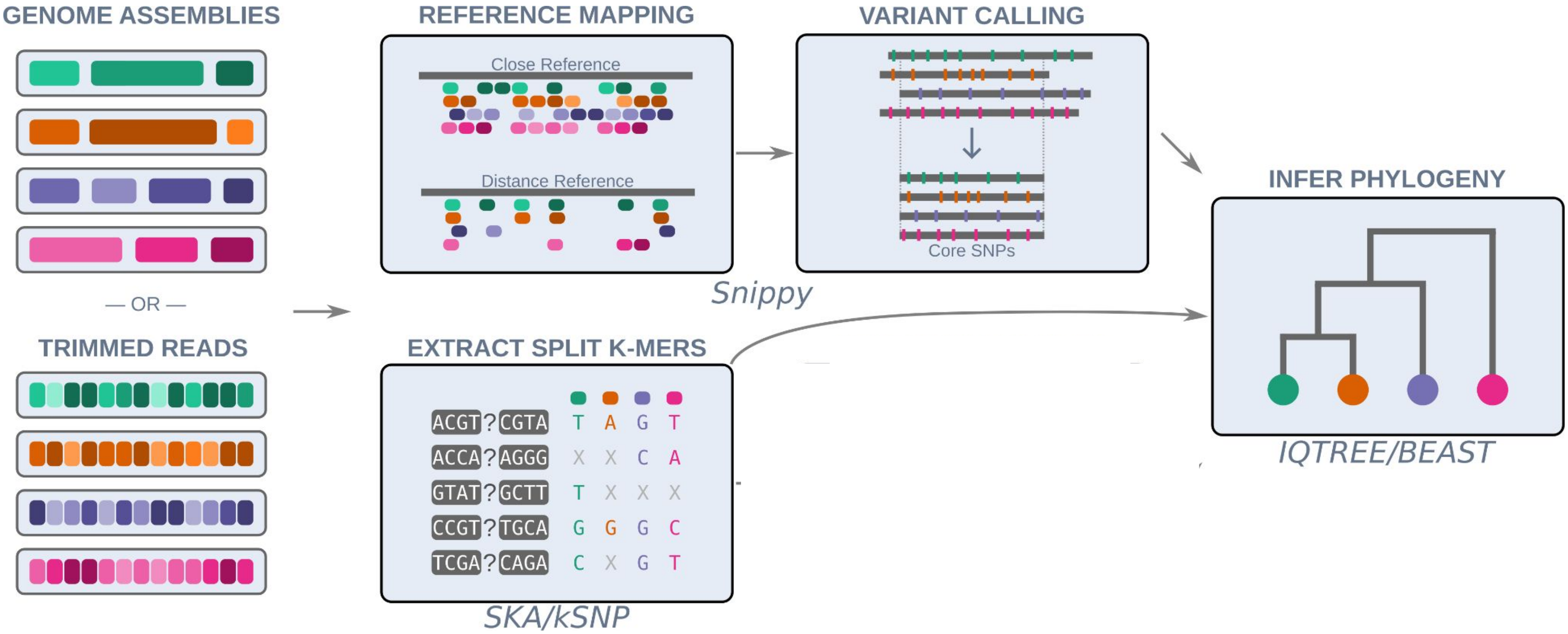


Read-mapping is relatively slow and very reliant on a good quality closely related reference

Split K-mers enables reference-free SNP analyses



Split K-mers enables reference-free SNP analyses



GENOME ASSEMBLIES

— OR —

TRIMMED READS

REFERENCE MAPPING

Close Reference

Distance Reference

VARIANT CALLING

Core SNPs

Snippy

EXTRACT SPLIT K-MERS

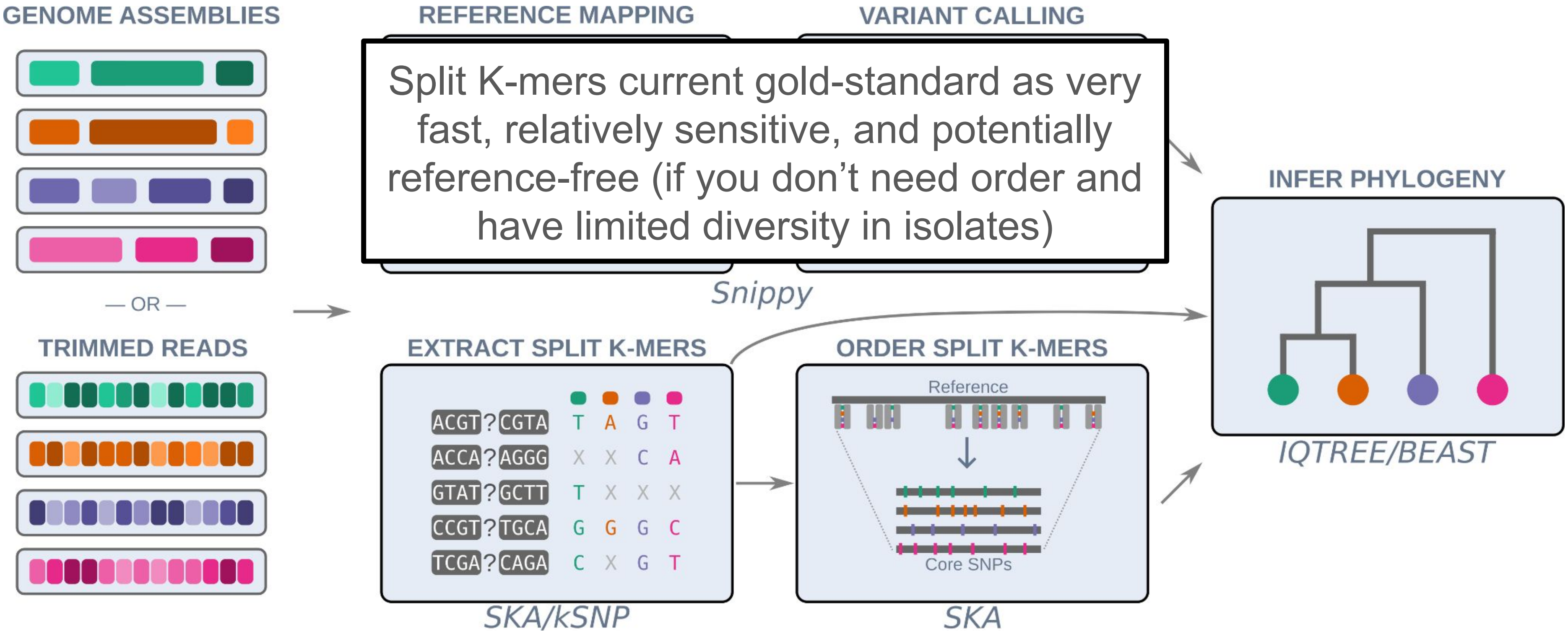
	🟩	🟧	🟪	🟥
ACGT?CGTA	T	A	G	T
ACCA?AGGG	X	X	C	A
GTAT?GCTT	T	X	X	X
CCGT?TGCA	G	G	G	C
TCGA?CAGA	C	X	G	T

SKA/kSNP

INFER PHYLOGENY

IQTREE/BEAST

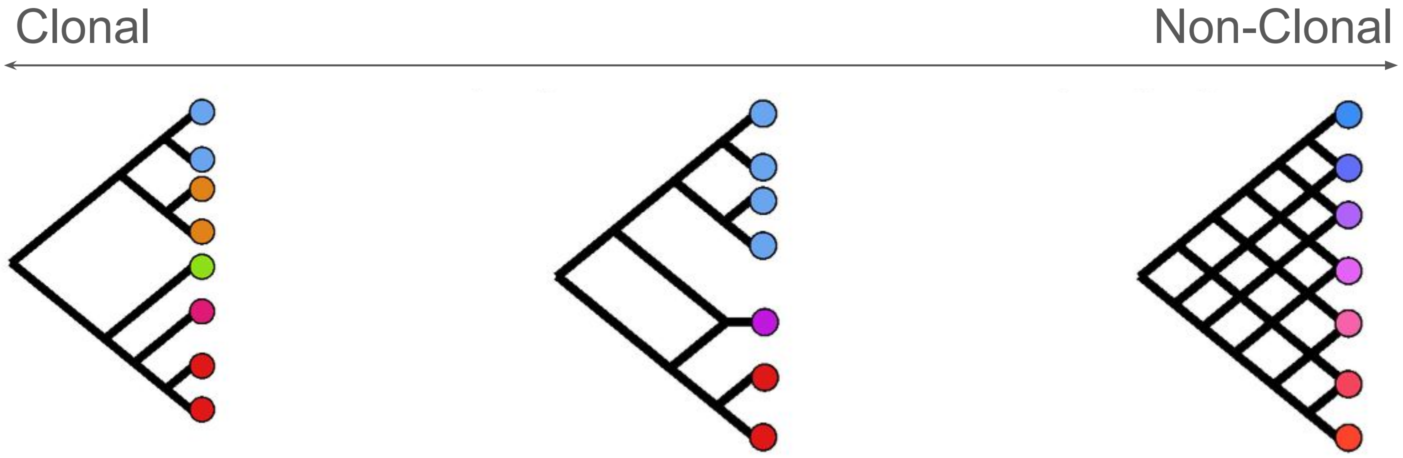
Split K-mers enable reference-free SNP analyses



Split K-mers current gold-standard as very fast, relatively sensitive, and potentially reference-free (if you don't need order and have limited diversity in isolates)

Why not just use SKA without reference-based SNP ordering?

Pathogens exist on a spectrum of tree-like evolution



Clonal

Non-Clonal

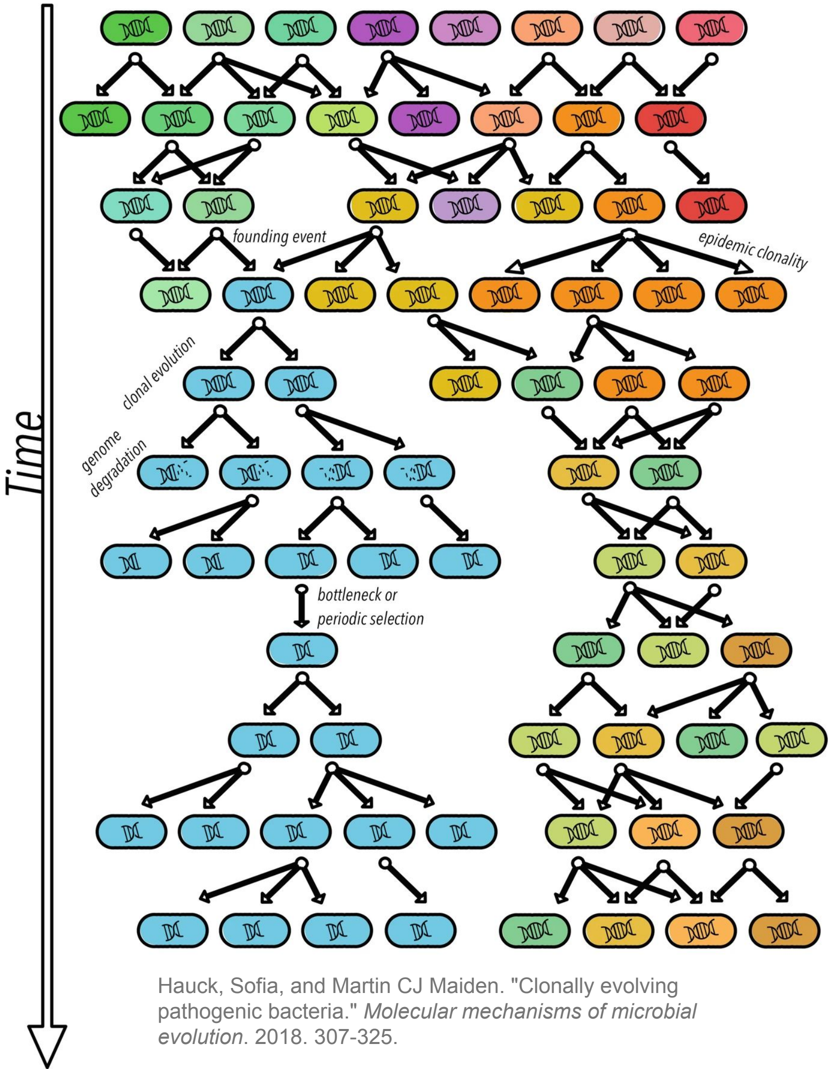
M. tuberculosis
B. anthracis
Y. pestis
Ebolavirus

E. coli
S. aureus
P. aeruginosa
SARS-CoV-2

C. jejuni
H. pylori
B. pseudomallei
HIV, HBV

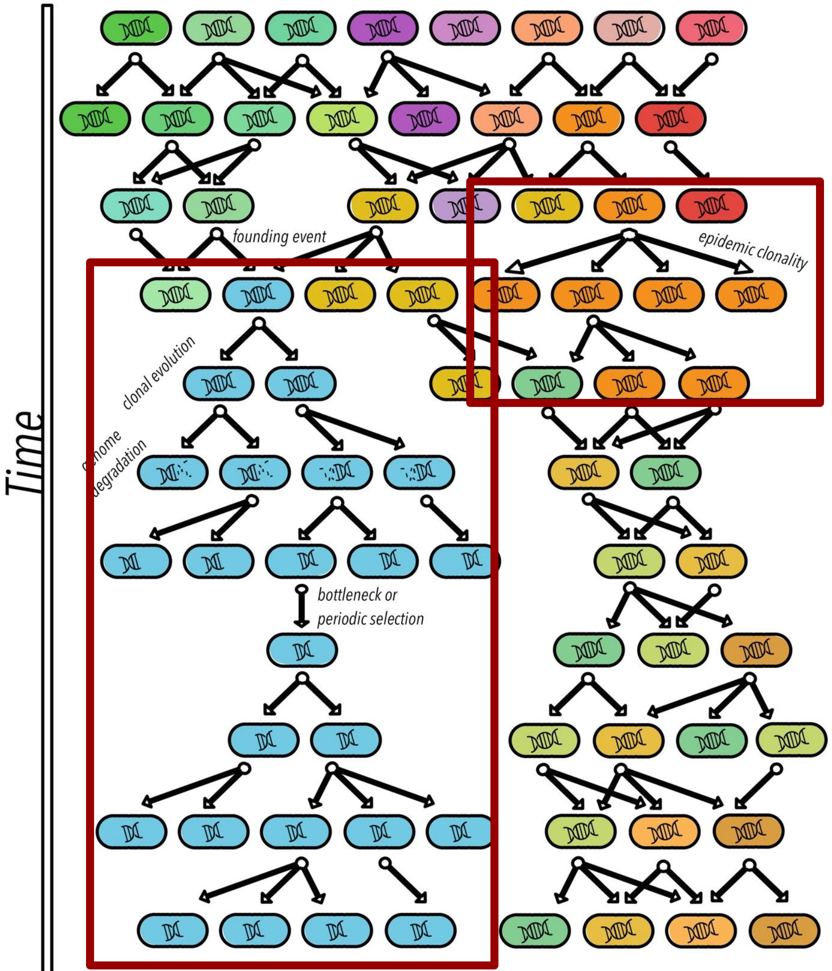
Adapted from: Bacigalupe, Rodrigo. Population genomic analysis of bacterial pathogen niche adaptation. Diss. University of Edinburgh, 2018. Tibayrenc, Michel, and Francisco J. Ayala. "Reproductive clonality of pathogens: a perspective on pathogenic viruses, bacteria, fungi, and parasitic protozoa." Proceedings of the National Academy of Sciences 109.48 (2012): E3305-E3313.

