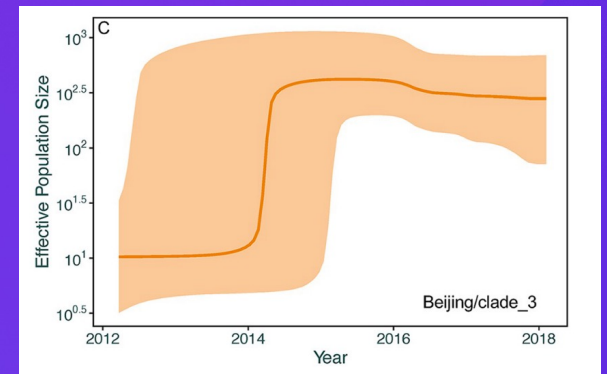
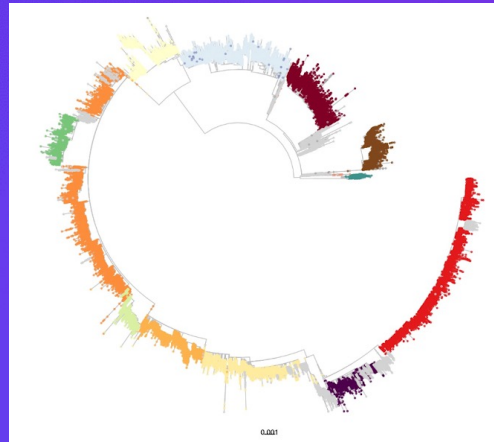


Genomic Analysis and Phylodynamics

Workshop: Simon Fraser University – 5th – 9th February 2024



Instructor: Dr. Ben Sobkowiak
Yale University / University College London

Lecture 3: Phylodynamics, phylogeography and molecular evolution

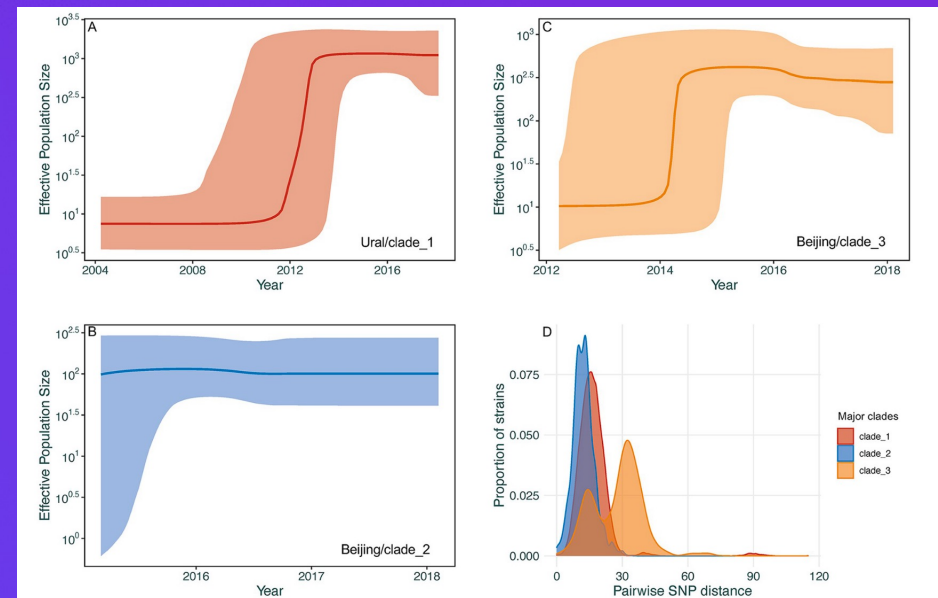
How can we use genomics and phylogenetics to learn more about pathogen populations and transmission?

