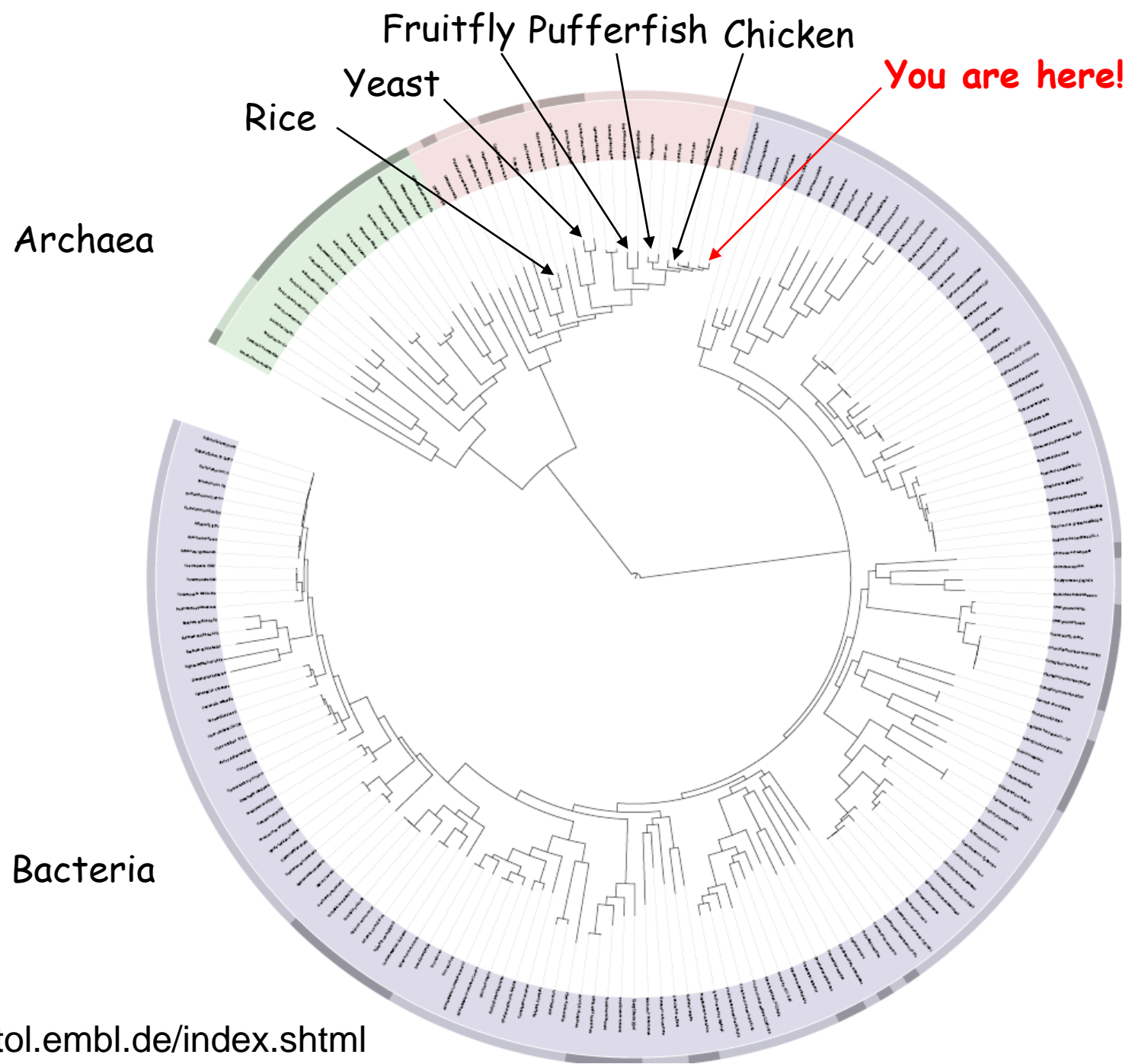


Why the evolution of bacteria is  
both interesting and important

Dick James

Ed Feil

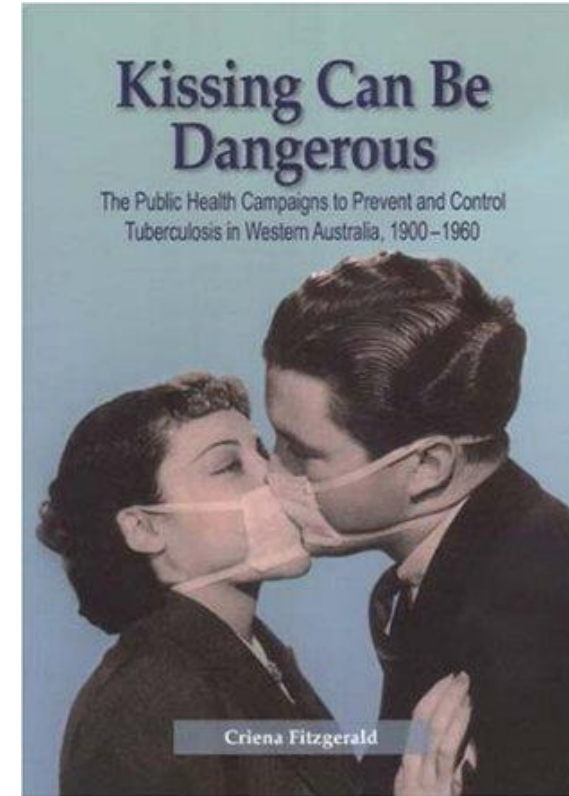




# Two biggest bacterial killers

*M. tuberculosis* – Tuberculosis. In 2014, 9.6 million people contracted TB, and 1.5 million died. One-third of the world's population is thought to be infected with TB. WHO declared TB a “global health emergency” in 1993.

*S. pneumoniae* - Pneumococcal disease (meningitis, pneumonia).  
1.6 Million deaths annually (0.7–1 million children aged under 5 years).