Outbreaks often clonal (mostly)



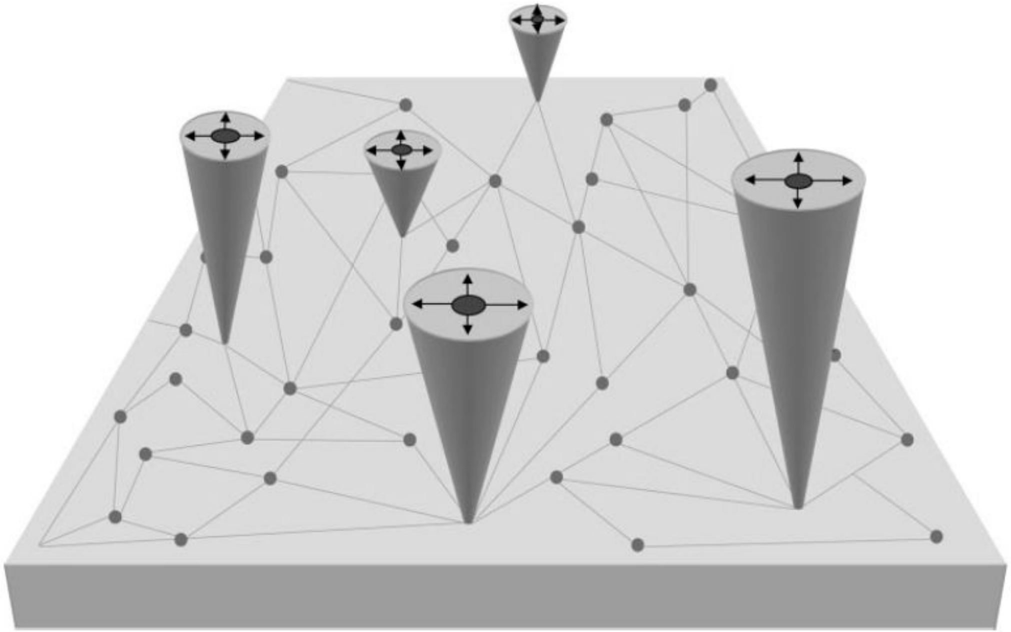
Hauck, Sofia, and Martin CJ Maiden. "Clonally evolving pathogenic bacteria." *Molecular mechanisms of microbial evolution*. 2018. 307-325.

Outbreaks often clonal (mostly)

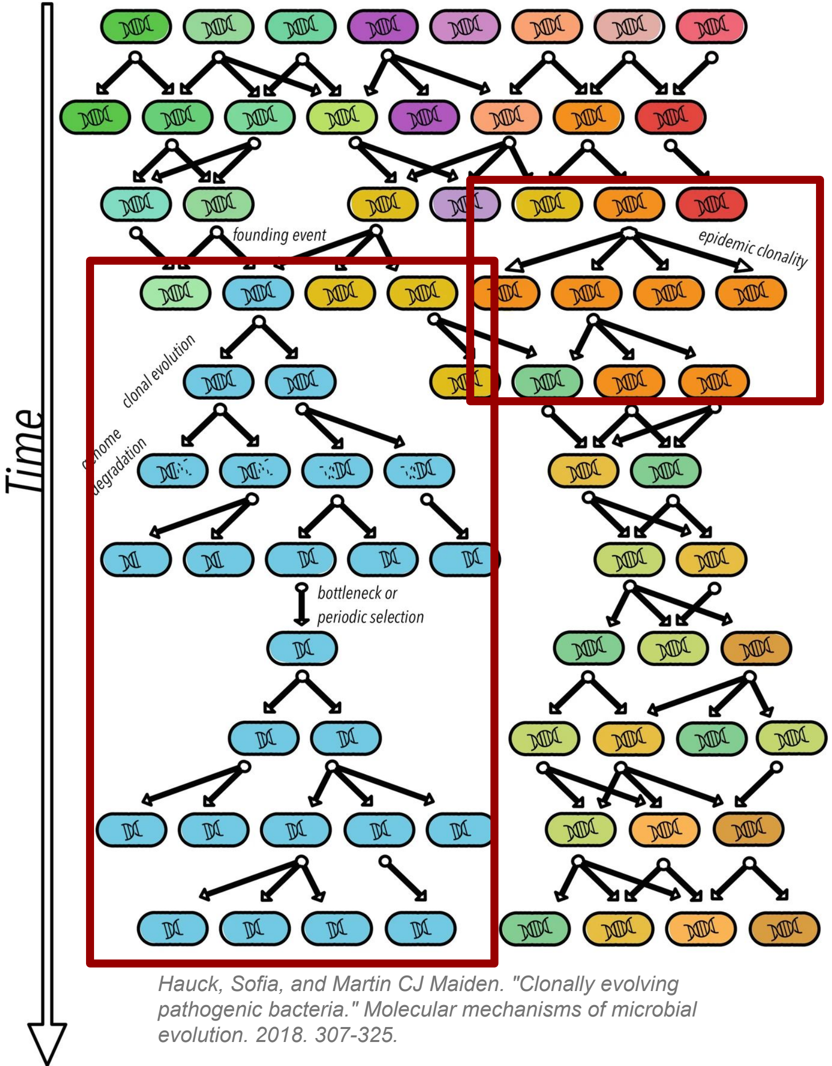


Hauck, Sofia, and Martin CJ Maiden. "Clonally evolving pathogenic bacteria." *Molecular mechanisms of microbial evolution*. 2018. 307-325.

Outbreaks often clonal (mostly)

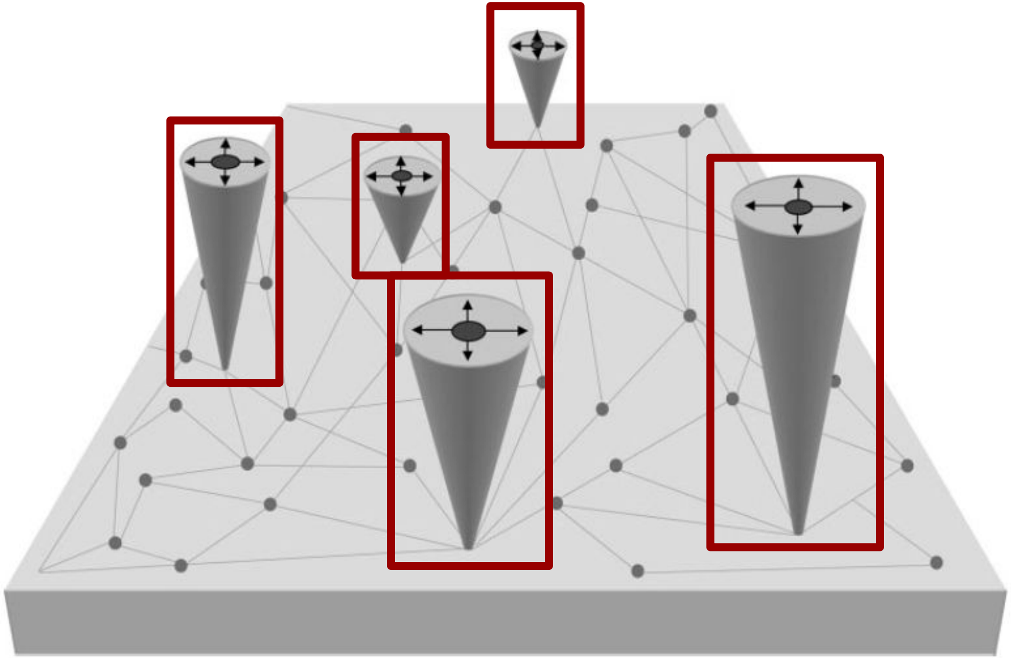


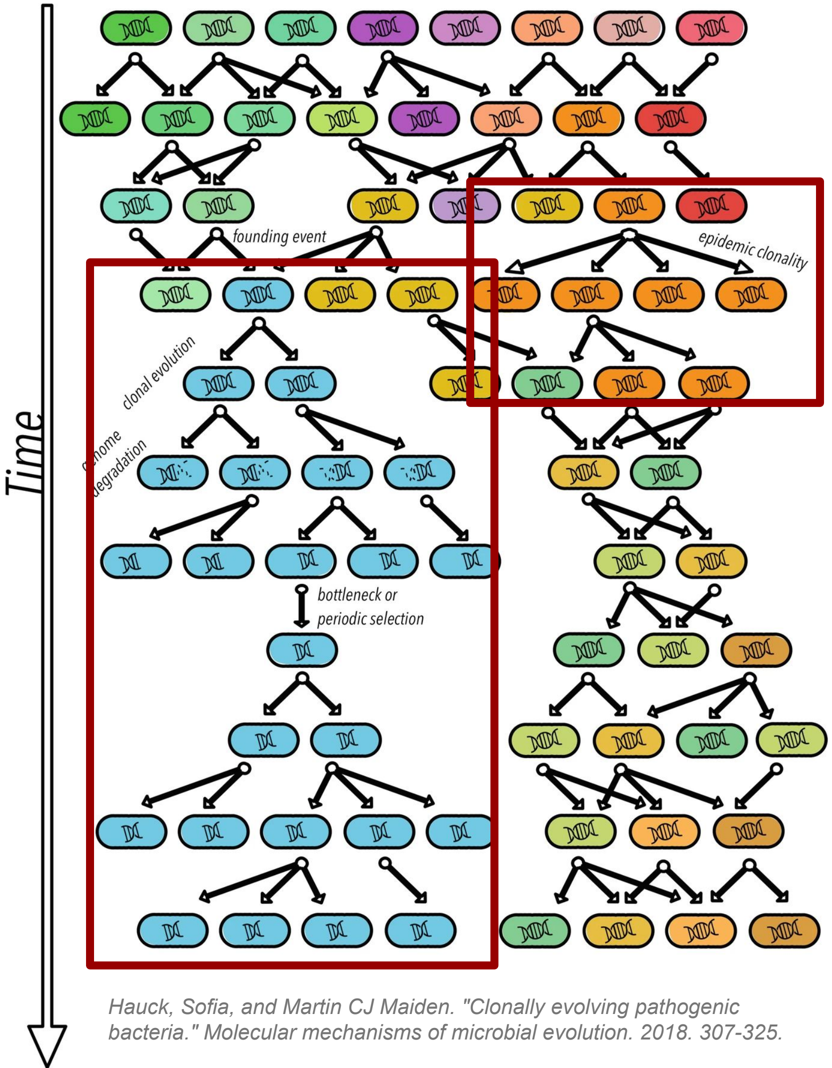
Feil, Edward J., and Brian G. Spratt. "Recombination and the population structures of bacterial pathogens." Annual Reviews in Microbiology 55.1 (2001): 561-590.



Hauck, Sofia, and Martin CJ Maiden. "Clonally evolving pathogenic bacteria." Molecular mechanisms of microbial evolution. 2018. 307-325.

Outbreaks often clonal (mostly)



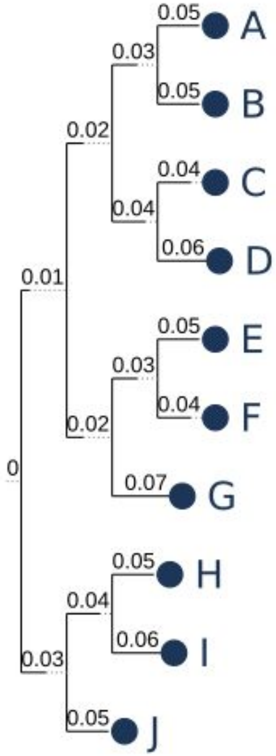


But recombination still commonly encountered

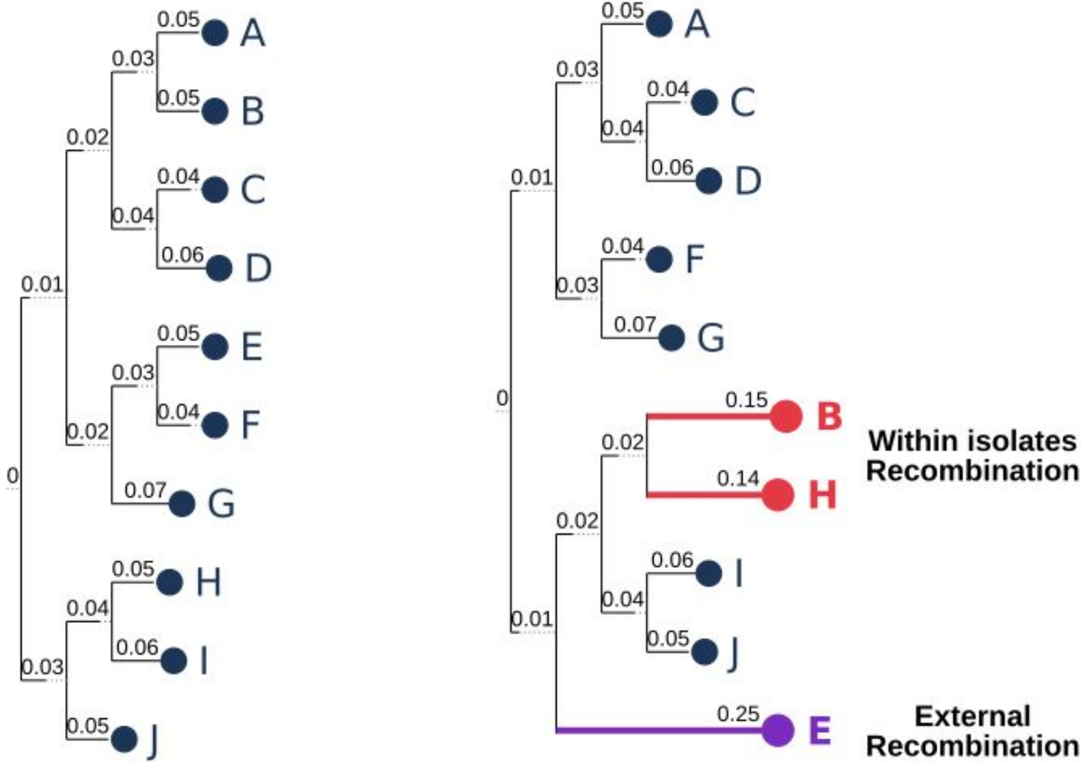
Feil, Edward J., and Brian G. Spratt. "Recombination and the population structures of bacterial pathogens." Annual Reviews in Microbiology 55.1 (2001): 561-590.