- Evolutionary history and phylodynamics
 - Transmission Dynamics
 - Phylogeography
 - Identify sites under selection or significantly associated with a trait
 - Drug Resistance
 - Pathogenicity and virulence
 - Host adaptation
-

Lecture 3: Phylodynamics, phylogeography and molecular evolution

Phylodynamics

- Phylodynamics is the study of how epidemiological and evolutionary processes act to shape phylogenies
- Phylodynamic approaches typically involve the reconstruction of phylogenetic trees, combined with mathematical models that describe evolutionary processes and population dynamics

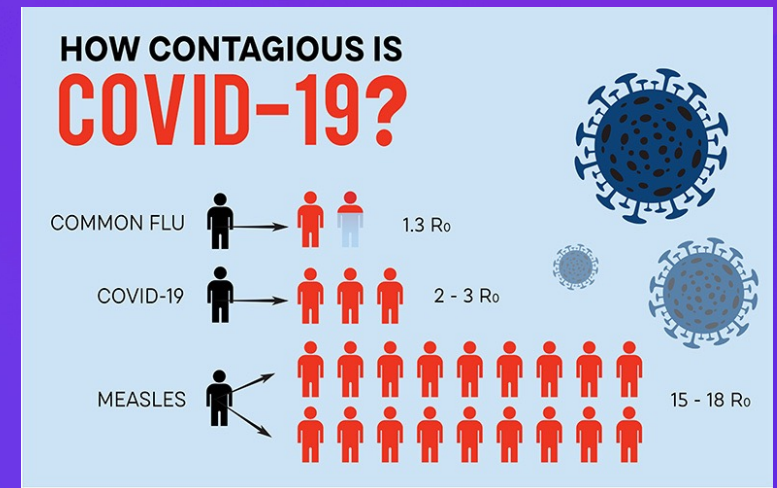


From Yang et. al. 2022

Lecture 3: Phylodynamics, phylogeography and molecular evolution

Phylodynamics

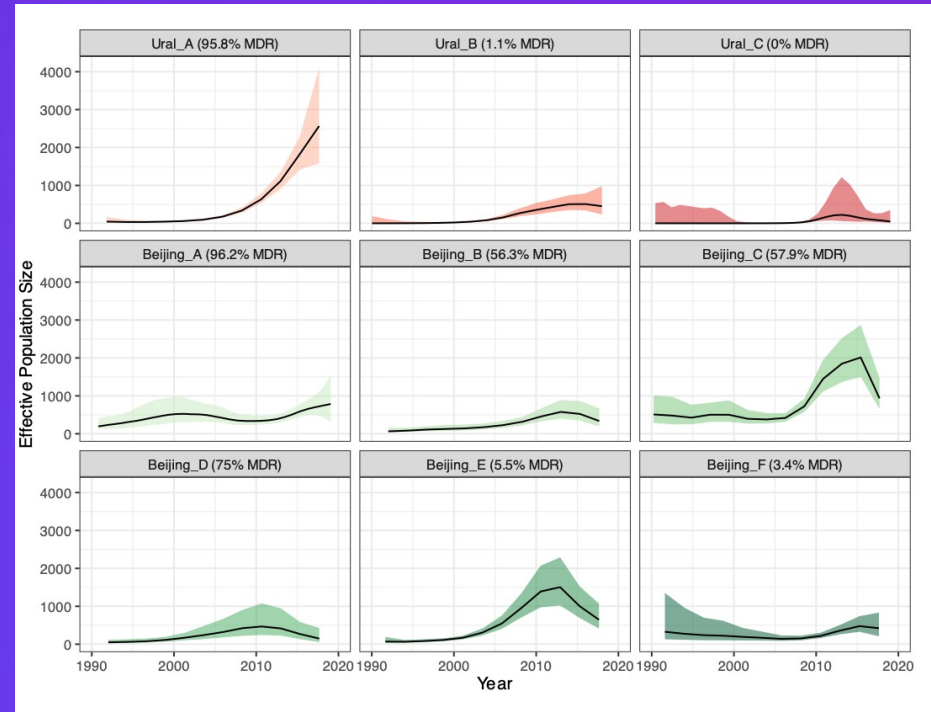
- These analyses can help us to estimate key epidemiological parameters:
 - R_0
 - Transmission rate
 - Force of infection
 - Serial intervals
- Infer past population demographics to estimate how lineages or clades on a tree have grown or shrunk through time



From endeavor.moffitt.org

Lecture 3: Phylodynamics, phylogeography and molecular evolution

- Reconstruct past population dynamics of 9 clades of M.tb in Moldova.
- Found evidence of recent, rapid expansion of one MDR clade
- Leads to more questions, what's driving this expansion?