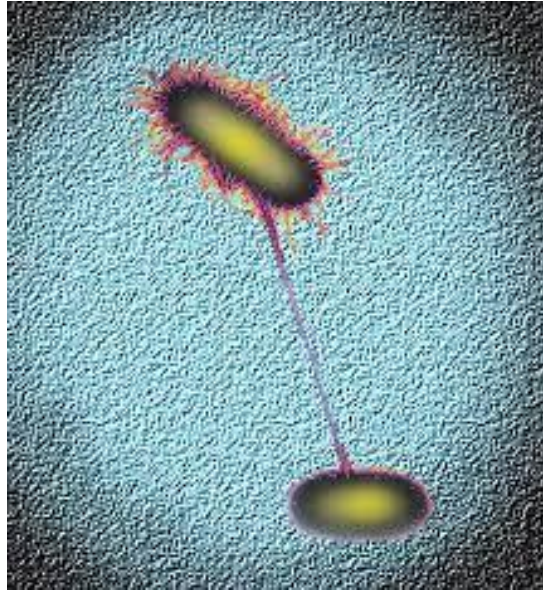
## The emergence and spread of antibiotic resistance

The most significant evolutionary change we have been able to observe in real time (both economically and in terms of human suffering)

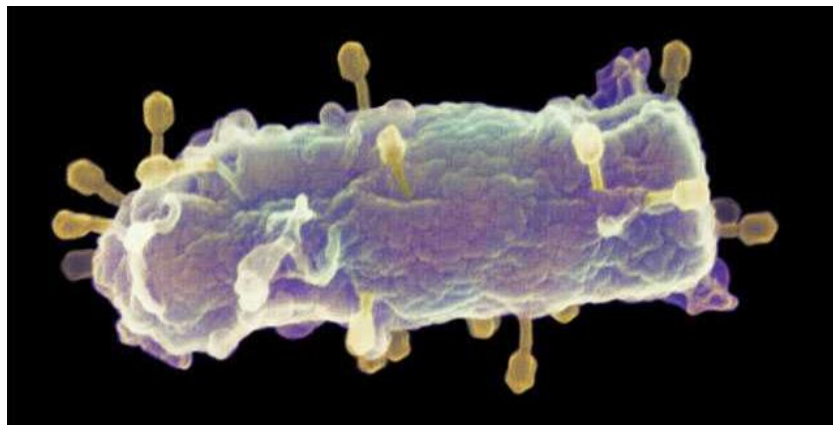
Illustrates how bacteria can adapt to new environments very rapidly:

- Large population sizes.
- Short generation times.
- The ability to “swap” genes (recombine).