Hauck, Sofia, and Martin CJ Maiden. "Clonally evolving pathogenic bacteria." Molecular mechanisms of microbial evolution. 2018. 307-325.

Recombination distorts phylogeny and temporal signal



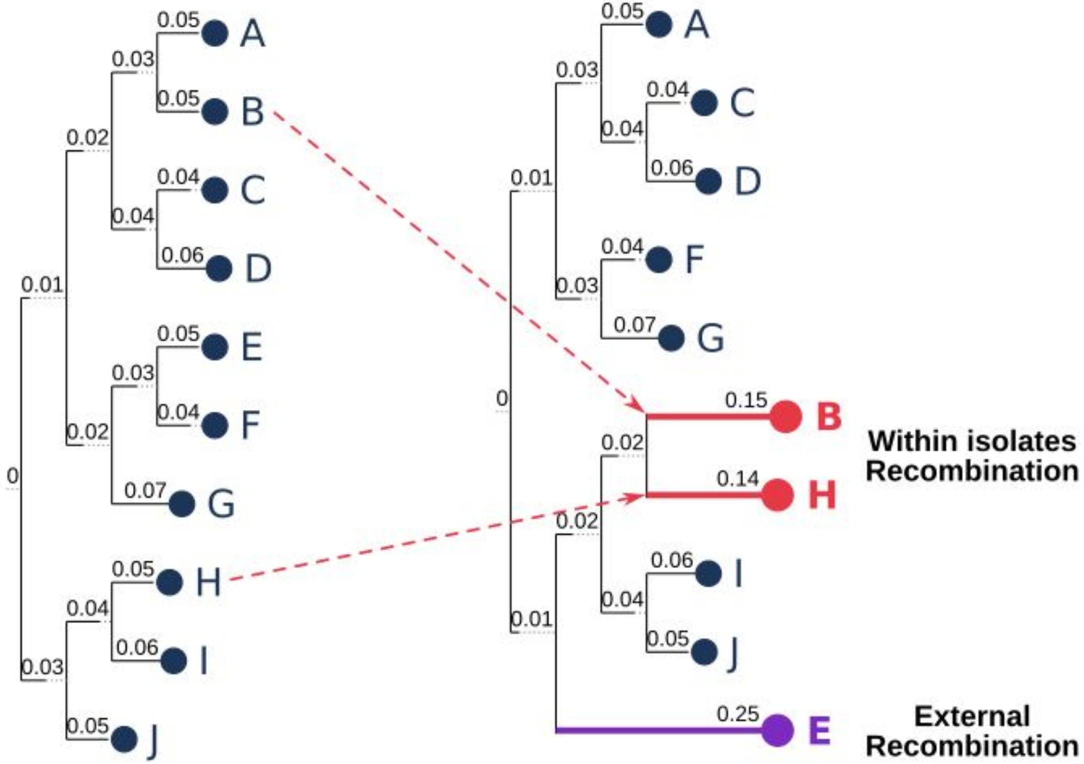
Recombination distorts phylogeny and temporal signal