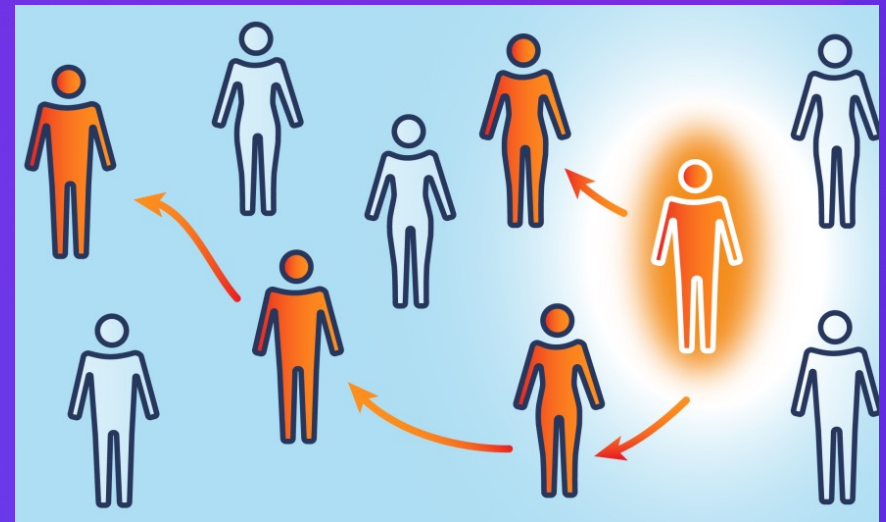


From Chitwood et. al. 2024, In submission

Lecture 3: Phylodynamics, phylogeography and molecular evolution

Transmission dynamics

- We can link phylogenetic trees with transmission to reconstruct transmission networks
- Phylogenetic tree \neq transmission tree in all instances, particularly when taxa have complex epidemiology
- Can use sophisticated approaches to incorporate epidemiological models with phylogenetics to estimate transmission trees



Johns Hopkins Bloomberg School of Public Health

Lecture 3: Phylodynamics, phylogeography and molecular evolution

Transmission dynamics

- Example software to estimate transmission networks from phylogenies or genomic data
 - TransPhylo
 - BEASTLIER
 - SCOTTi
 - Outbreaker
- Employ different underlying models and parameters that can result in different inferred transmission networks

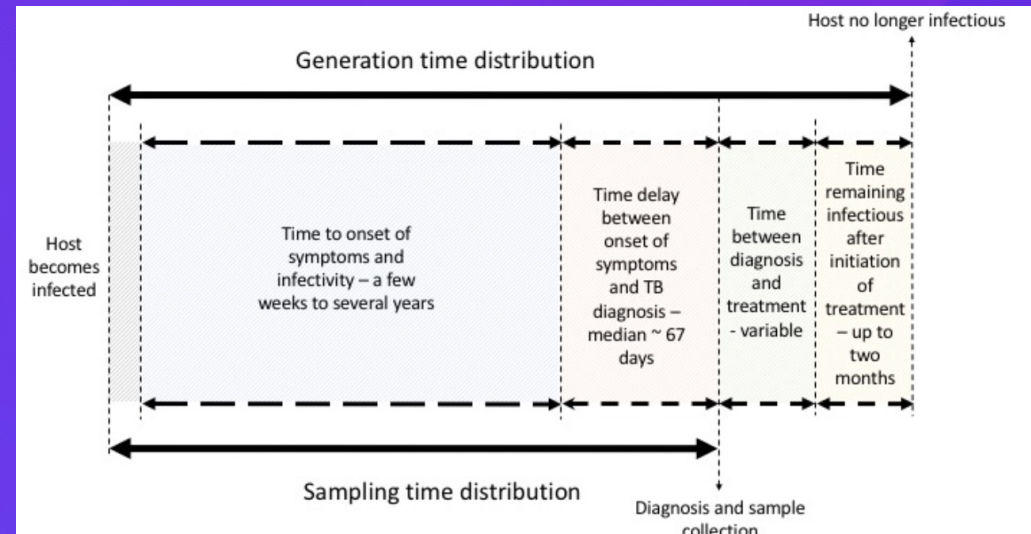


From Sobkowiak et al, 2023

Lecture 3: Phylodynamics, phylogeography and molecular evolution

TransPhylo

- Bayesian method for inferring transmission events from a timed phylogeny
- Epidemiological parameters including:
 - Generation time distribution
 - Sampling time distribution
 - Sampling density
 - Within-host coalescent rate