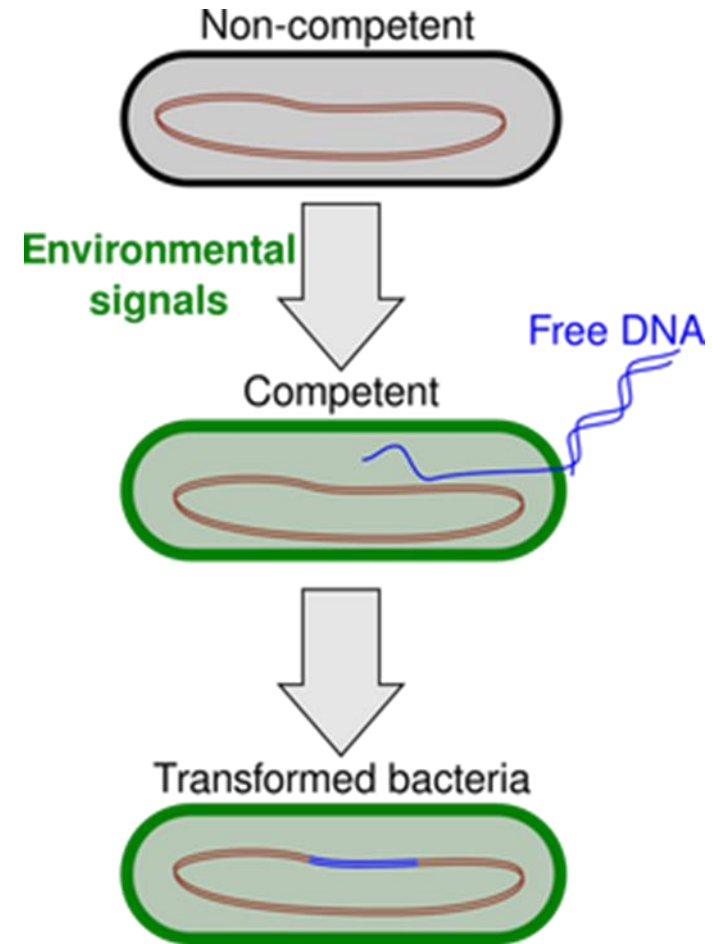
## Conjugation



## Transduction

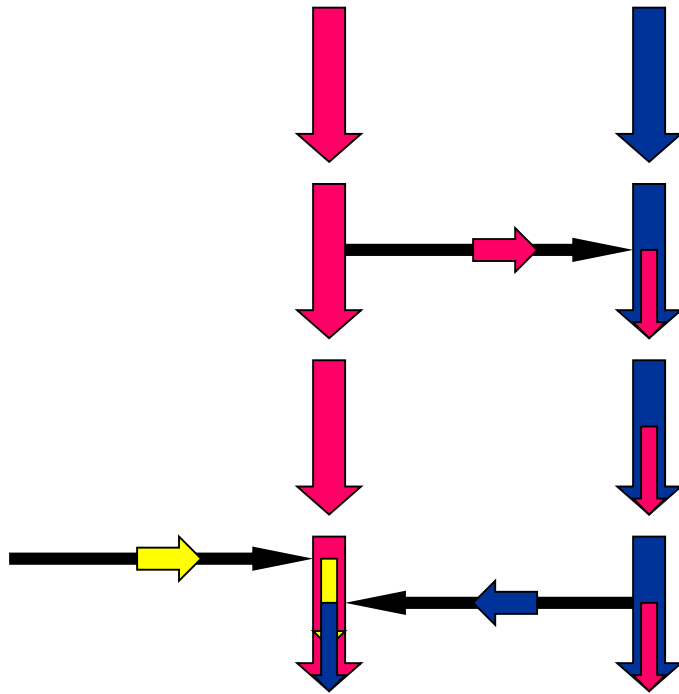


## Transformation

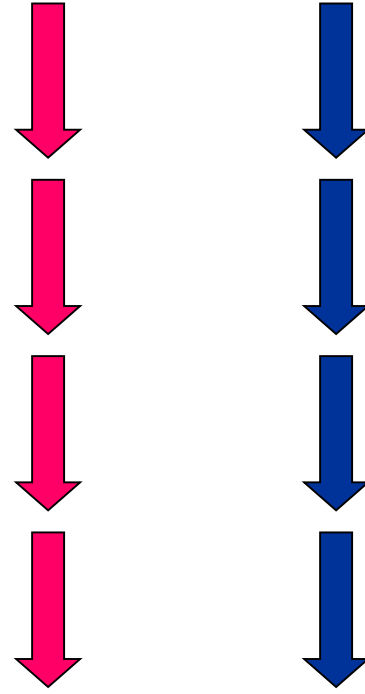


# Recombination in Bacteria

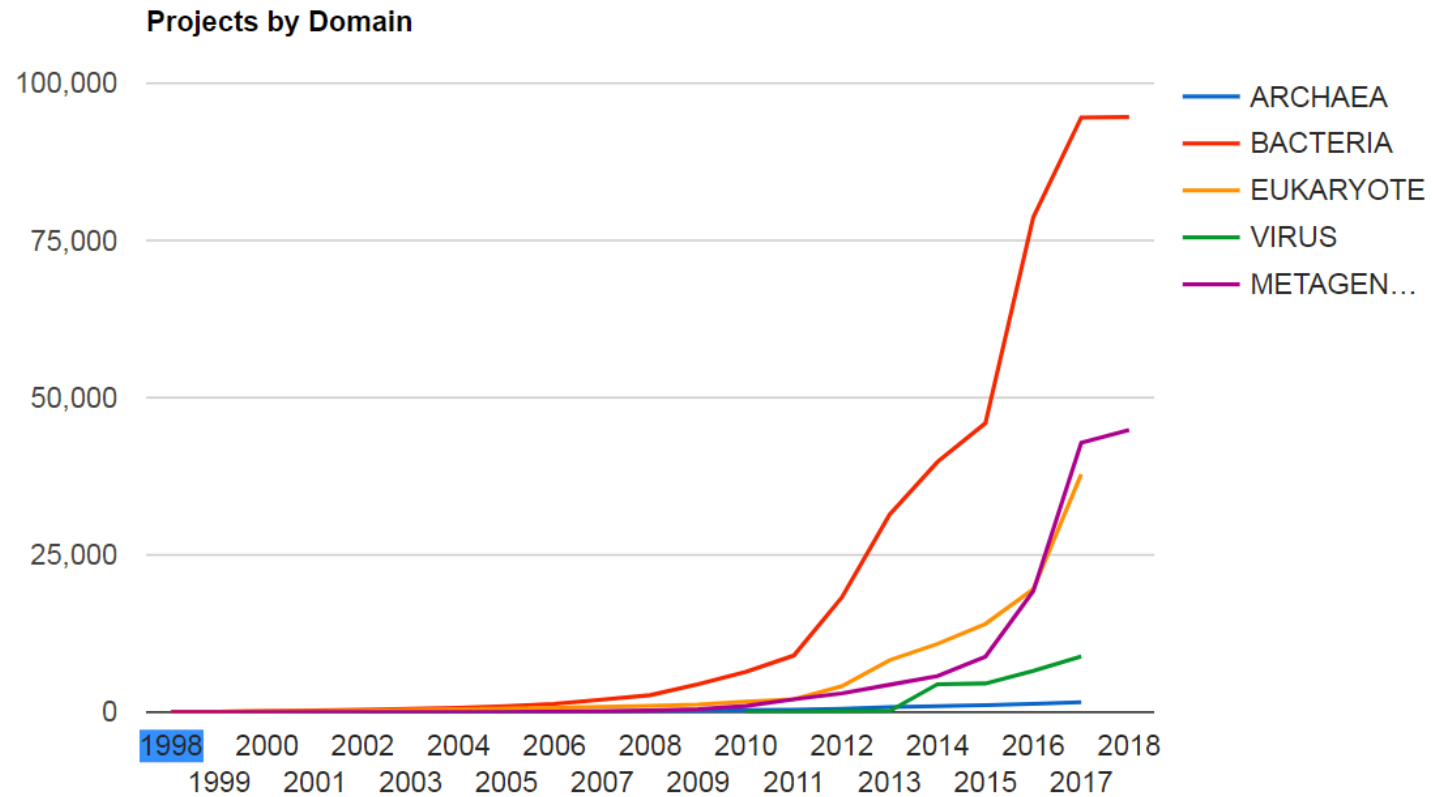
Recombination



No Recombination



## Genome sequences in the public domain



<http://www.genomesonline.org> (Jan 2018)



# The many ways in which bacterial genomes differ

Broad Phylogenetic Comparisons

Shallow Phylogenetic Comparisons



Genome “shape”  
(Circular / Linear)

Number of  
chromosomes

GC Content

Genome Size

Gene content diversity  
(MGEs)

Gene Order  
(Rearrangements)

Rates of recombination  
 (“population structure”)

Intra-species:

Single nucleotide  
polymorphisms

Small INDELS

Rate of mutation  
(hypermutation)

Number of plasmids

DNA modification

# Whole genome sequencing as a high resolution typing tool

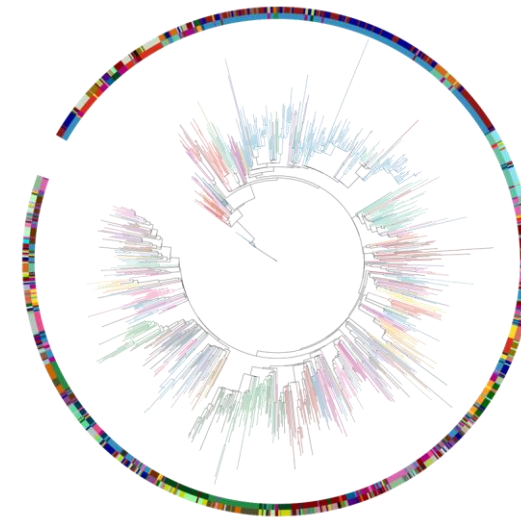
## Illumina reads mapped to a reference genome