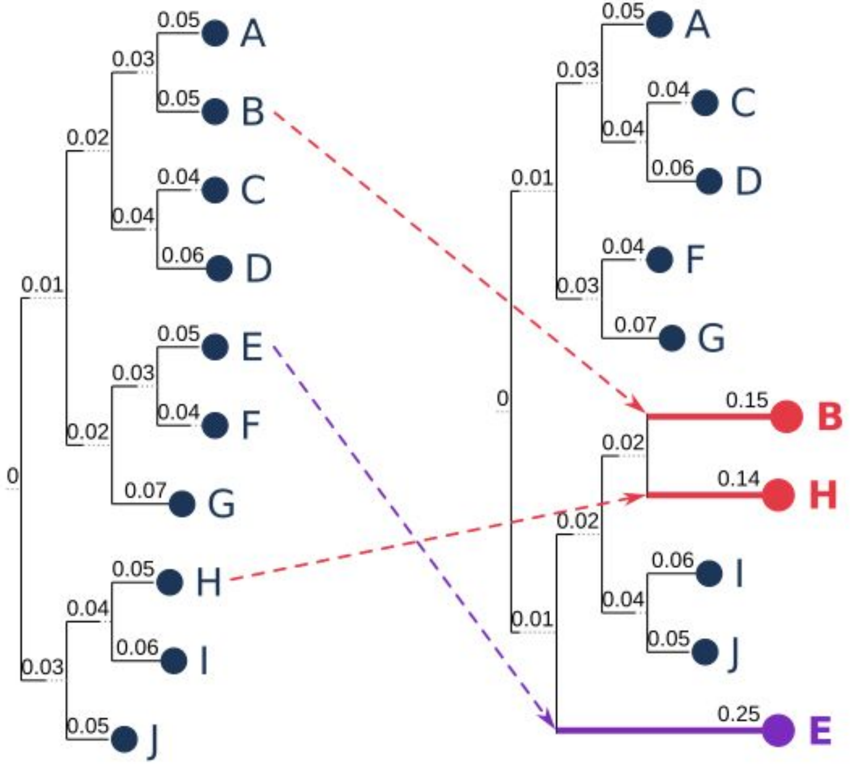
Within isolates Recombination

External Recombination

Recombination distorts phylogeny and temporal signal



Recombination distorts phylogeny and temporal signal

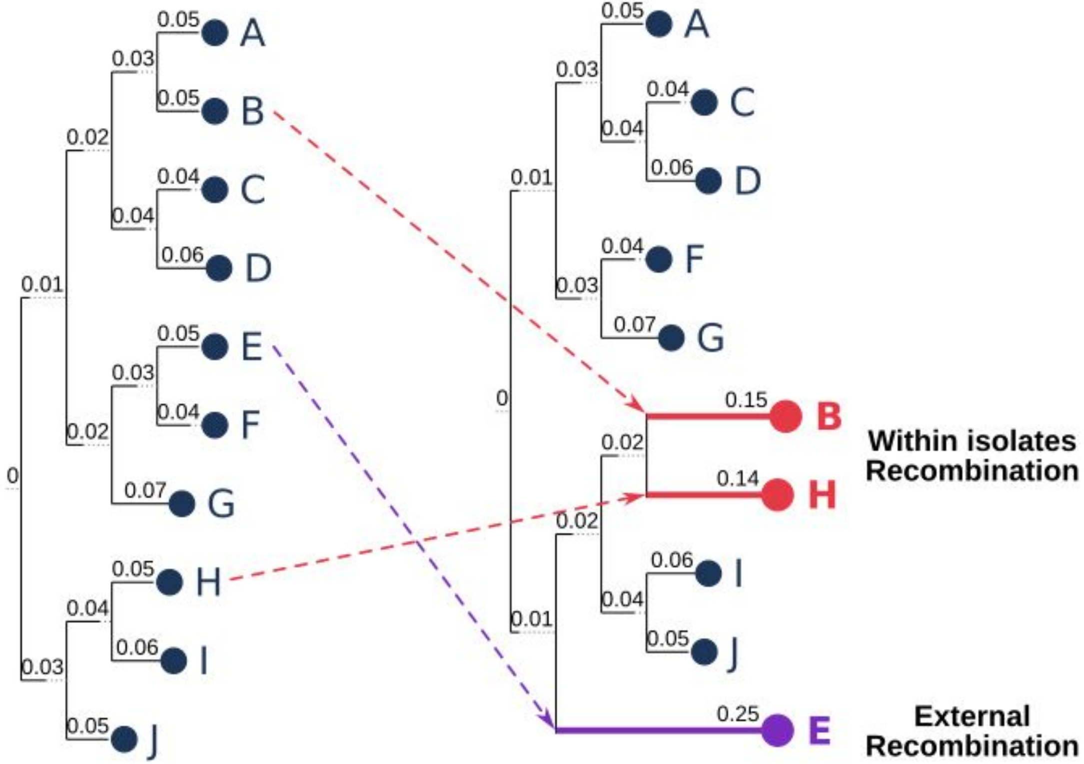


Within isolates Recombination

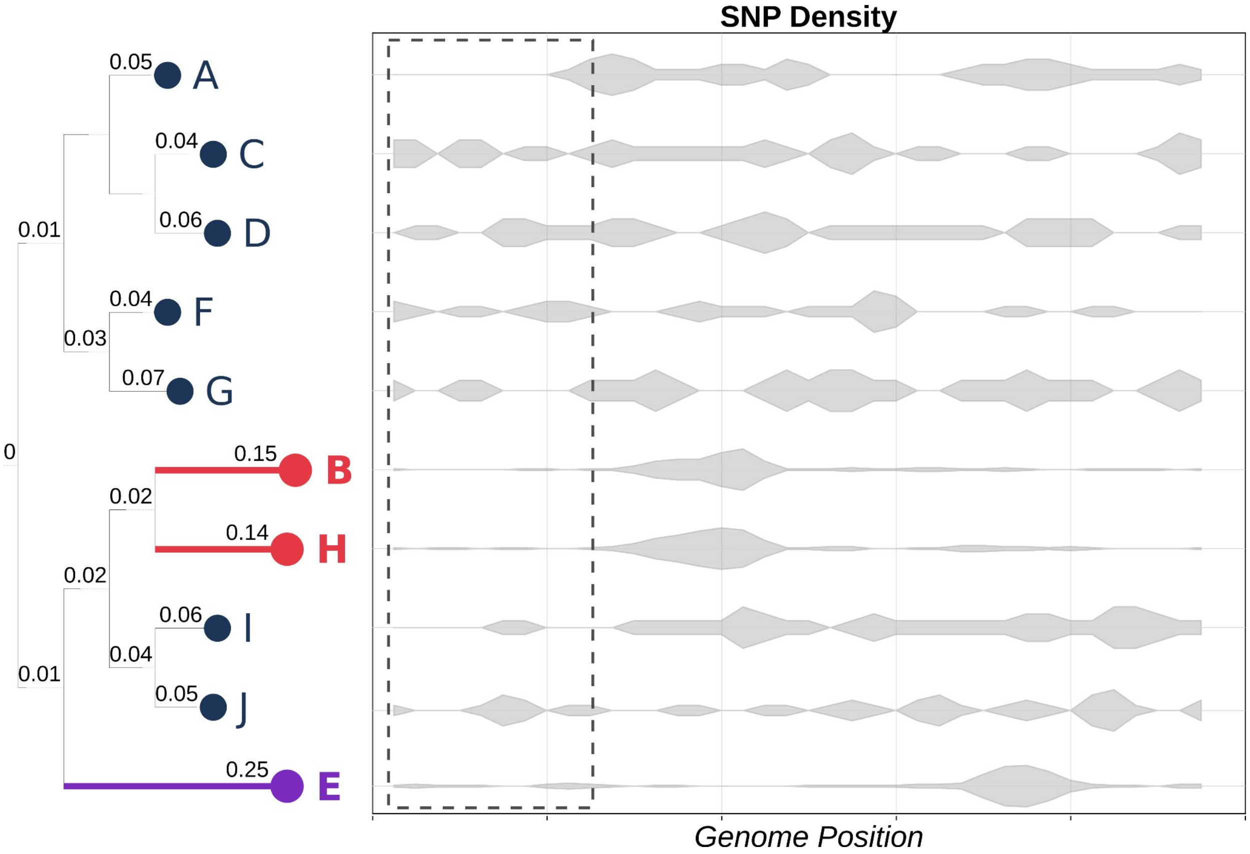
External Recombination

Recombination distorts phylogeny and temporal signal

Accurate branch lengths important for temporal inference

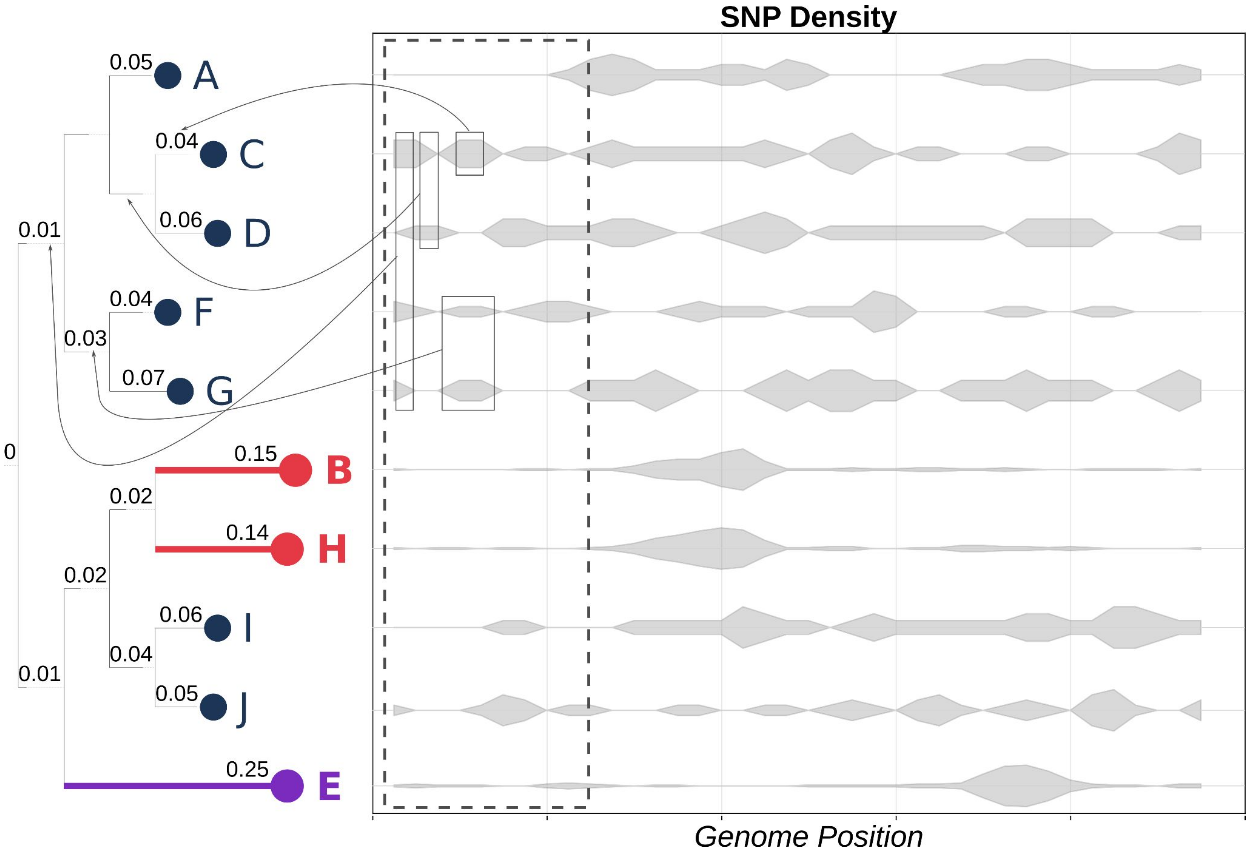


Mask recombination by finding elevated mutation density

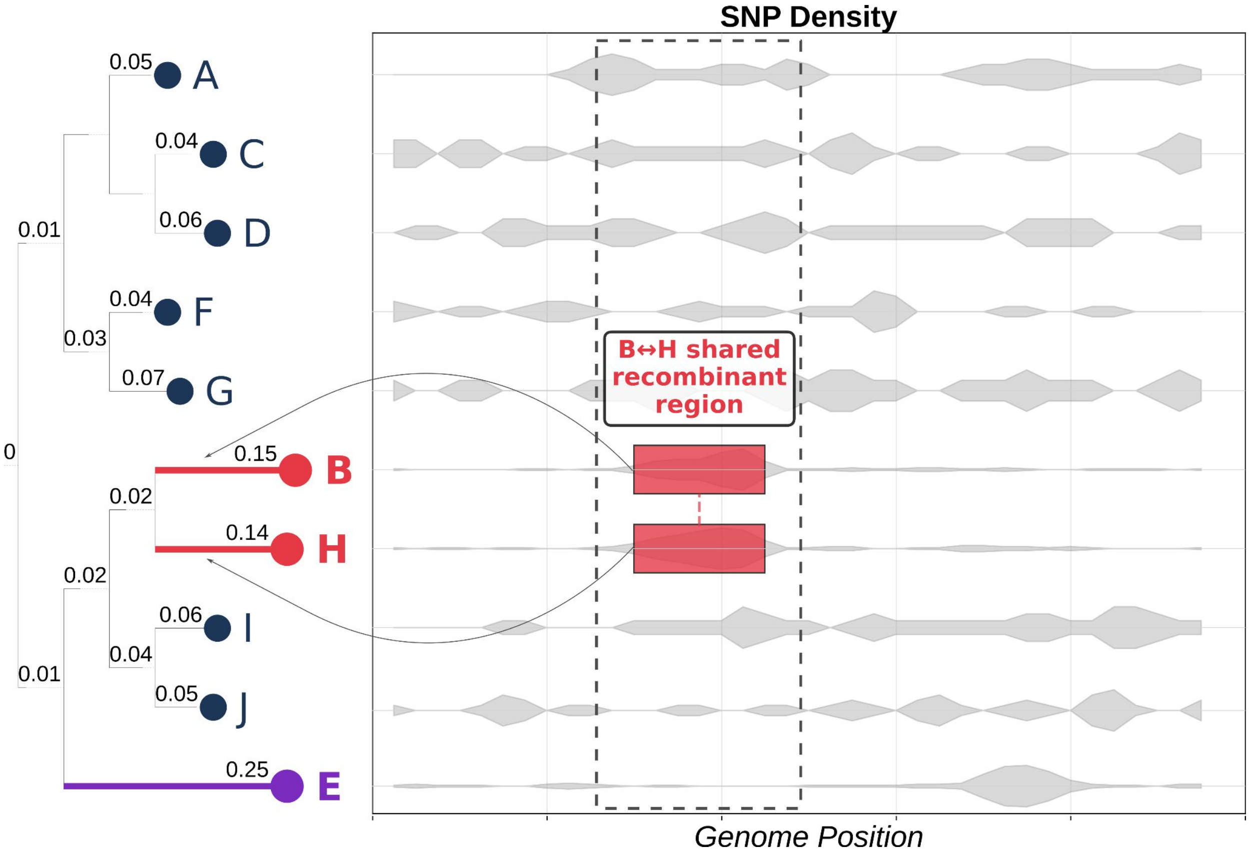


Requires SNP positions => reference-based SKA analysis

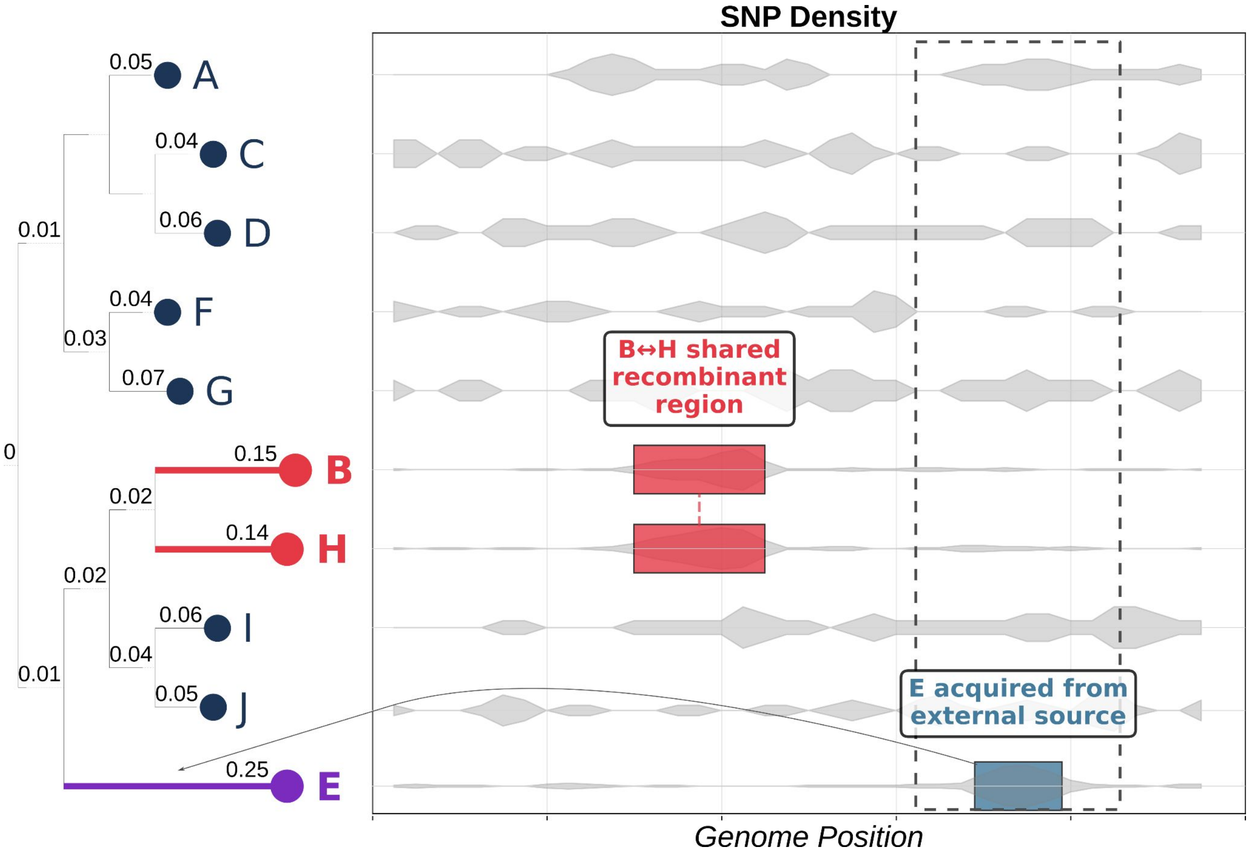
Mask recombination by finding elevated mutation density



Mask recombination by finding elevated mutation density



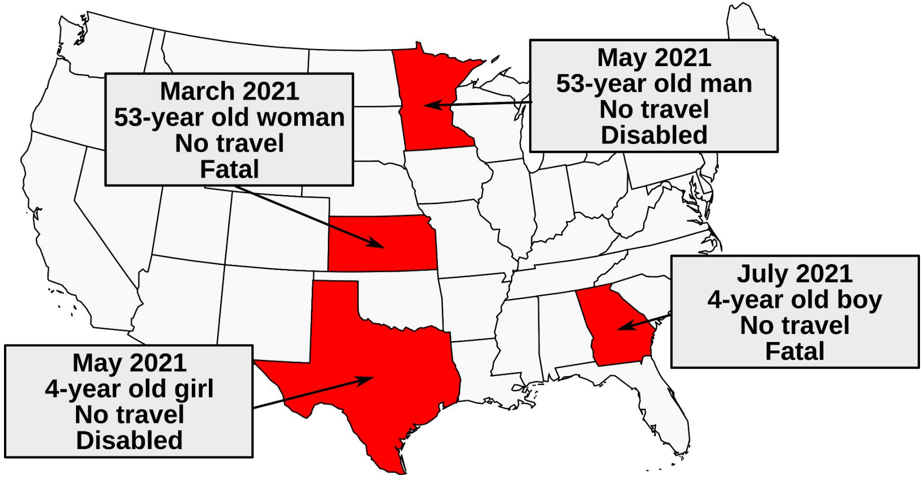
Mask recombination by finding elevated mutation density



So, how are refined phylogenies useful?

Phylogenies provide vital information in outbreaks

- *Burkholderia pseudomallei* infections **4 people** in **4 states** over **4 months**
- Genomics showed **same strain** of *B. pseudomallei* (linked to South Asia)
- **No travel history**/common behaviour

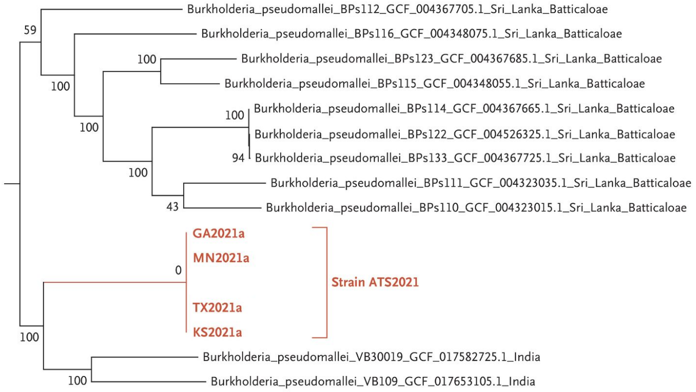


March 2021
53-year old woman
No travel
Fatal

May 2021
53-year old man
No travel
Disabled

May 2021
4-year old girl
No travel
Disabled

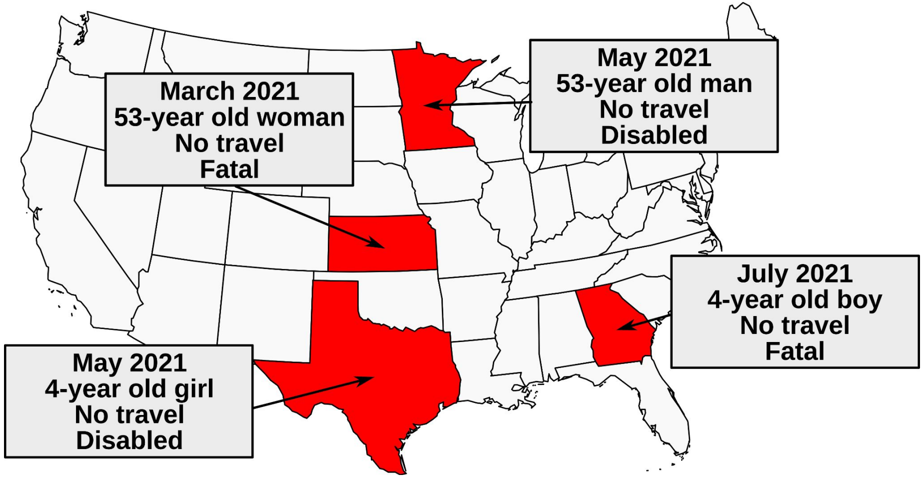
July 2021
4-year old boy
No travel
Fatal



Gee, Jay E., et al. "Multistate outbreak of melioidosis associated with imported aromatherapy spray." *New England Journal of Medicine* 386.9 (2022): 861-868.