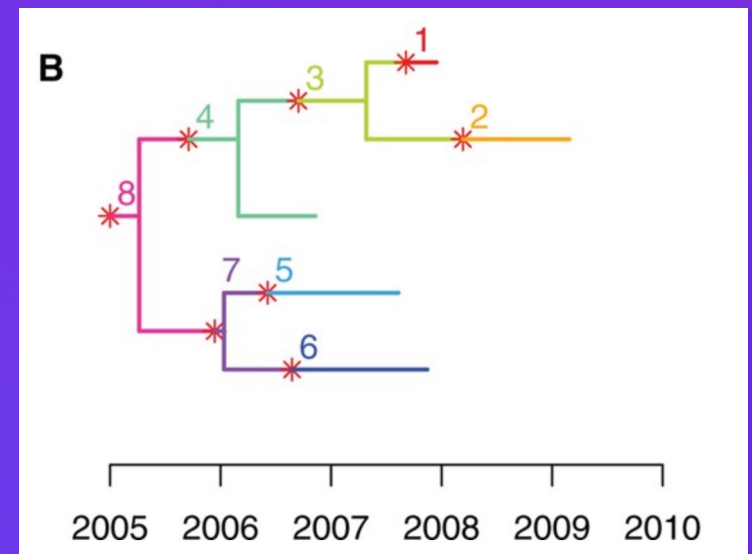


From Sobkowiak et. al. 2020

Lecture 3: Phylodynamics, phylogeography and molecular evolution

TransPhylo

- Colours the phylogeny by transmission events
- Each new sampled or unsampled case is given a unique colour and transmission events illustrated with the star