Gather SNPs from other isolates

## Phylogenetic tree

Core SNPs used to construct a tree



A single SNP can distinguish closely related isolates

Retrospective Studies

VS

Real-time genomic surveillance

# UKCRC project

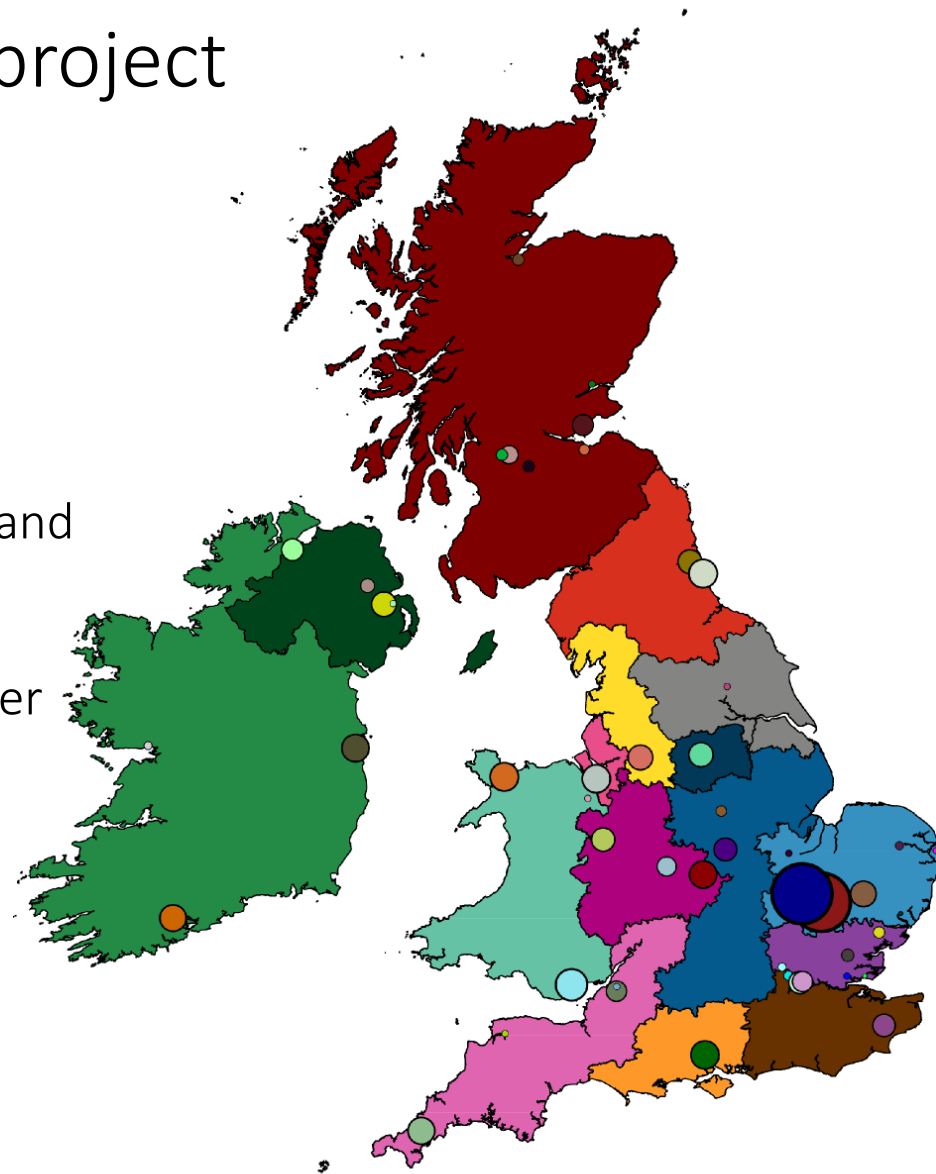
2046 clinical *Staphylococcus aureus* isolates

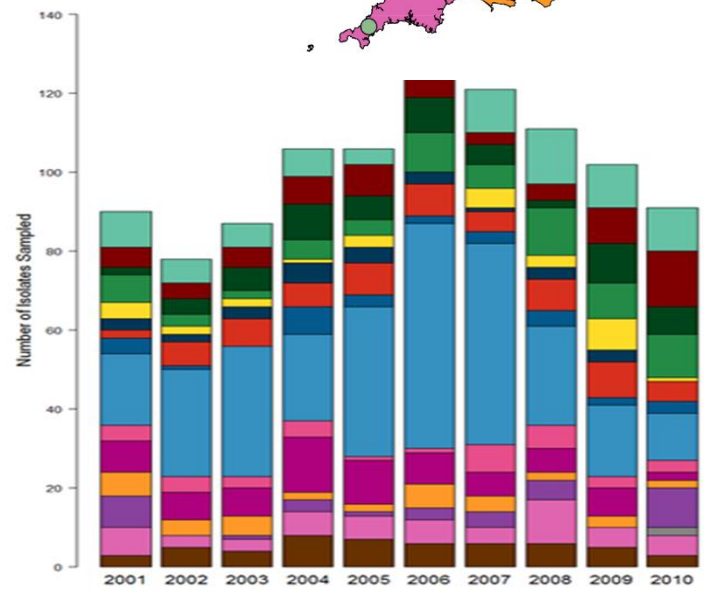
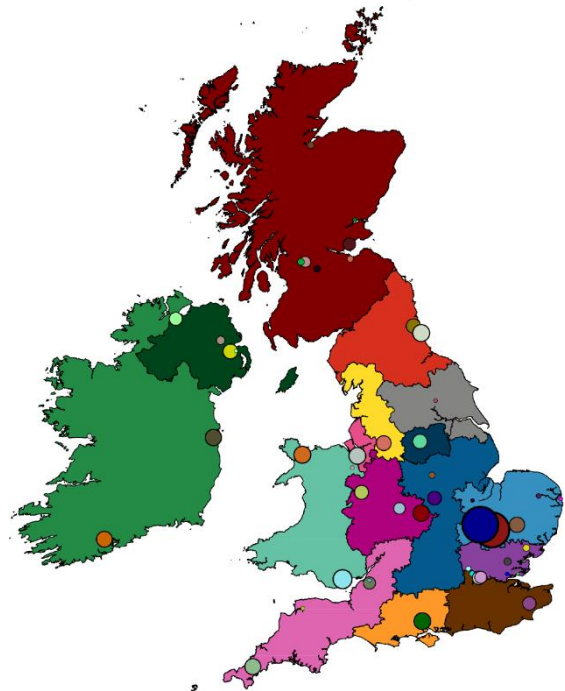
Retrospectively from clinical collections representing 46 hospitals in the UK and Ireland from 2001 to 2012.

Combined with data on patient transfers over the same period.

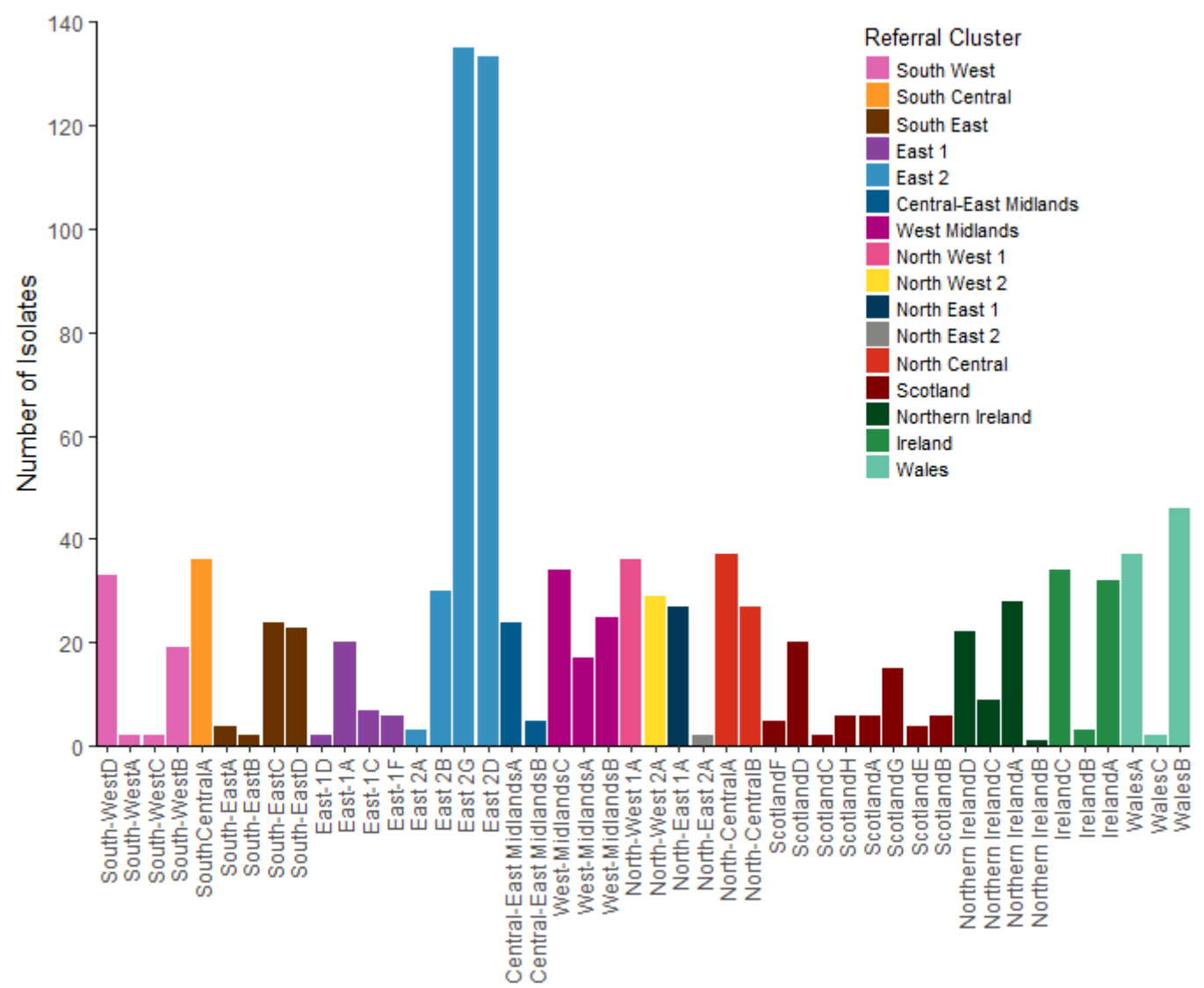
Colours represent “referral regions”