Phylogenies provide vital information in outbreaks

- *Burkholderia pseudomallei* infections **4 people** in **4 states** over **4 months**
- Genomics showed **same strain** of *B. pseudomallei* (linked to South Asia)
- **No travel history**/common behaviour
- **>200 items** tested as potential topical/ ingestion/inhaled source (aerosols, food, cleaning & beauty products, medications, soil/water)

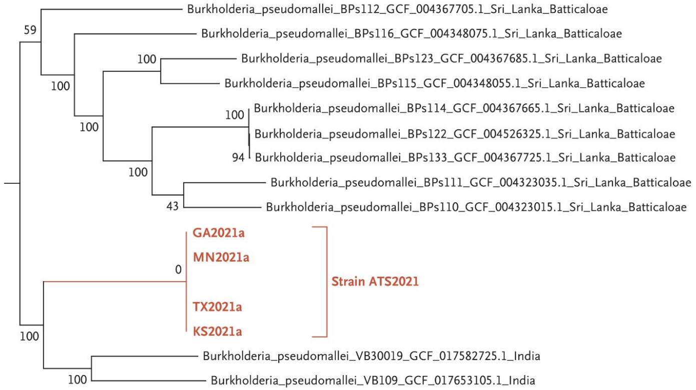


March 2021
53-year old woman
No travel
Fatal

May 2021
53-year old man
No travel
Disabled

May 2021
4-year old girl
No travel
Disabled

July 2021
4-year old boy
No travel
Fatal

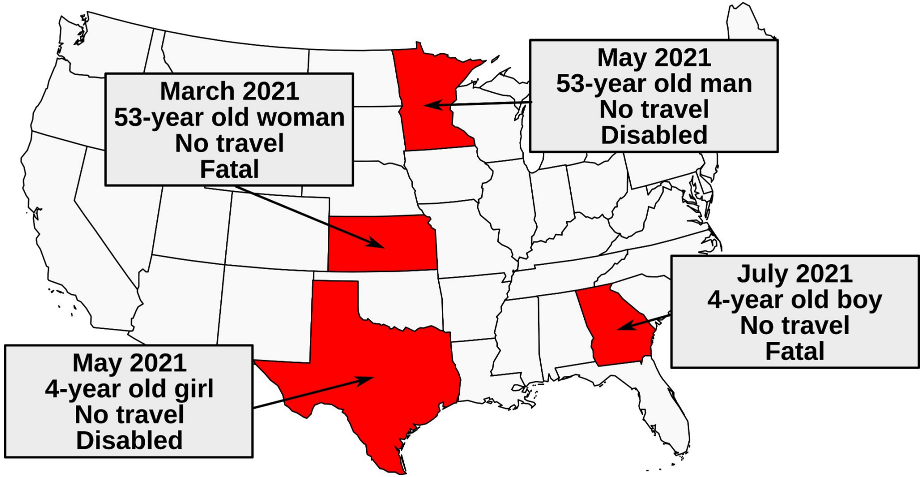


Gee, Jay E., et al. "Multistate outbreak of melioidosis associated with imported aromatherapy spray." *New England Journal of Medicine* 386.9 (2022): 861-868.

Phylogenies provide vital information in outbreaks

- *Burkholderia pseudomallei* infections **4 people** in **4 states** over **4 months**
- Genomics showed **same strain** of *B. pseudomallei* (linked to South Asia)
- **No travel history**/common behaviour
- **>200 items** tested as potential topical/ ingestion/inhaled source (aerosols, food, cleaning & beauty products, medications, soil/water)

- *Essential Oil and Semi-precious Stone Infused Aromatherapy Room Spray* positive for **genomically identical** *B. pseudomallei*

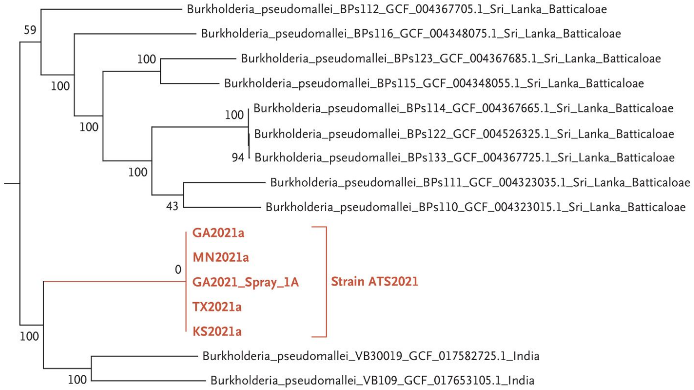


March 2021
53-year old woman
No travel
Fatal

May 2021
53-year old man
No travel
Disabled

July 2021
4-year old boy
No travel
Fatal

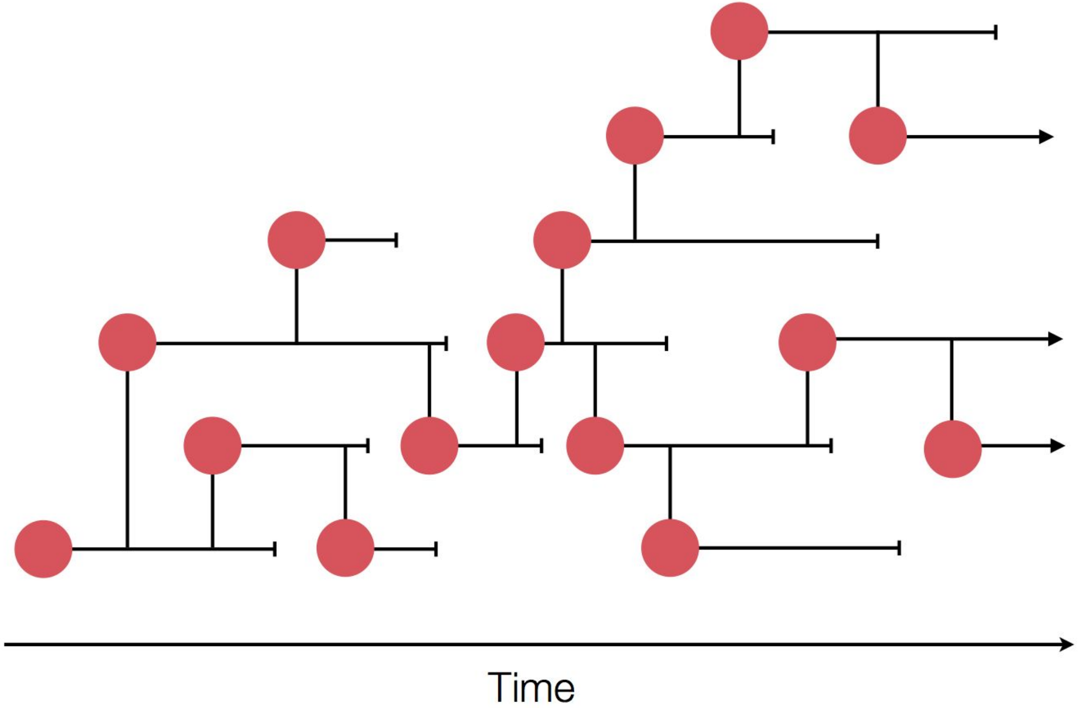
May 2021
4-year old girl
No travel
Disabled



Gee, Jay E., et al. "Multistate outbreak of melioidosis associated with imported aromatherapy spray." *New England Journal of Medicine* 386.9 (2022): 861-868.

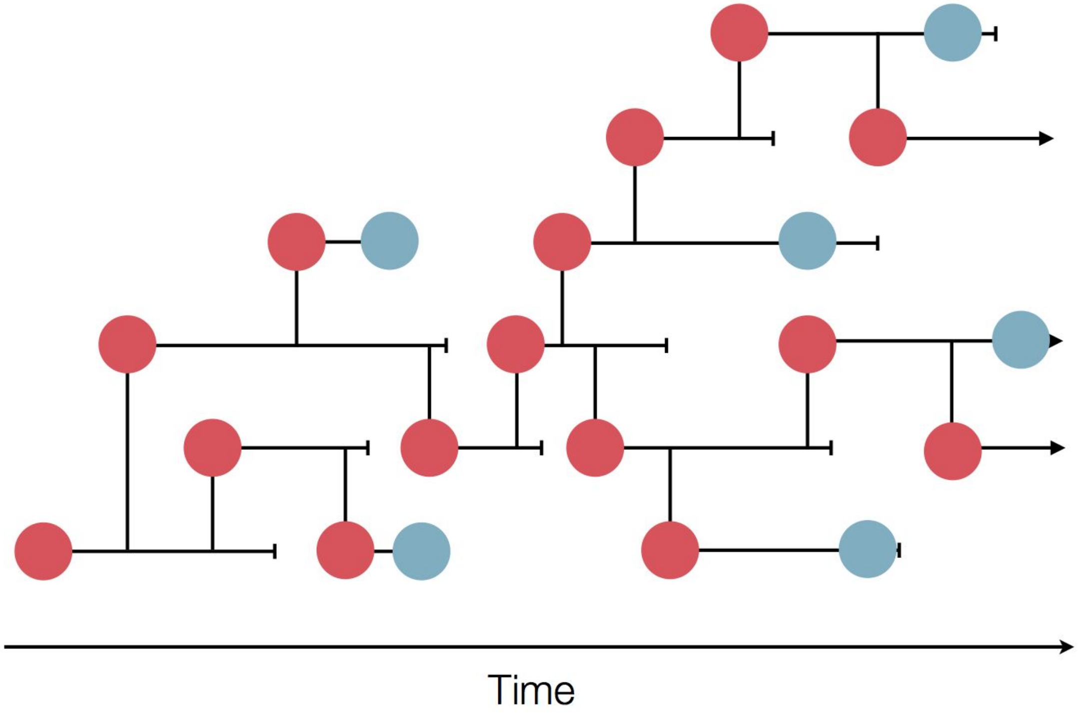
Does our refined phylogeny tell us the transmission events?

Phylogeny is a sampling of the underlying epidemic process



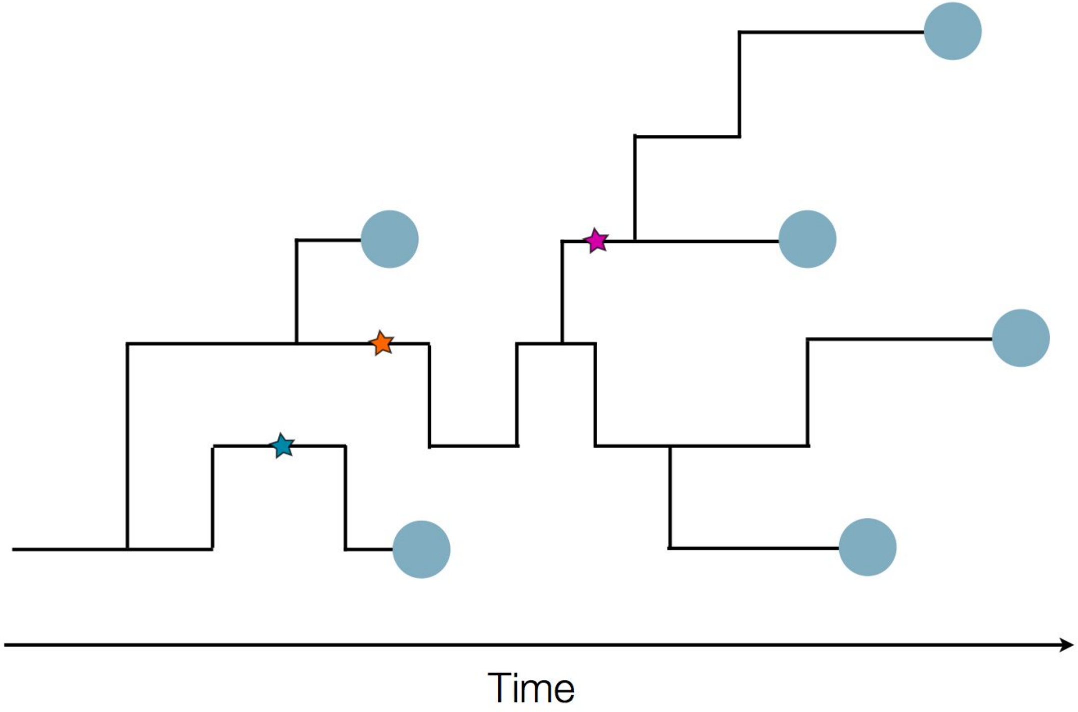
https://github.com/trvrb/phylodynamics-lecture

Phylogeny is a sampling of the underlying epidemic process



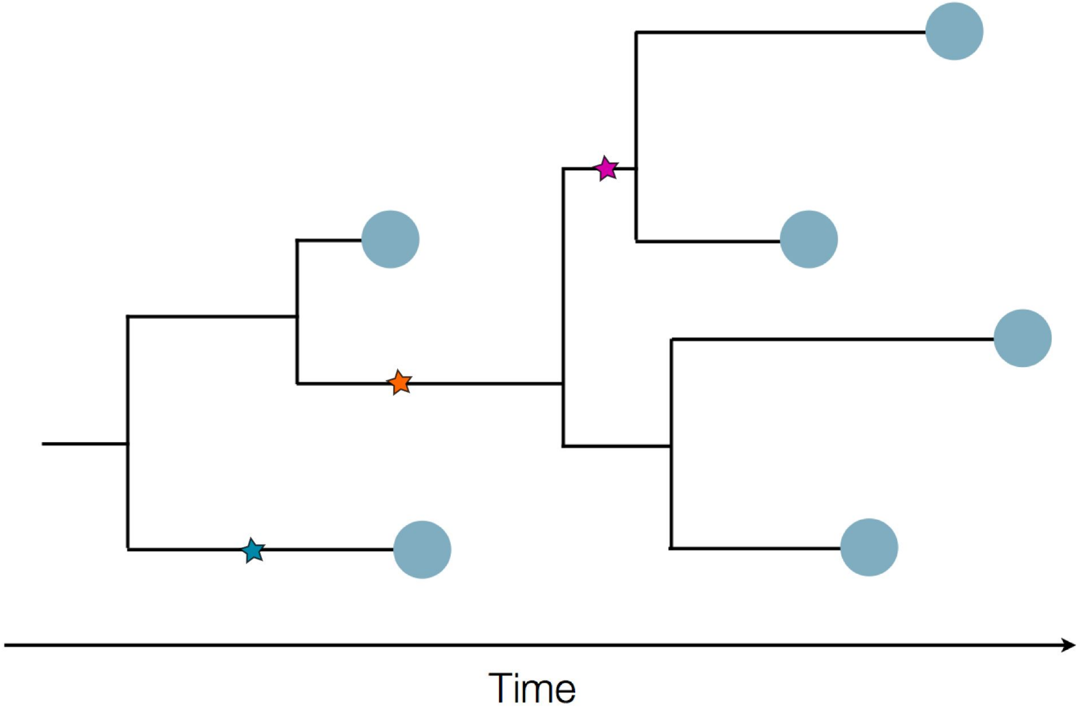
https://github.com/trvrb/phylodynamics-lecture