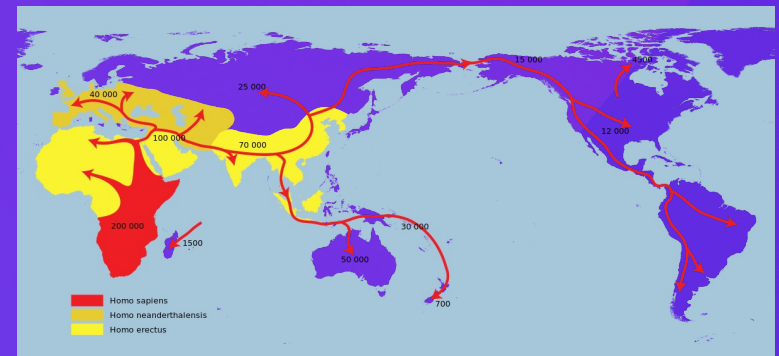


From Didelot et. al. 2017

Lecture 3: Phylodynamics, phylogeography and molecular evolution

Phylogeography

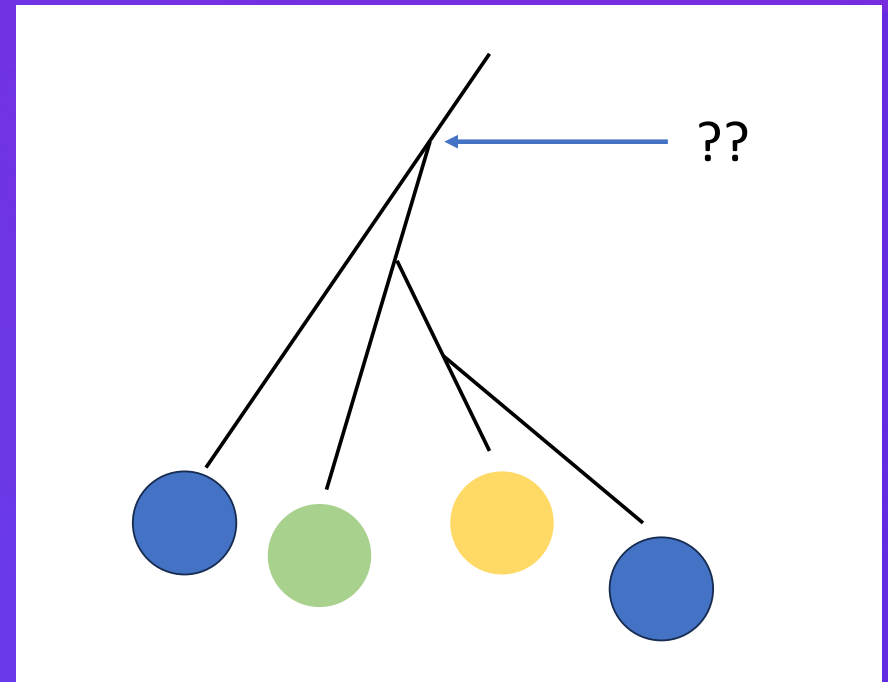
- Phylogeography allows us to reconstruct when and where lineages or clades were present
- Estimate the location of the emergence of new strains or lineages
- Track the migration of different strains or the flow of particular genes or traits



Lecture 3: Phylodynamics, phylogeography and molecular evolution

Phylogeography

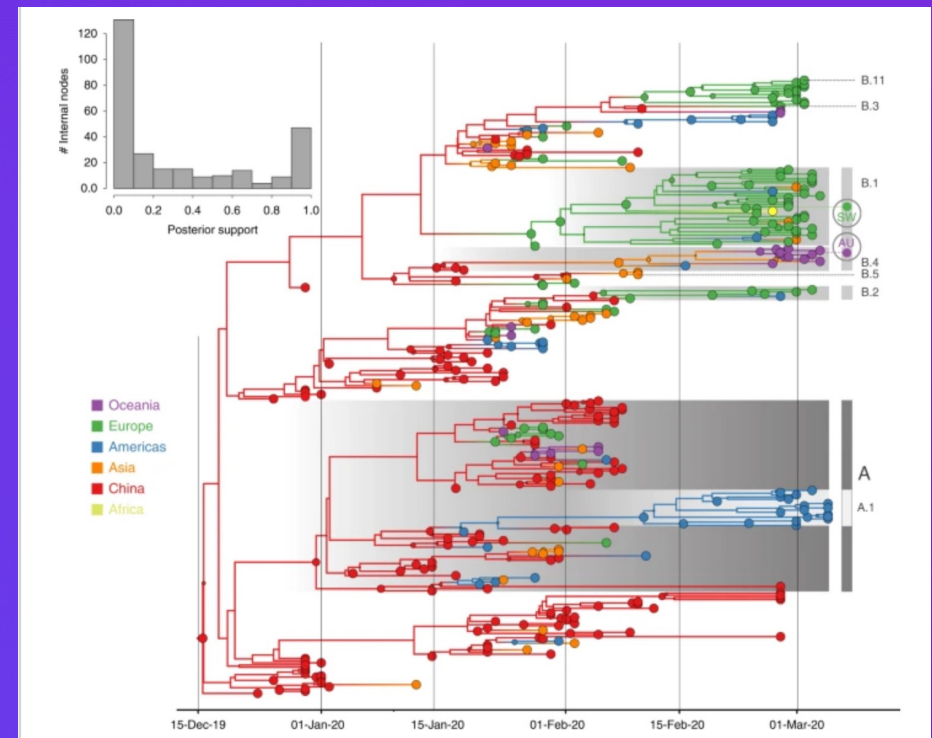
- In which location were past ancestors?
- Inferred using phylogenetic trees and the location data at tips using ancestral state reconstruction
- Employs probabilistic models to infer the most likely states of ancestral nodes, taking into account:
 - the observed character at tips
 - the topology of the phylogenetic tree
 - evolutionary processes governing the character evolution