Defined by the amount of patient sharing.



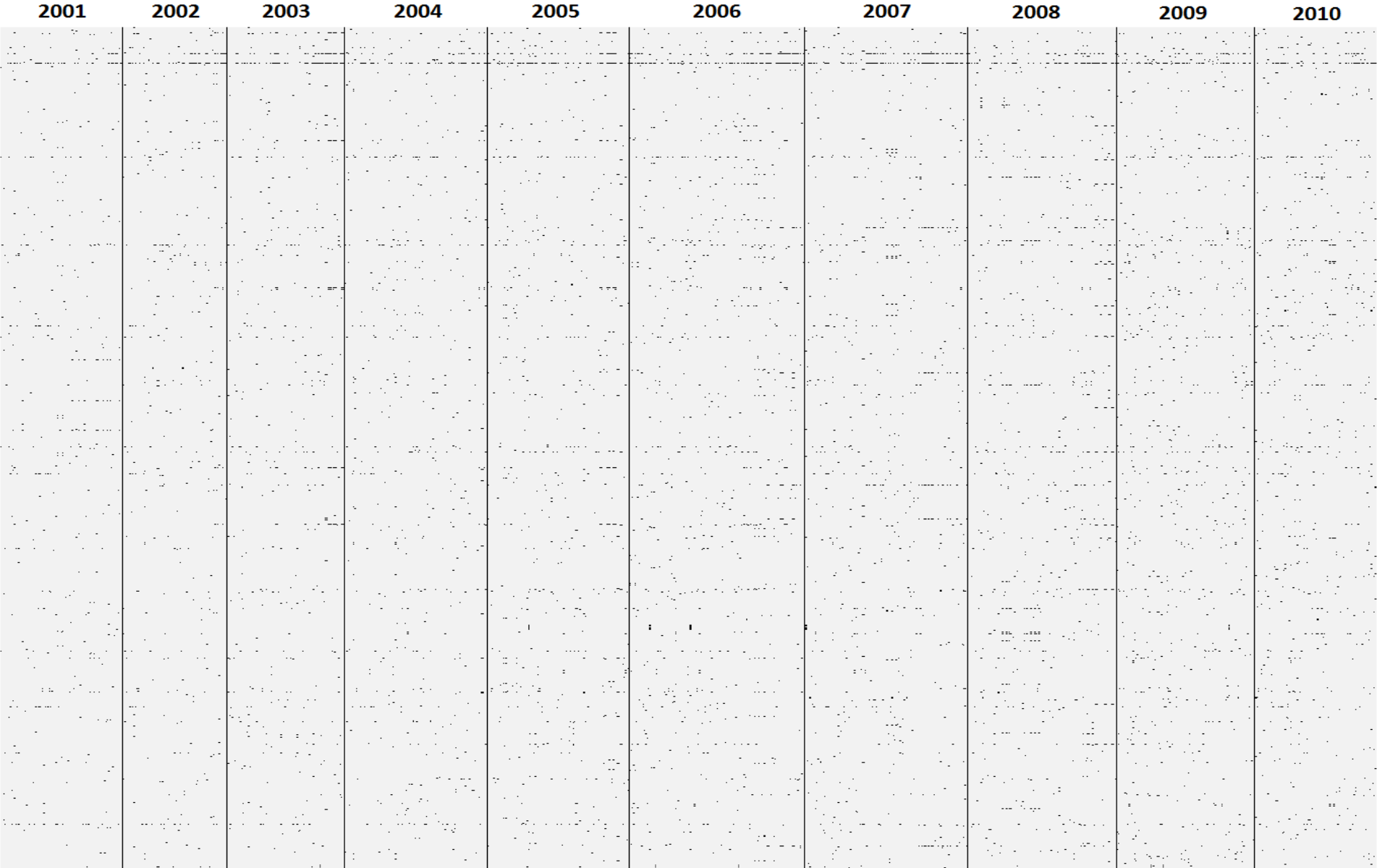


Sampling distribution between years



Sampling distribution between locations

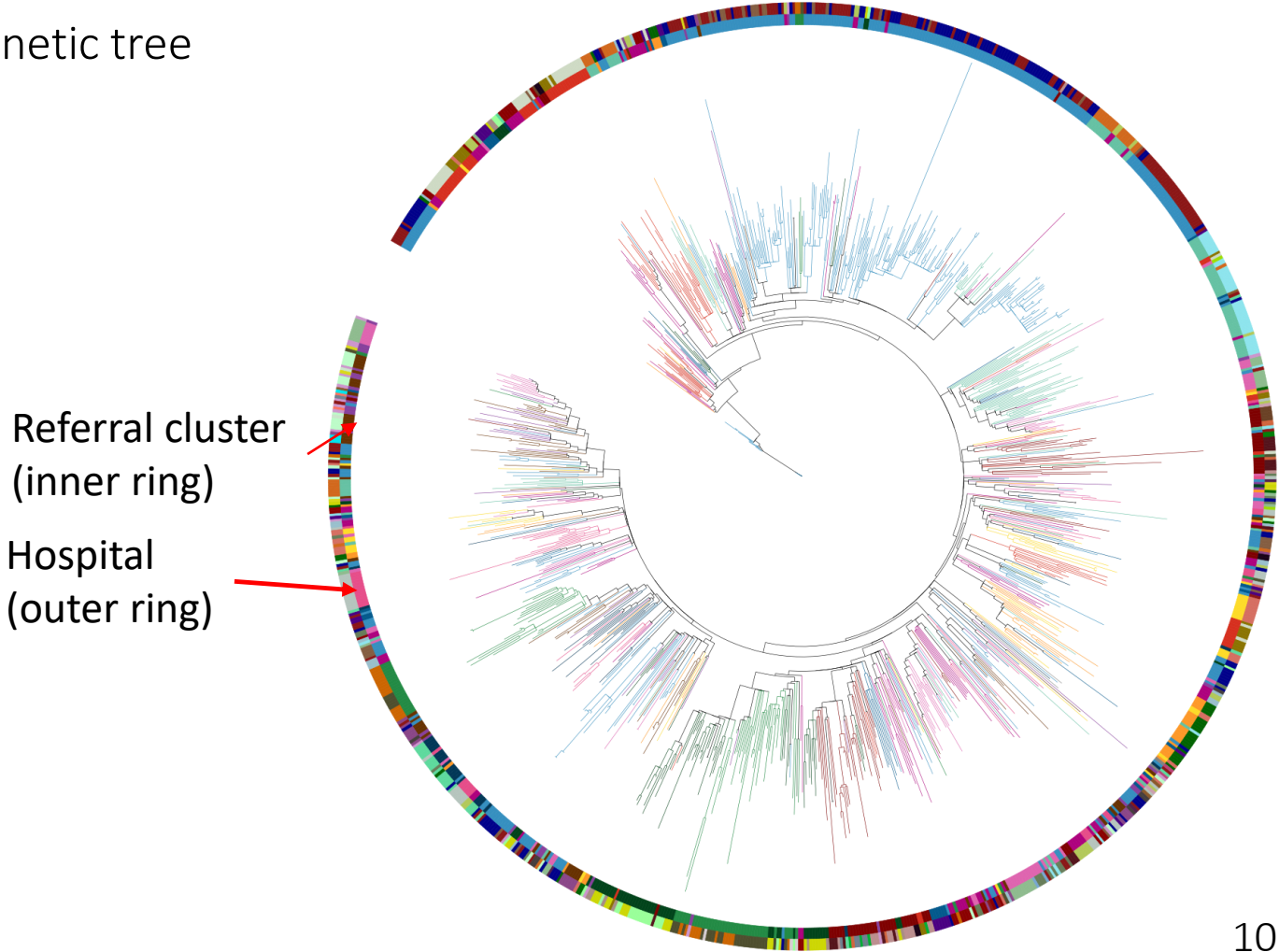