Phylogeny is a sampling of the underlying epidemic process



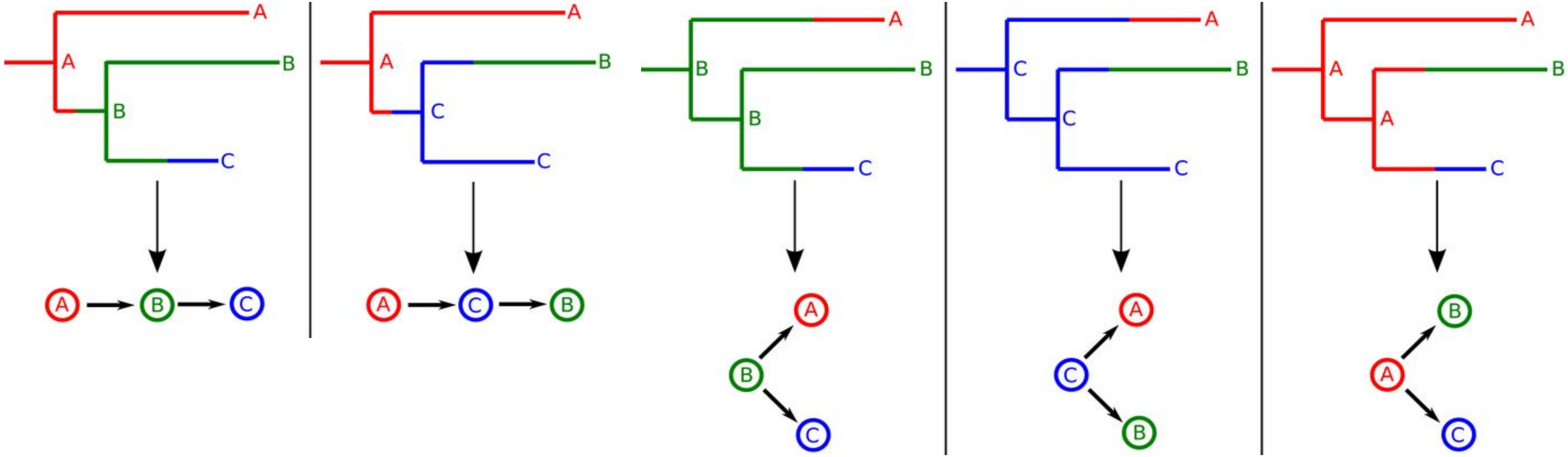
https://github.com/trvrb/phylodynamics-lecture

Phylogeny is a sampling of the underlying epidemic process



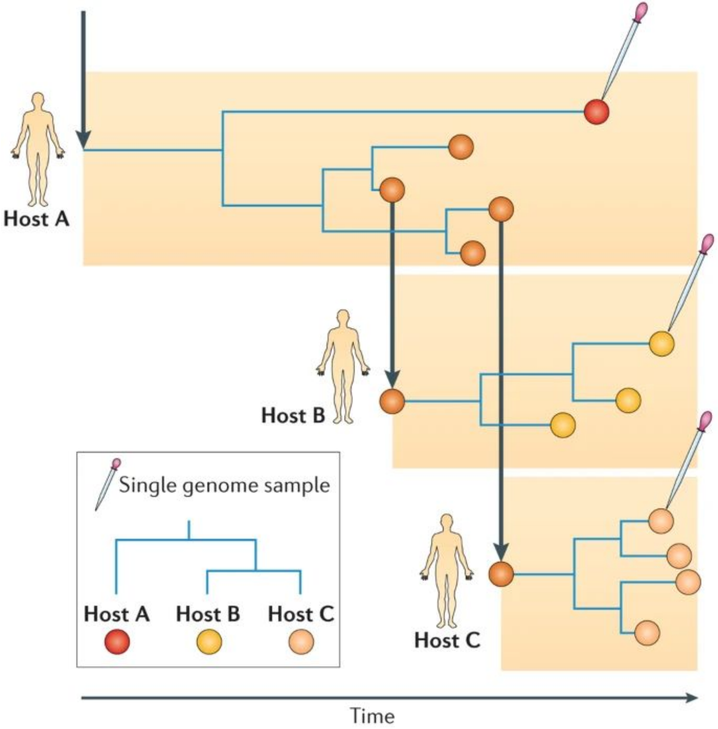
https://github.com/trvrb/phylodynamics-lecture

Same tree can be compatible with different transmissions



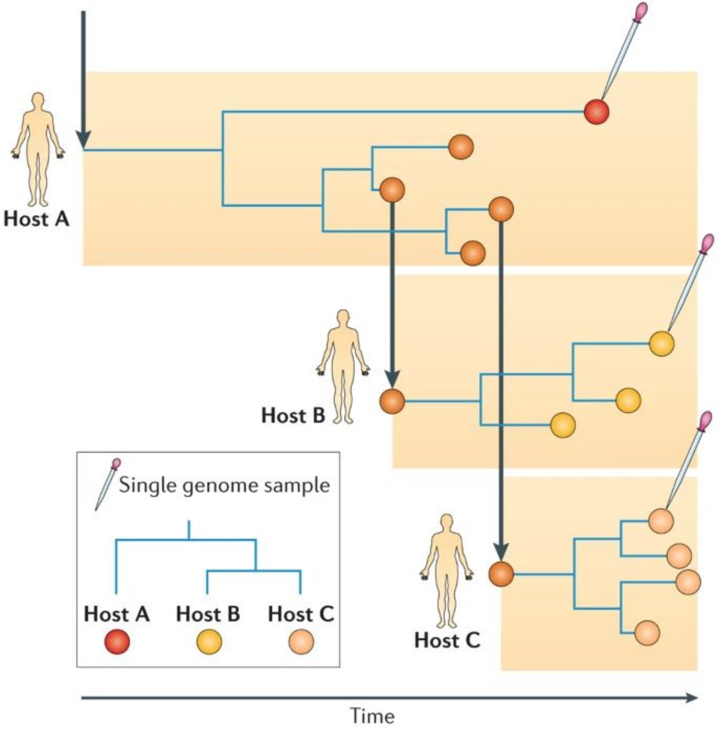
Hall, Matthew, Mark Woolhouse, and Andrew Rambaut. "Epidemic reconstruction in a phylogenetics framework: transmission trees as partitions of the node set." PLoS computational biology 11.12 (2015): e1004613.

Within-host diversity & missing hosts add to complications

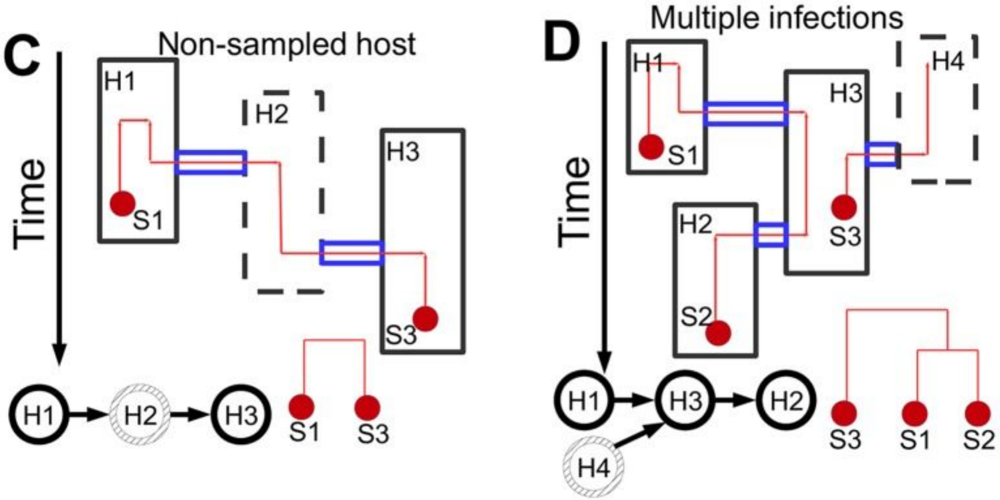


Didelot, Xavier, et al. "Within-host evolution of bacterial pathogens." Nature Reviews Microbiology 14.3 (2016): 150-162.

Within-host diversity & missing hosts add to complications



Didelot, Xavier, et al. "Within-host evolution of bacterial pathogens." Nature Reviews Microbiology 14.3 (2016): 150-162.



De Maio, Nicola, Chieh-Hsi Wu, and Daniel J. Wilson. "SCOTTI: efficient reconstruction of transmission within outbreaks with the structured coalescent." PLoS computational biology 12.9 (2016): e1005130.

So, how do we infer transmission?

SeqTrack provides a simple parsimony-based inference

Host A
ACGTACT
Day 1

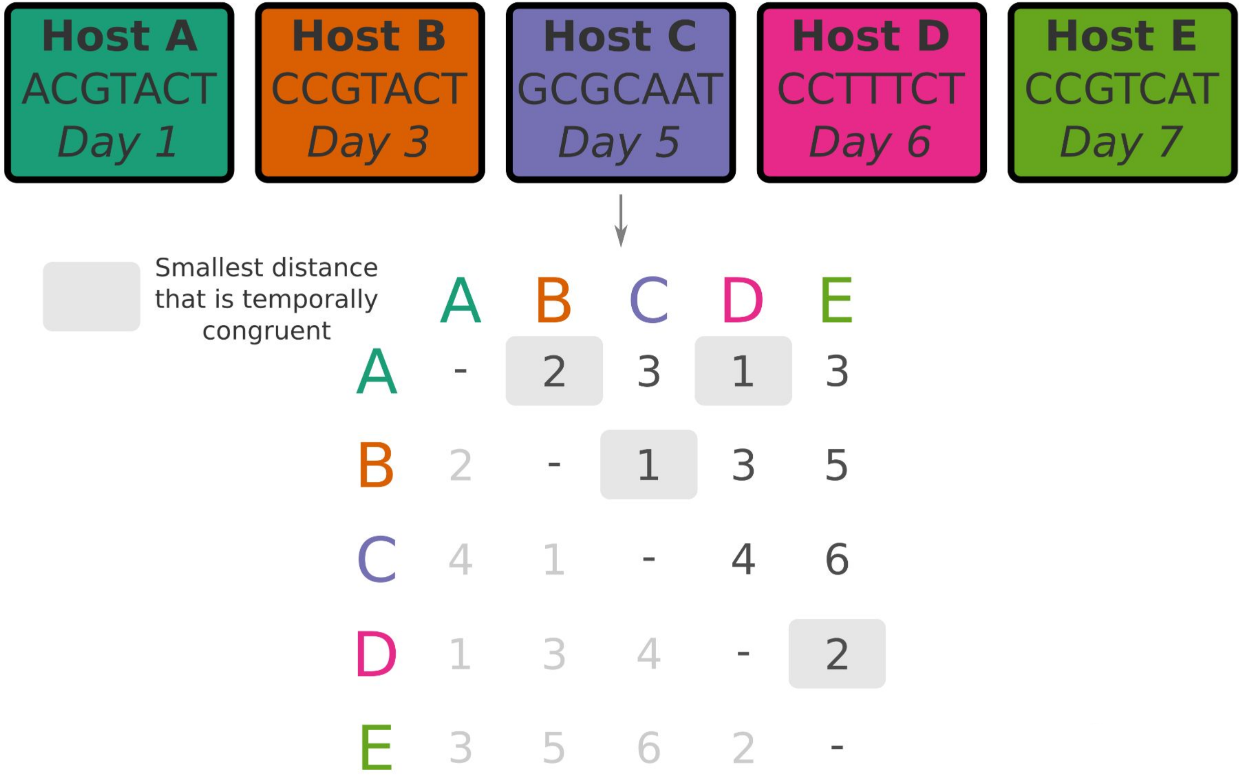
Host B
CCGTACT
Day 3

Host C
GCGCAAT
Day 5

Host D
CCTTTCT
Day 6

Host E
CCGTCAT
Day 7

SeqTrack provides a simple parsimony-based inference



Host A
ACGTACT
Day 1

Host B
CCGTACT
Day 3

Host C
GCGCAAT
Day 5

Host D
CCTTTCT
Day 6

Host E
CCGTCAT
Day 7

▨ Smallest distance that is temporally congruent

	A	B	C	D	E
A	-	2	3	1	3
B	2	-	1	3	5
C	4	1	-	4	6
D	1	3	4	-	2
E	3	5	6	2	-

SeqTrack provides a simple parsimony-based inference

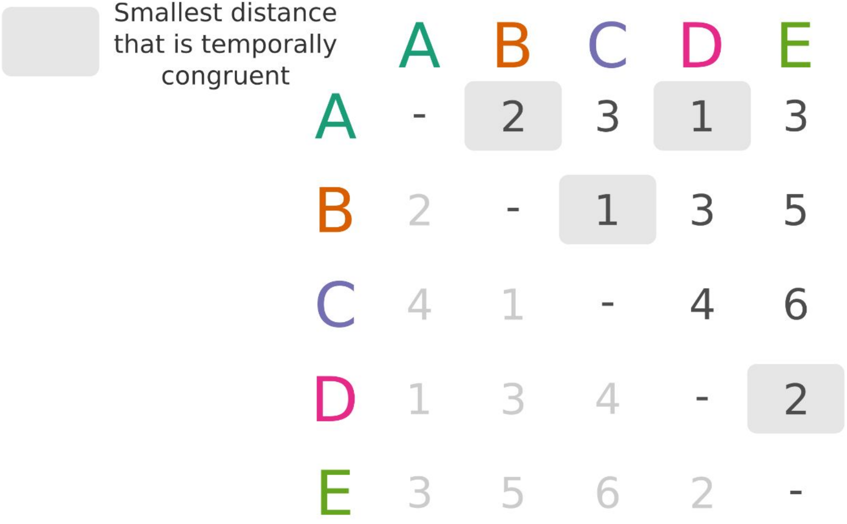
Host A
ACGTACT
Day 1

Host B
CCGTACT
Day 3

Host C
GCGCAAT
Day 5

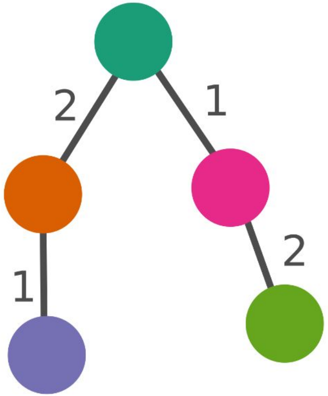
Host D
CCTTTCT
Day 6

Host E
CCGTCAT
Day 7



▨ Smallest distance that is temporally congruent

	A	B	C	D	E
A	-	2	3	1	3
B	2	-	1	3	5
C	4	1	-	4	6
D	1	3	4	-	2
E	3	5	6	2	-



SeqTrack provides a simple parsimony-based inference

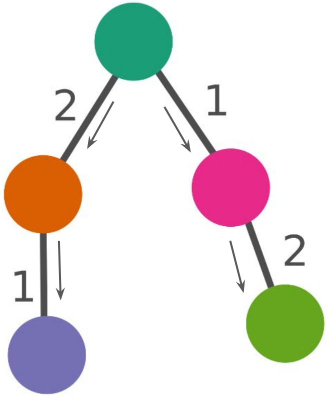
Host A	**Host B**	**Host C**	**Host D**	**Host E**
ACGTACT	CCGTACT	GCGCAAT	CCTTTCT	CCGTCAT
Day 1	*Day 3*	*Day 5*	*Day 6*	*Day 7*

