Challenges and opportunities for whole-genome sequencing–based surveillance of antibiotic resistance

Prof. Willem van Schaik
Professor in Microbiology and Infection

Institute of Microbiology and Infection
University of Birmingham, United Kingdom
w.vanschaik@bham.ac.uk
🐦 WvSchaik

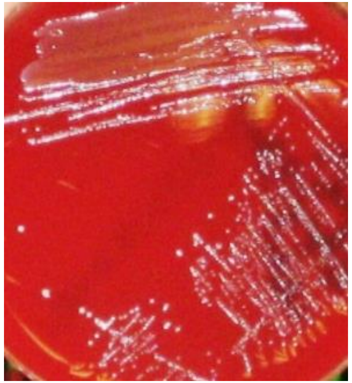
UNIVERSITY OF
BIRMINGHAM

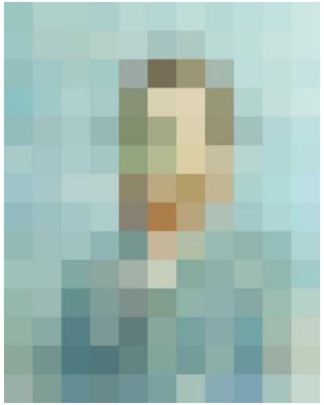
Sequence-based pathogen surveillance

use sequence data to infer relatedness of strains

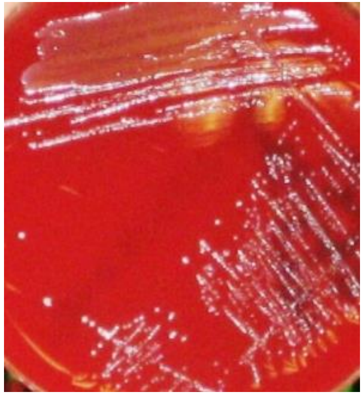
MLST
(allelic profile of 5 – 7 housekeeping genes)

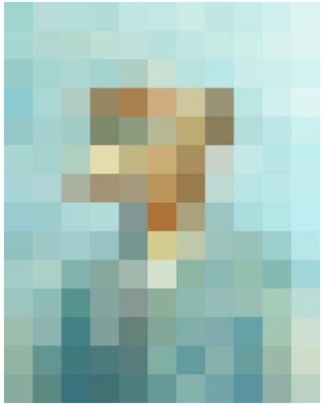
Strain A





Strain B





Low resolution

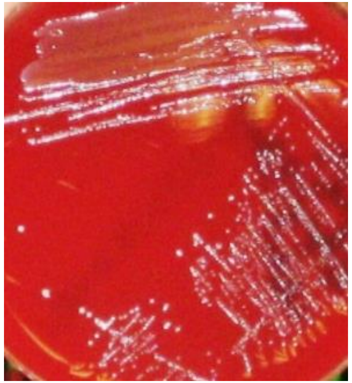
Differences may be apparent, but not clear

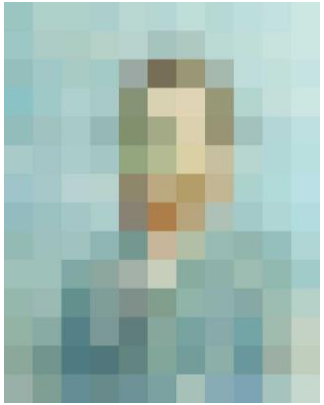
Provides a nomenclature to describe clones (e.g. *E. coli* ST131)

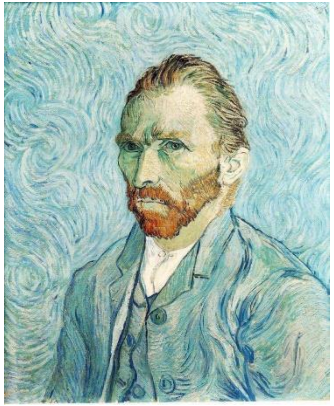
MLST

WGS

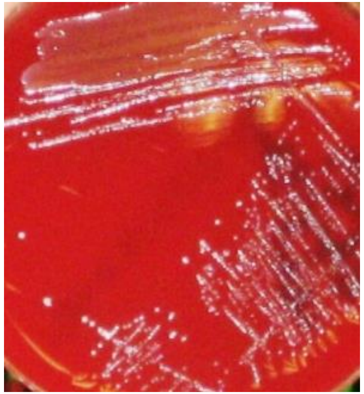
Strain A

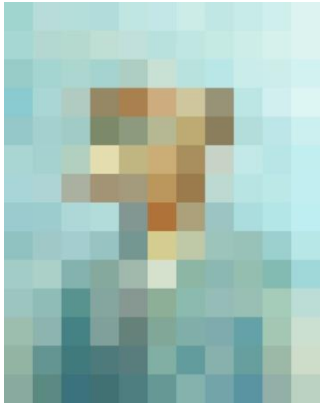


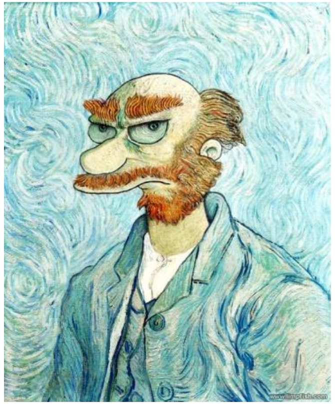




Strain B







High resolution

Core genome phylogeny (SNP-based, core genome MLST)