Database: 1022 isolates; 5469 loci; 46 hospitals



Isolate

SNP

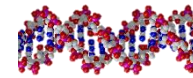
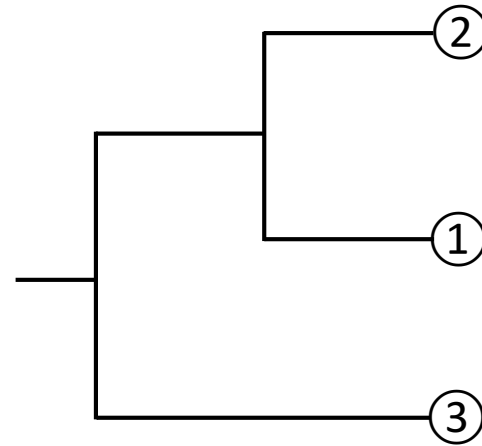
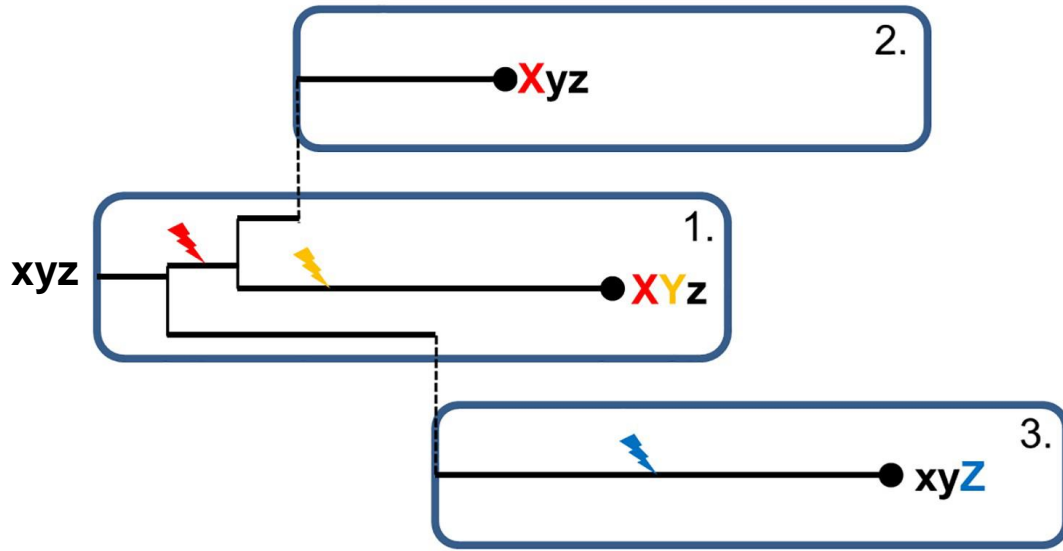
Maximum Likelihood phylogenetic tree