▨ Smallest distance that is temporally congruent

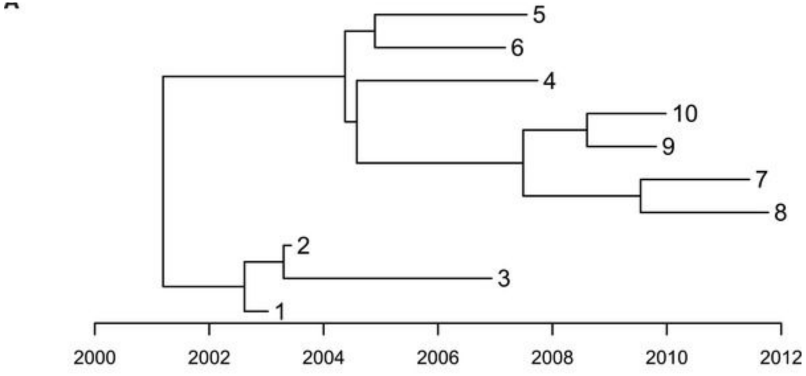
	A	B	C	D	E
A	-	2	3	1	3
B	2	-	1	3	5
C	4	1	-	4	6
D	1	3	4	-	2
E	3	5	6	2	-

Assumptions:

- All hosts sampled
- t(Ancestor) < t(Descendent)
- Closest = source

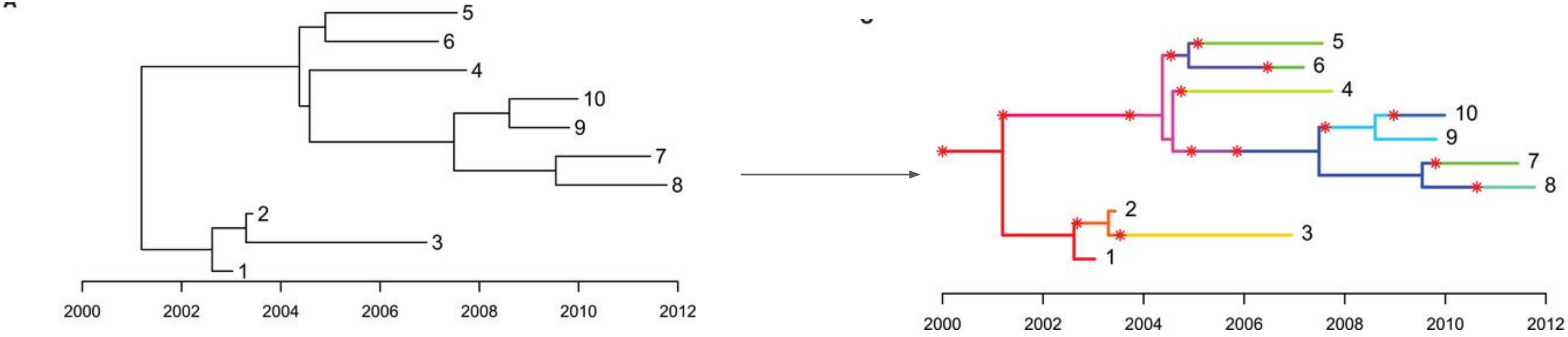


TransPhylo colors a dated tree using Bayesian inference



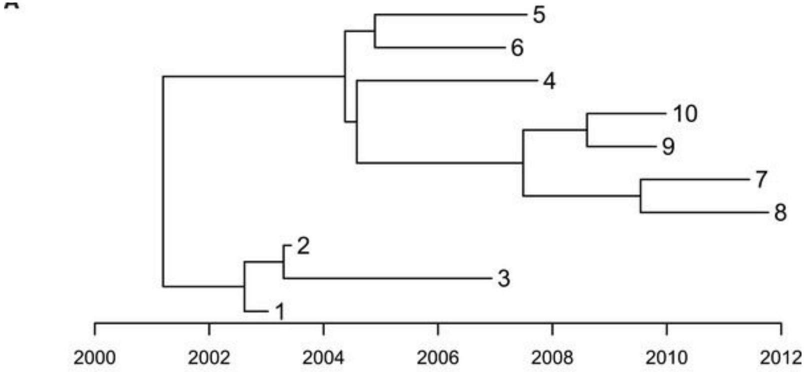
Didelot, Xavier, et al. "Genomic epidemiology analysis of infectious disease outbreaks using TransPhylo." Current protocols 1.2 (2021): e60.

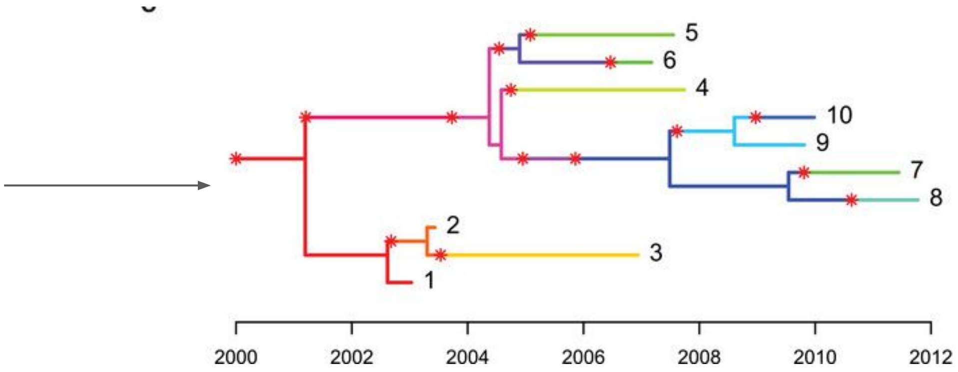
TransPhylo colors a dated tree using Bayesian inference

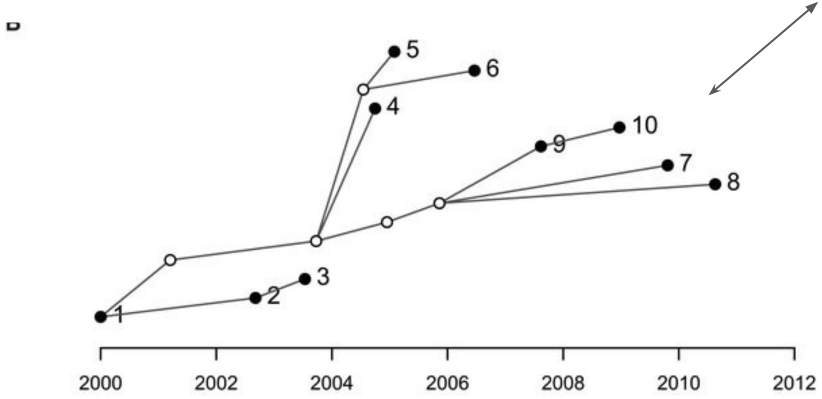


Didelot, Xavier, et al. "Genomic epidemiology analysis of infectious disease outbreaks using TransPhylo." Current protocols 1.2 (2021): e60.

TransPhylo colors a dated tree using Bayesian inference

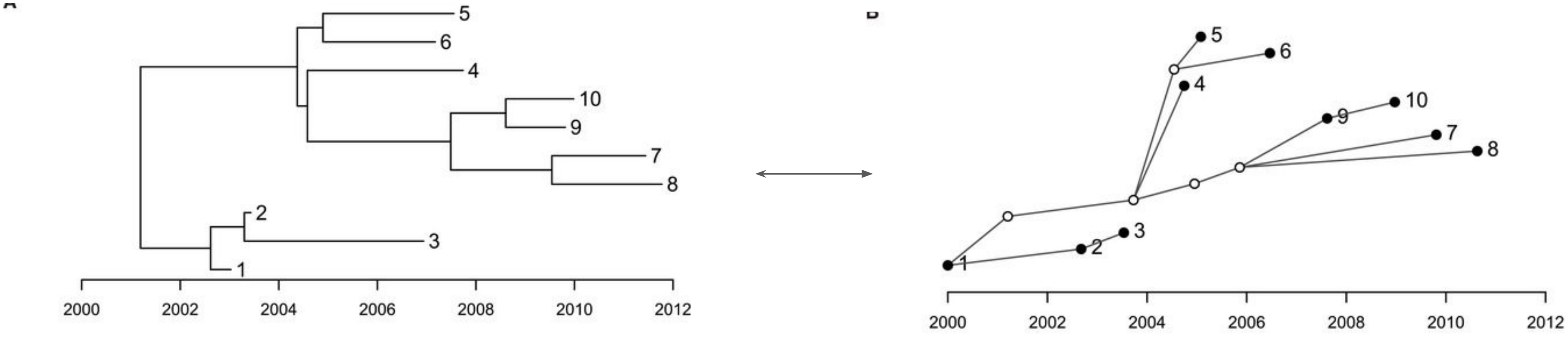






Didelot, Xavier, et al. "Genomic epidemiology analysis of infectious disease outbreaks using TransPhylo." Current protocols 1.2 (2021): e60.

PhyBreak & SCOTTI jointly infer phylogeny & transmission

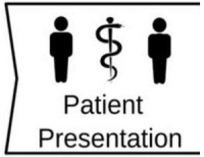


More on these types of analyses in phylodynamics

What about mobile gene outbreaks?

Mobile gene outbreaks make this complicated

Patient A: *Proteus*
Patient B: *Enterobacter*
Patient B: *Pseudomonas*



Link? ✓

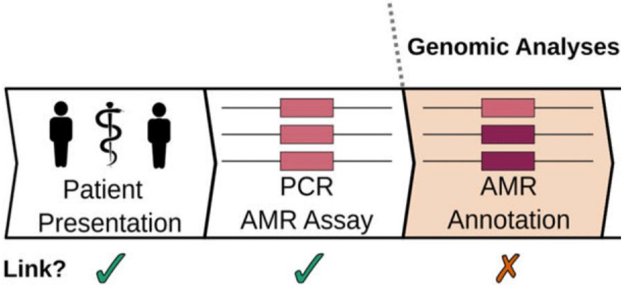
Mahoney, David Burke James, et al. "Utility of hybrid whole genome sequencing in assessing potential nosocomial VIM transmission." Antimicrobial Stewardship & Healthcare Epidemiology 4.1 (2024): e106.

Mobile gene outbreaks make this complicated

Mahoney, David Burke James, et al. "Utility of hybrid whole genome sequencing in assessing potential nosocomial VIM transmission." Antimicrobial Stewardship & Healthcare Epidemiology 4.1 (2024): e106.

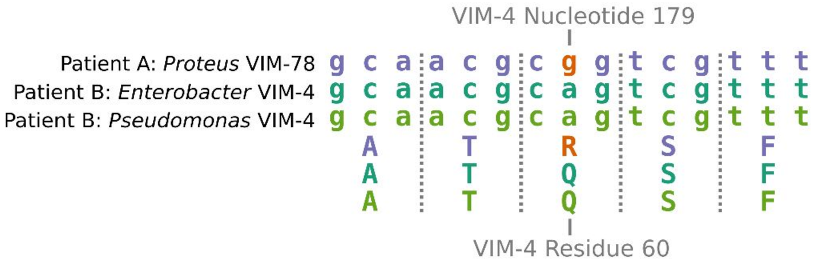
Mobile gene outbreaks make this complicated

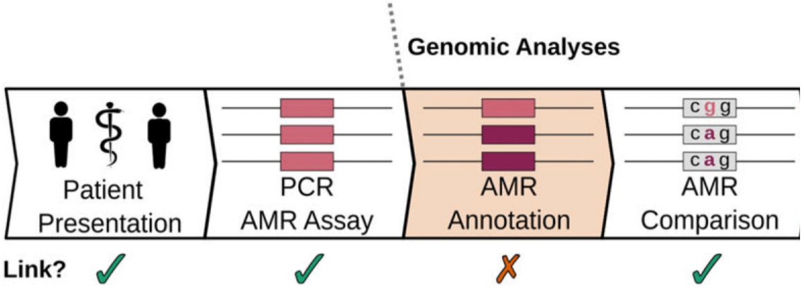
Patient A: *Proteus* VIM-78
Patient B: *Enterobacter* VIM-4
Patient B: *Pseudomonas* VIM-4



Mahoney, David Burke James, et al. "Utility of hybrid whole genome sequencing in assessing potential nosocomial VIM transmission." Antimicrobial Stewardship & Healthcare Epidemiology 4.1 (2024): e106.