Whole genome phylogeny (whole genome MLST)

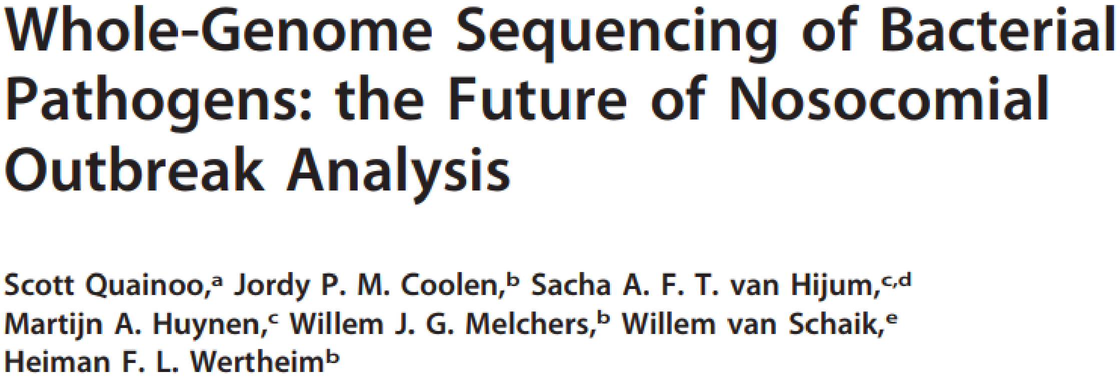
Many tools

Which ones to choose?

Or use the tools developed by Torsten Seemann (Univ of Melbourne)

AMERICAN
SOCIETY FOR
MICROBIOLOGY

Clinical Microbiology
Reviews®



Whole-Genome Sequencing of Bacterial Pathogens: the Future of Nosocomial Outbreak Analysis

Scott Quainoo,[a] Jordy P. M. Coolen,[b] Sacha A. F. T. van Hijum,[c,d]
Martijn A. Huynen,[c] Willem J. G. Melchers,[b] Willem van Schaik,[e]
Heiman F. L. Wertheim[b]

Clin Microbiol Rev 30:1015–1063

But there is more in a whole genome sequence…

Antibiotic resistance genes

Presence of mobile genetic elements

Outbreak vancomycin-resistant *Enterococcus faecium*

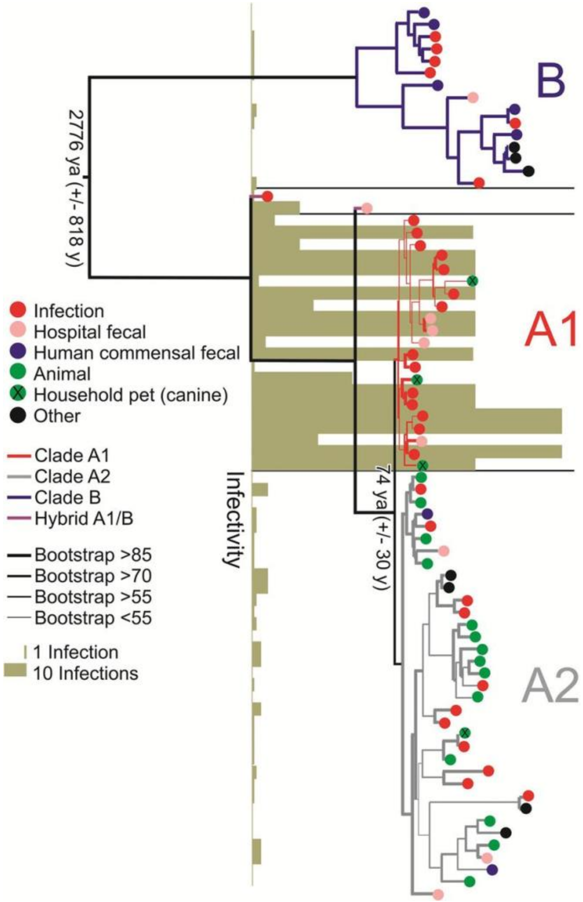
Heartlands Hospital, Birmingham, United Kingdom

Patients colonised and infected with vancomycin-resistant *Enterococcus faecium*

Clonal outbreak of strains?