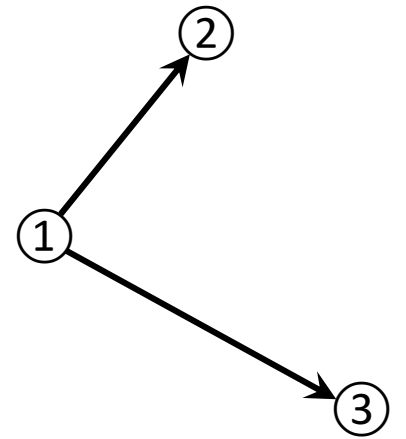
1022 leaves

Colour indicates location where sample was taken from

Object of real interest is the transmission tree



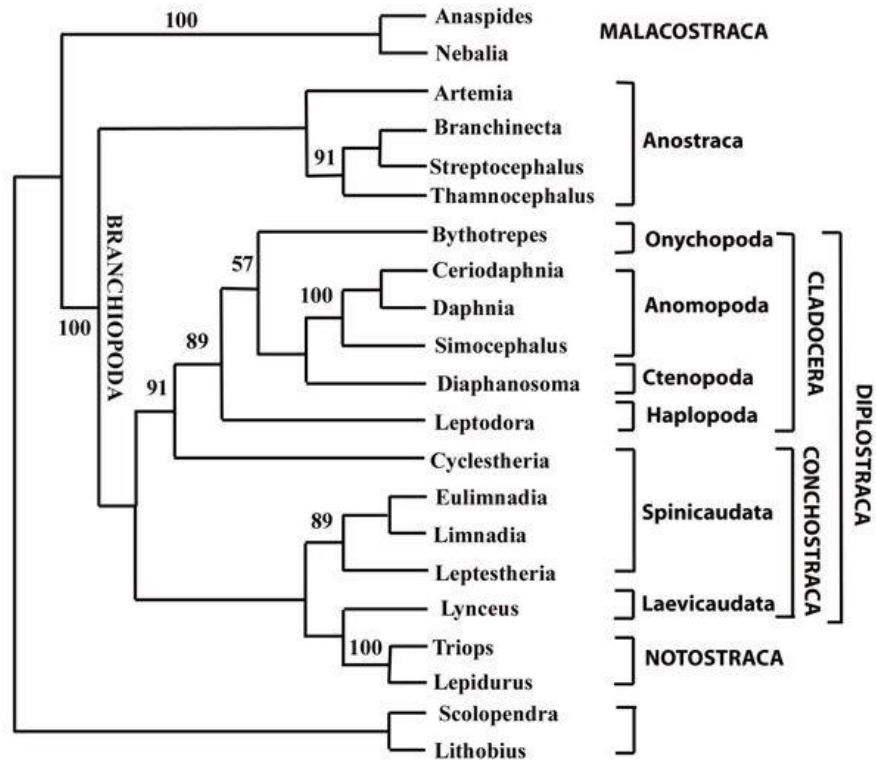
Phylogenetic tree



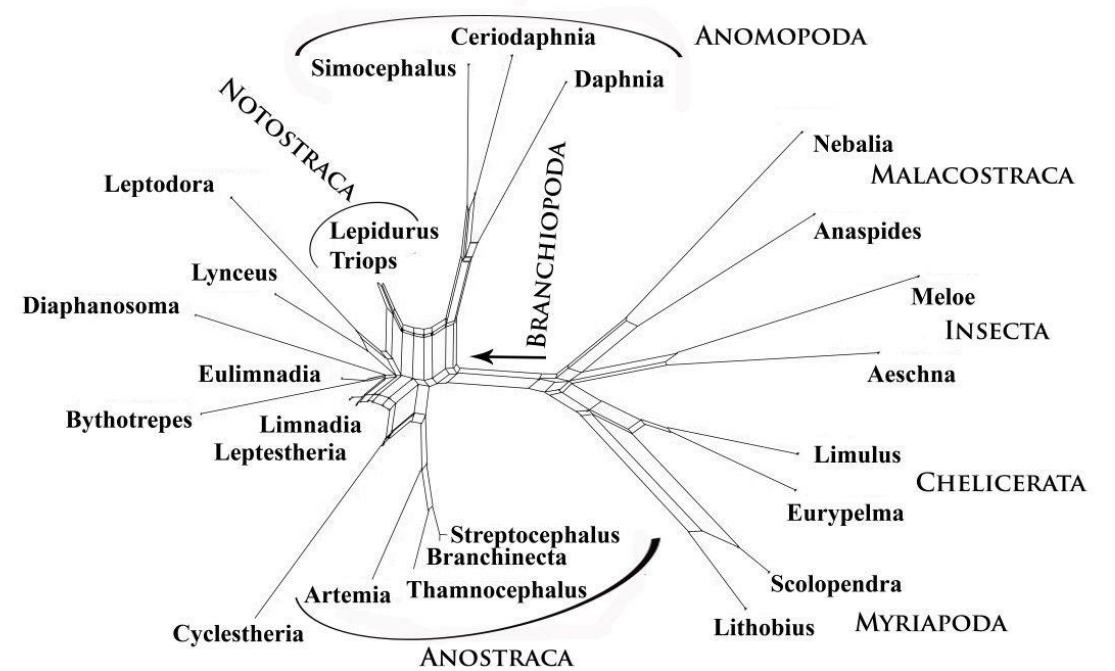
Transmission tree



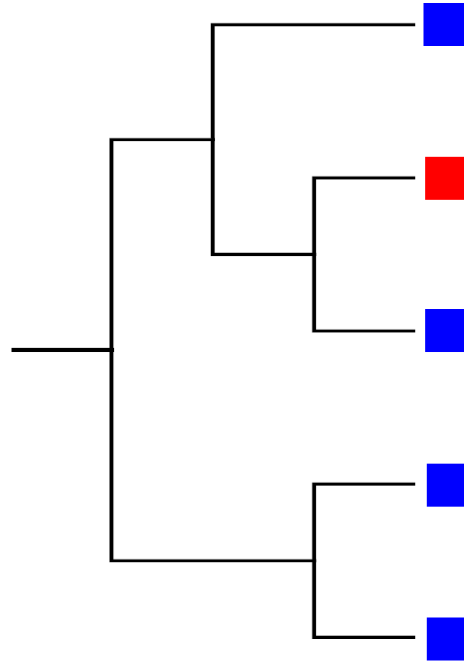
Phylogenetic tree...