Lecture 3: Phylodynamics, phylogeography and molecular evolution

Phylogeography

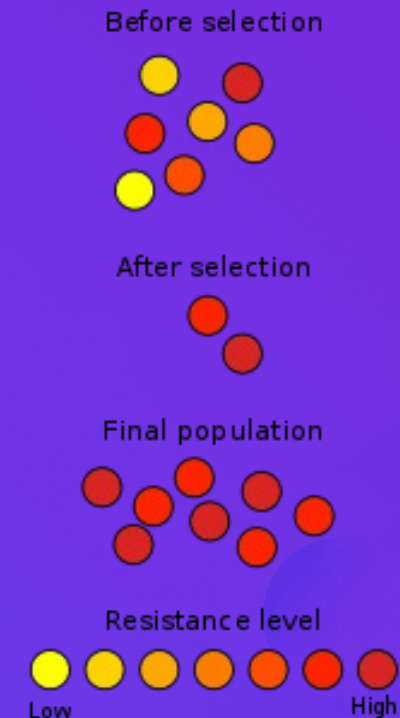
- Lemay et. al. 2020 reconstructed the location of the emergence of the COVID-19 pandemic
- Also inferred the dates in which there were first introductions to other regions



Lecture 3: Phylodynamics, phylogeography and molecular evolution

Testing for sites under selection

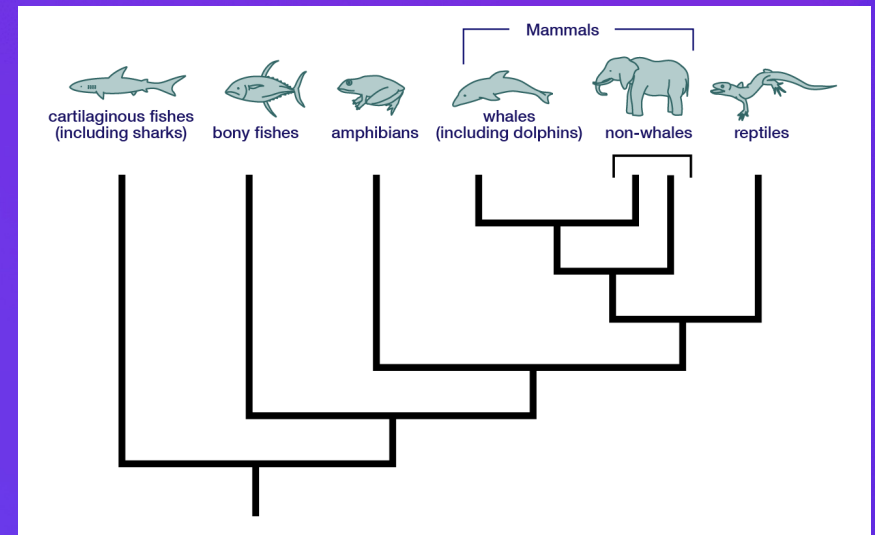
- Selection refers to the process by which certain heritable traits become more or less common in a population over time
- This process occurs because individuals with advantageous traits are more likely to survive and reproduce
- Selection can lead to adaptation and the evolution of new traits or the fixation of particular mutations



Lecture 3: Phylodynamics, phylogeography and molecular evolution

Convergent evolution of genomic variants

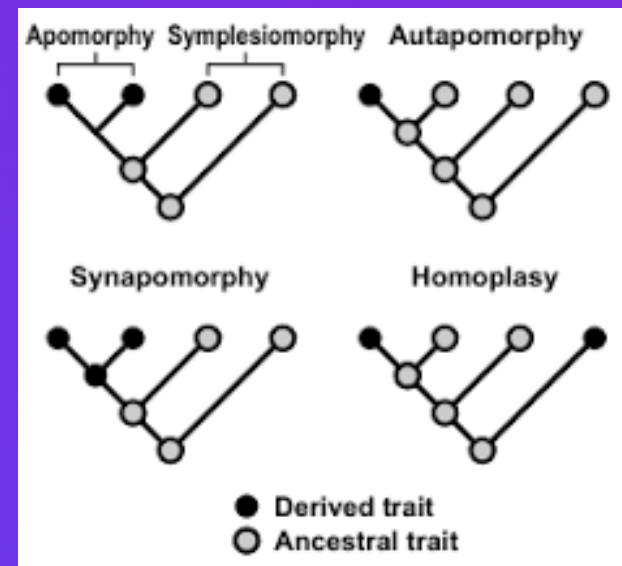
- The independent evolution of mutations or traits in distantly related individuals due to selective pressures or environmental constraints
- In genetics, sites under convergent evolution may represent instances of positive selection
- Mutations that enhance the fitness or adaptability of organisms to similar conditions