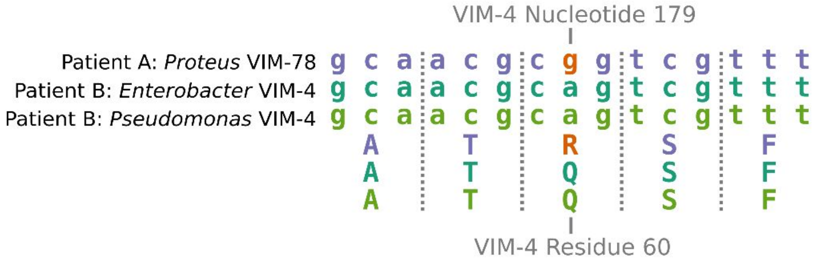
Mobile gene outbreaks make this complicated

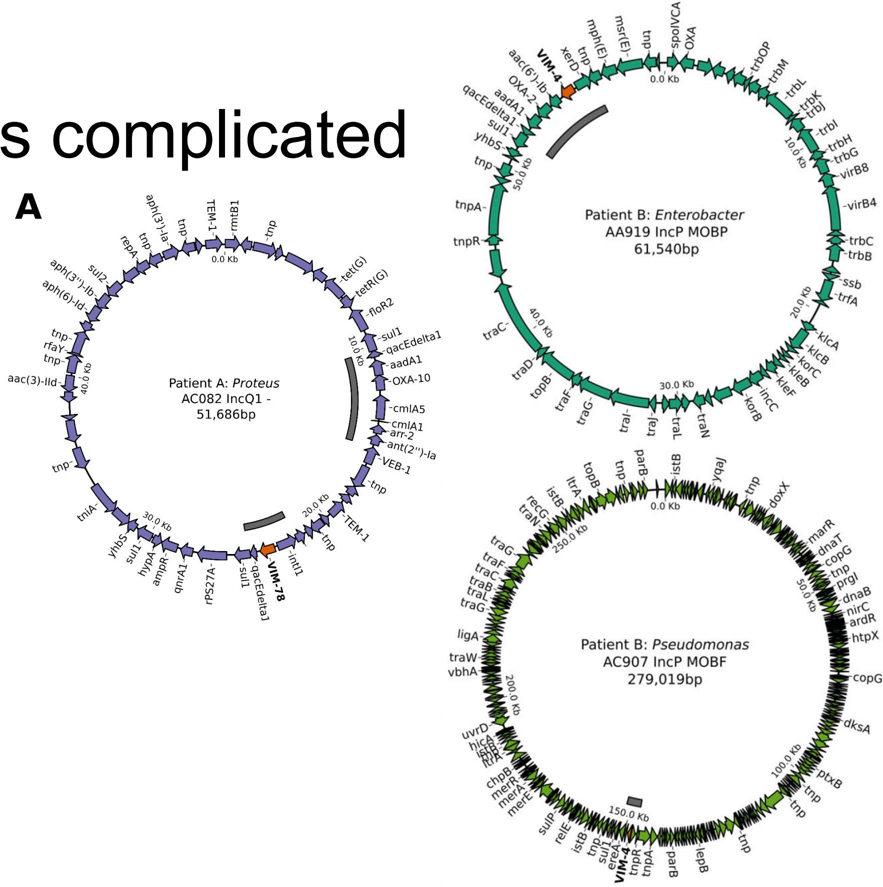


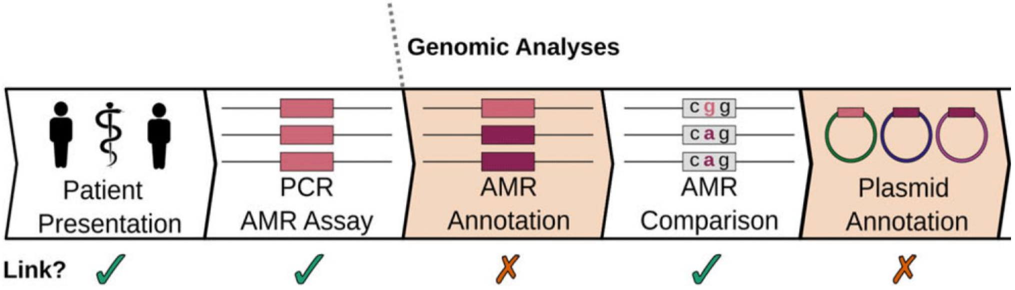


Mahoney, David Burke James, et al. "Utility of hybrid whole genome sequencing in assessing potential nosocomial VIM transmission." Antimicrobial Stewardship & Healthcare Epidemiology 4.1 (2024): e106.

Mobile gene outbreaks make this complicated

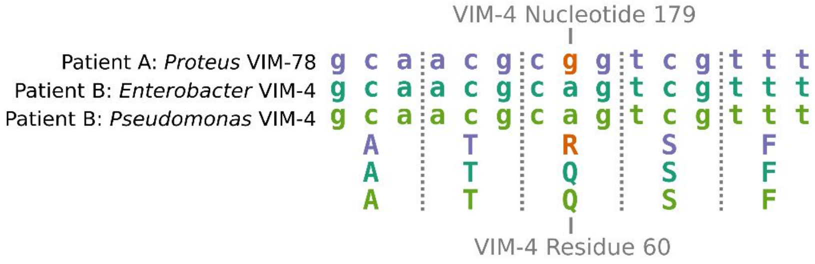


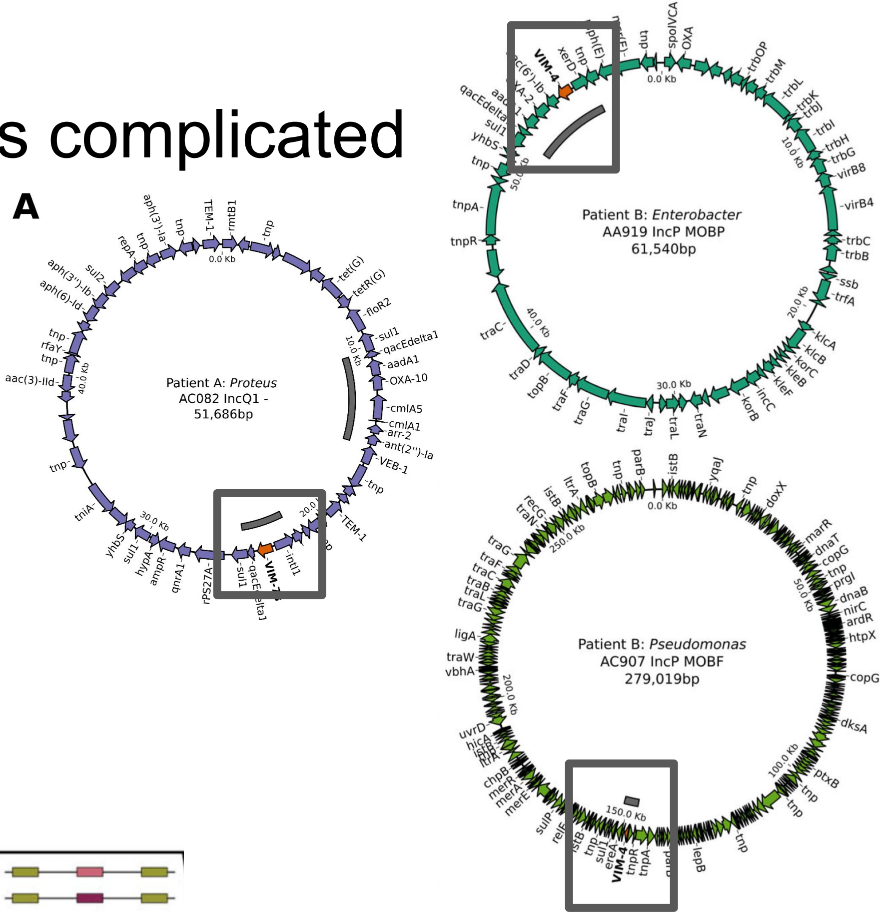


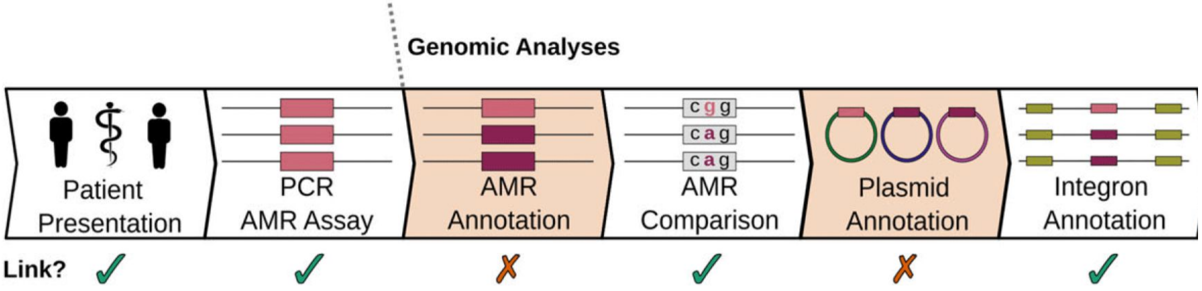


Mahoney, David Burke James, et al. "Utility of hybrid whole genome sequencing in assessing potential nosocomial VIM transmission." Antimicrobial Stewardship & Healthcare Epidemiology 4.1 (2024): e106.

Mobile gene outbreaks make this complicated

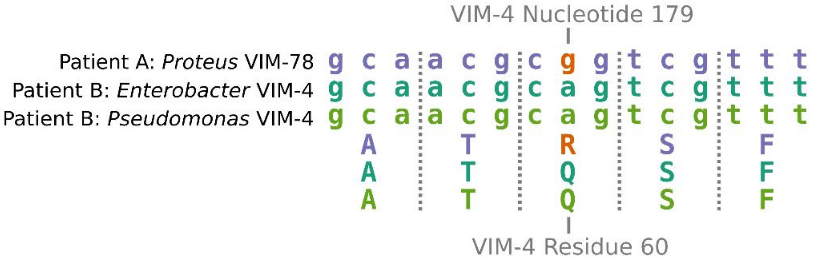


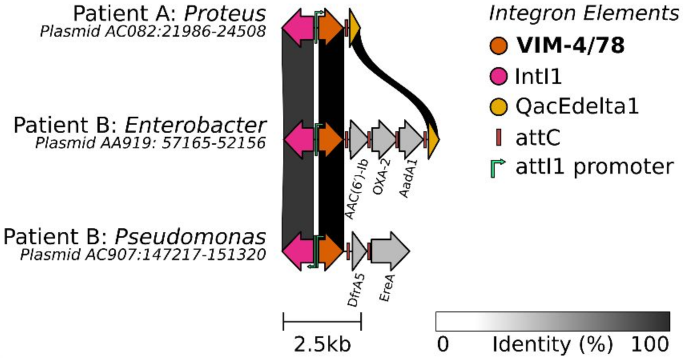


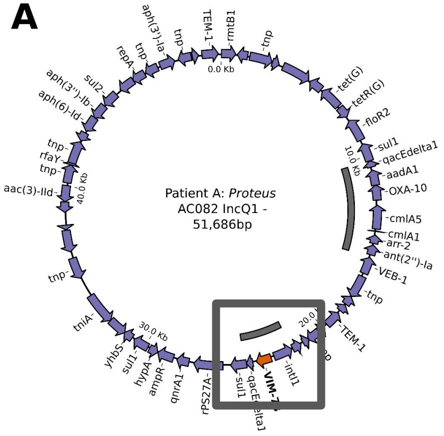


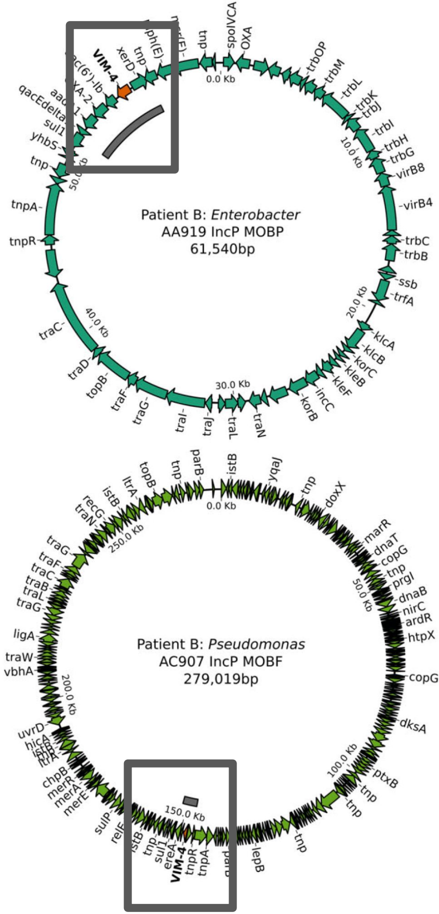
Mahoney, David Burke James, et al. "Utility of hybrid whole genome sequencing in assessing potential nosocomial VIM transmission." Antimicrobial Stewardship & Healthcare Epidemiology 4.1 (2024): e106.

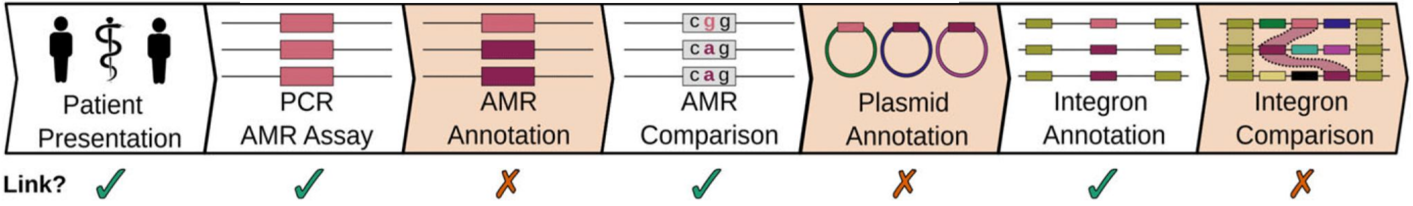
Mobile gene outbreaks make this complicated











	Patient Presentation	PCR AMR Assay	AMR Annotation	AMR Comparison	Plasmid Annotation	Integron Annotation	Integron Comparison
Link?	✓	✓	✗	✓	✗	✓	✗

Summary

- Outbreaks are defined relatively to baseline surveillance and can encompass a wide-range of scales/types
- Genomics (diagnostics, clustering/typing, phylogenetics, transmission inference, phylodynamic) can inform nearly all stages of an outbreak investigation
- Fast whole genome typing systems like PopPUNK can be used to rapidly find and refine outbreak isolates
- SKA-based phylogenetics enable rapid granular SNP phylogenies (and can be reference-free in low recombination contexts)
- Recombination distorts outbreak phylogenies and can be masked by identifying unexpected SNP distributions
- Phylogenies are compatible with multiple different transmission networks (within-host evolution, substitution-rate variability, and incomplete sampling add to complexity)
- Transmission inference ranges from simple parsimony approaches (SeqTrack) to probabilistic colouring of a dated phylogeny (TransPhylo) to complete joint Bayesian inferences (SCOTTI, PhyBreak)