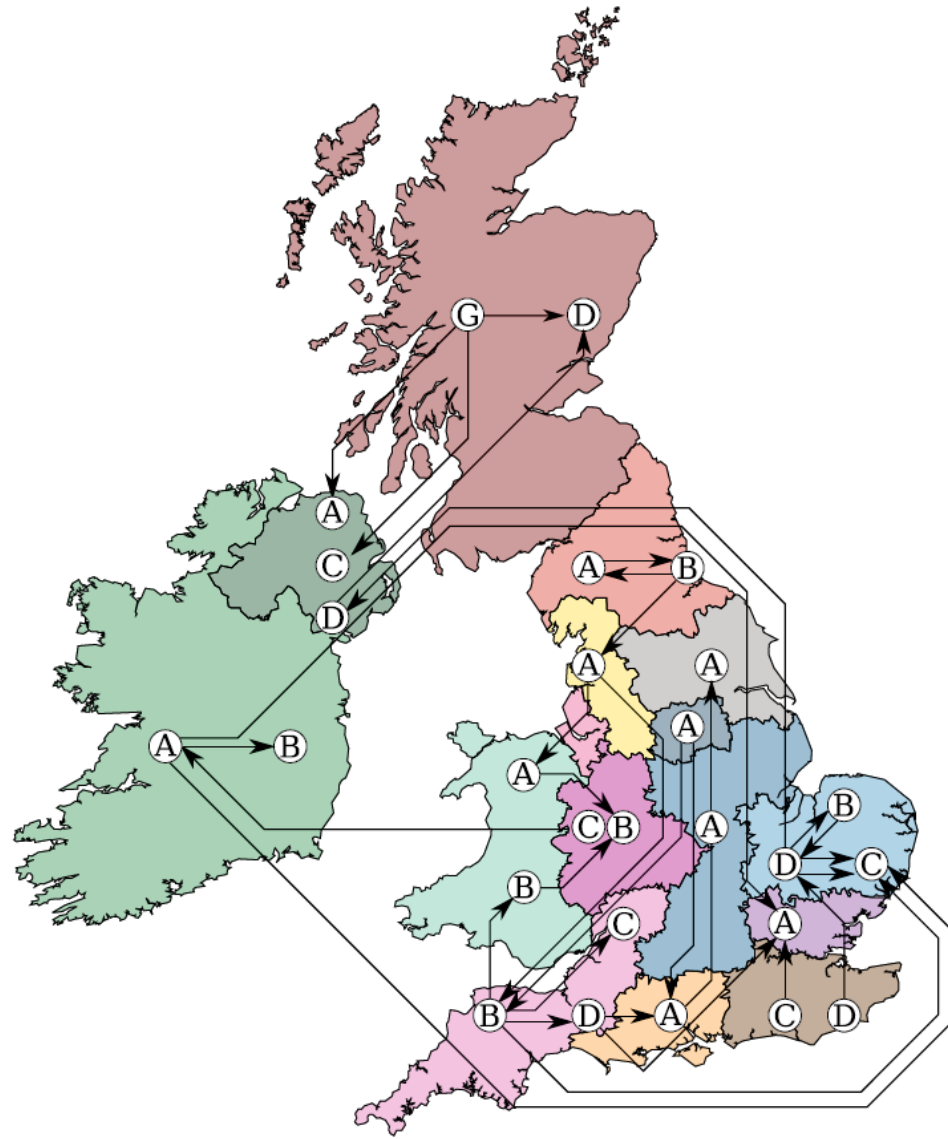
... or network?



Inferring inter-hospital transmission from the phylogenetic tree....

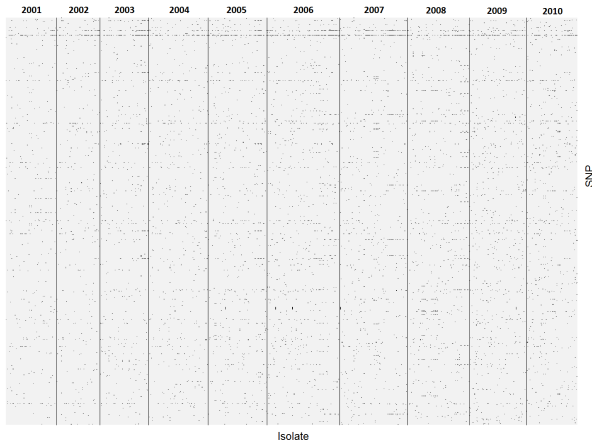


Transmission from the blue location to the red location

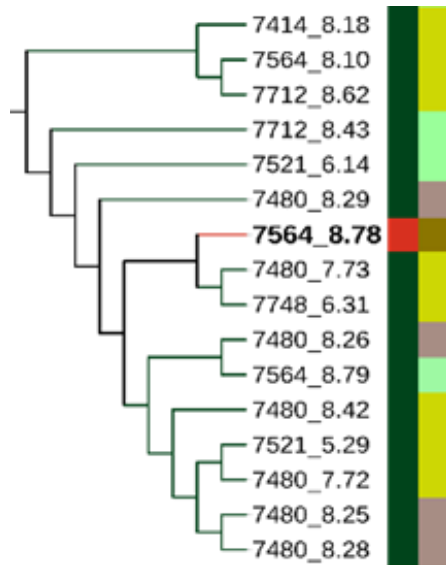
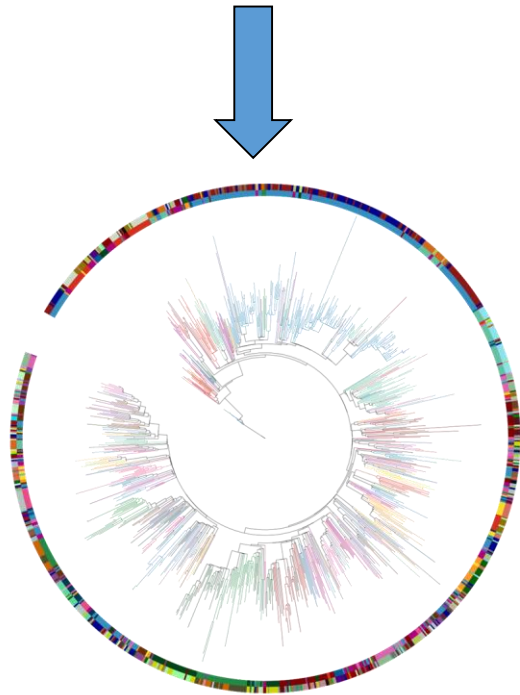


Donker et al, Microbial Genomics

# Q1: Real-time tree-independent genomic surveillance?



← Direct from here instead?



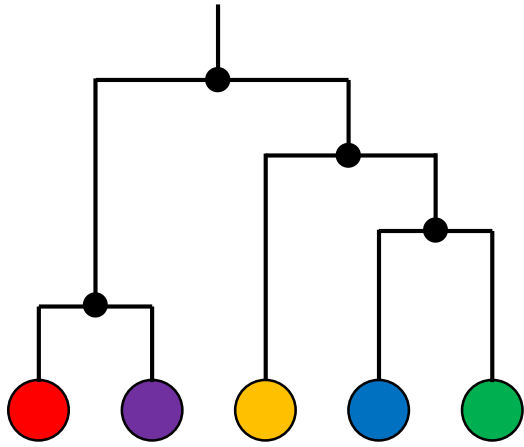
# Rapid statistical methods for inferring intra- and inter-hospital transmission of nosocomial pathogens from whole genome sequence data


Marianne Aspbury, James Sciberras, Jukka Corander, Sion C. Bayliss, Tjibbe Donker, Edward J. Feil, Richard James  
[e.feil@bath.ac.uk](mailto:e.feil@bath.ac.uk)

Milner Centre for Evolution  
Dept. Biology and Biochemistry,  
University of Bath, Bath, UK



Q2: How tree-like is a given network?



	$\begin{bmatrix} 0 & 1 & 4 & 4 & 4 \\ 1 & 0 & 4 & 4 & 4 \\ 4 & 4 & 0 & 3 & 3 \\ 4 & 4 & 3 & 0 & 2 \\ 4 & 4 & 3 & 2 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 4 & 1 & 1 & 1 \\ 4 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 2 & 2 \\ 1 & 1 & 2 & 0 & 3 \\ 1 & 1 & 2 & 3 & 0 \end{bmatrix}$
---	---	---