Lecture 3: Phylodynamics, phylogeography and molecular evolution

Homoplasy

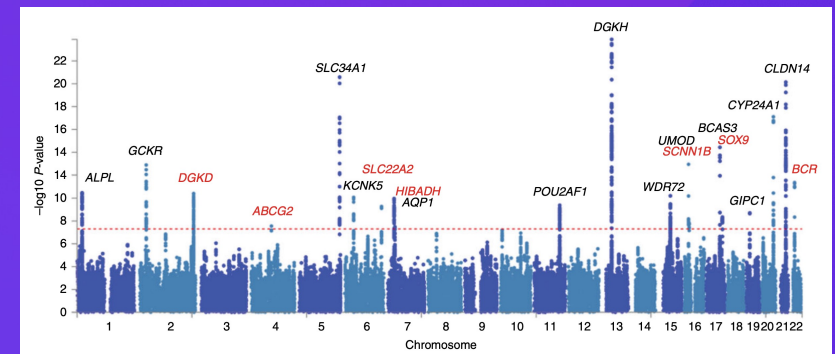
- Convergent evolution can cause homoplasies on a phylogenetic tree
- Shared traits or mutations among taxa that do not accurately reflect their evolutionary relationships
- Most methods use parsimony to infer the most likely ancestral states at internal nodes of the tree
- Identify branches where reversals or independent acquisitions occur, indicating homoplasy



Lecture 3: Phylodynamics, phylogeography and molecular evolution

Genome Wide Association Studies

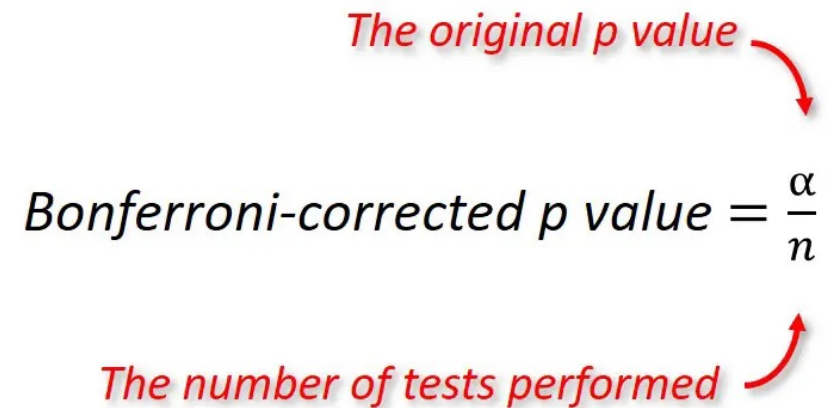
- GWAS is a test to identify genetic variants that are associated with a particular trait
- This may be mutations that cause antibiotic resistance, increase virulence or transmissibility, or evolve with host adaptation
- Advancements in sequencing technologies and bioinformatics have significantly enhanced the ability to conduct GWAS in microbial populations



Lecture 3: Phylodynamics, phylogeography and molecular evolution

Genome Wide Association Studies

- GWAS statistically analyzes genetic variants to identify any correlations with specific traits or diseases
- Multiple testing correction (such as Bonferroni) is applied to account for potentially thousands of sites being tested
- Also need to account for population structure to remove the confounding effect of genetic substructure in the population



The original p value

$$\text{Bonferroni-corrected } p \text{ value} = \frac{\alpha}{n}$$

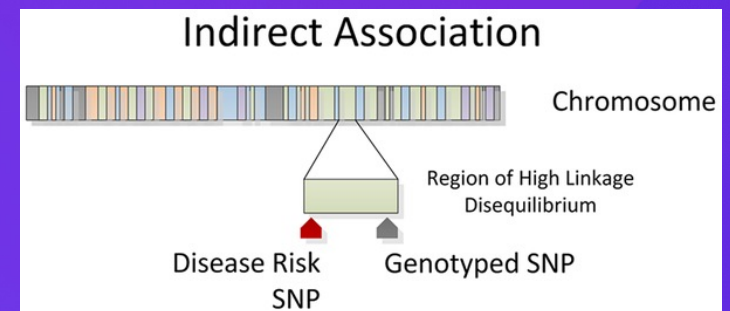
The number of tests performed

The diagram shows the formula for the Bonferroni-corrected p value. A red arrow points from the text 'The original p value' to the alpha symbol in the numerator. Another red arrow points from the text 'The number of tests performed' to the n symbol in the denominator.

Lecture 3: Phylodynamics, phylogeography and molecular evolution

Genome Wide Association Studies

- GWAS can be complicated in microbial populations for the following reasons:
 - Complex population structures, including clonal lineages, recombination events
 - Linkage disequilibrium, where alleles at different loci are inherited together due to limited recombination leading to spurious associations
 - Causal Inference, traits can be caused by multiple genes and environmental factors



From Bush & Moore, 2012

Lecture 3: Phylodynamics, phylogeography and molecular evolution

Genome Wide Association Studies

- GWAS has led to the discovery of novel variants associated with a variety of traits in a range of pathogens









Research article | [Open access](#) | Published: 05 June 2021

Genome-wide association studies reveal candidate genes associated to bacteraemia caused by ST93-IV CA-MRSA

[Stanley Pang](#) , [Denise A Daley](#), [Shafi Sahibzada](#), [Shakeel Mowlaboccus](#), [Marc Stegger](#) & [Geoffrey W Coombs](#)

BMC Genomics 22, Article number: 418 (2021) | [Cite this article](#)

Genome-Wide Association Studies for the Detection of Genetic Variants Associated With Daptomycin and Ceftaroline Resistance in *Staphylococcus aureus*

 Robert E. Weber¹  Stephan Fuchs²  Franziska Layer¹  Anna Sommer¹
 Jennifer K. Bender¹  Andrea Thürmer³  Guido Werner¹
 Birgit Strommenger^{1*}

¹ Department of Infectious Diseases, Robert Koch-Institute, Wernigerode, Germany

² Methodology and Research Infrastructure, Bioinformatics, Robert Koch-Institute, Berlin, Germany

³ Methodology and Research Infrastructure, Genome Sequencing, Robert Koch-Institute, Berlin, Germany

PLOS GENETICS

 OPEN ACCESS  PEER-REVIEWED

RESEARCH ARTICLE

Genome wide association study of *Escherichia coli* bloodstream infection isolates identifies genetic determinants for the portal of entry but not fatal outcome

[Erick Denamur](#) , [Bénédicte Condamine](#), [Marina Esposito-Farèse](#), [Guilhem Royer](#), [Olivier Clermont](#), [Cédric Laouenan](#), [Agnès Lefort](#), [Victoire de Lastours](#), [Marco Galardini](#) , the COLIBAFI , SEPTICOLI groups 

Published: March 24, 2022 • <https://doi.org/10.1371/journal.pgen.1010112>

nature communications

[Explore content](#) ▾ [About the journal](#) ▾ [Publish with us](#) ▾

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

Article | [Open access](#) | Published: 13 May 2019

GWAS for quantitative resistance phenotypes in *Mycobacterium tuberculosis* reveals resistance genes and regulatory regions

[Maha R. Farhat](#) , [Luca Freschi](#), [Roger Calderon](#), [Thomas Iøerger](#), [Matthew Snyder](#), [Conor J. Meehan](#), [Bouke de Jong](#), [Leen Rigouts](#), [Alex Sloutsky](#), [Devinder Kaur](#), [Shamil Sunyaev](#), [Dick van Soolingen](#), [Jay Shendure](#), [Jim Sacchettini](#) & [Megan Murray](#)

Nature Communications 10, Article number: 2128 (2019) | [Cite this article](#)

Practical 4: Phylogeography and phylodynamic analysis

1. Phylodynamic analysis in BEAST2
 2. Phylogeography using ancestral state reconstruction
 3. Identifying sites under selection
 4. Genome Wide Association Studies
-