Spread of identical vancomycin resistance plasmids?

Vancomycin-resistant *Enterococcus faecium*



"Human commensal"

"Clinical isolates"

"Animal isolates"

Maximum likelihood tree: concatenated alignment of 1344 core genes in 73 *E. faecium* genomes

Lebreton, van Schaik, McGuire *et al.*, 2013. Mbio 4:e00534-13

WHO priority list of pathogens

Priority 1: CRITICAL#

Acinetobacter baumannii, carbapenem-resistant

Pseudomonas aeruginosa, carbapenem-resistant

Enterobacteriaceae*, carbapenem-resistant, 3rd generation cephalosporin-resistant

Priority 2: HIGH

Enterococcus faecium, vancomycin-resistant

Staphylococcus aureus, methicillin-resistant, vancomycin intermediate and resistant

Helicobacter pylori, clarithromycin-resistant

Campylobacter, fluoroquinolone-resistant

Salmonella *spp.*, fluoroquinolone-resistant

Neisseria gonorrhoeae, 3rd generation cephalosporin-resistant, fluoroquinolone-resistant

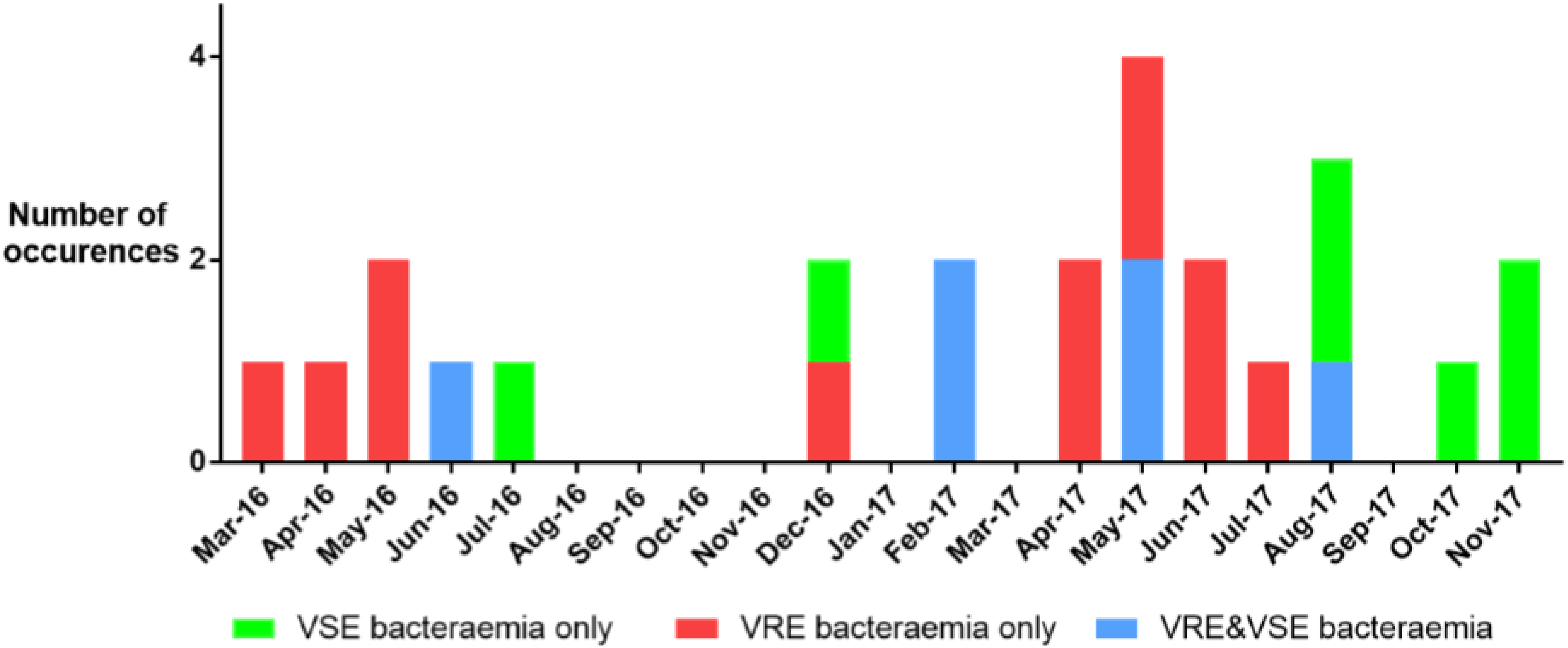
Priority 3: MEDIUM

Streptococcus pneumoniae, penicillin-non-susceptible

Haemophilus influenzae, ampicillin-resistant

Shigella *spp.*, fluoroquinolone-resistant

VRE Outbreak Heartlands Hospital



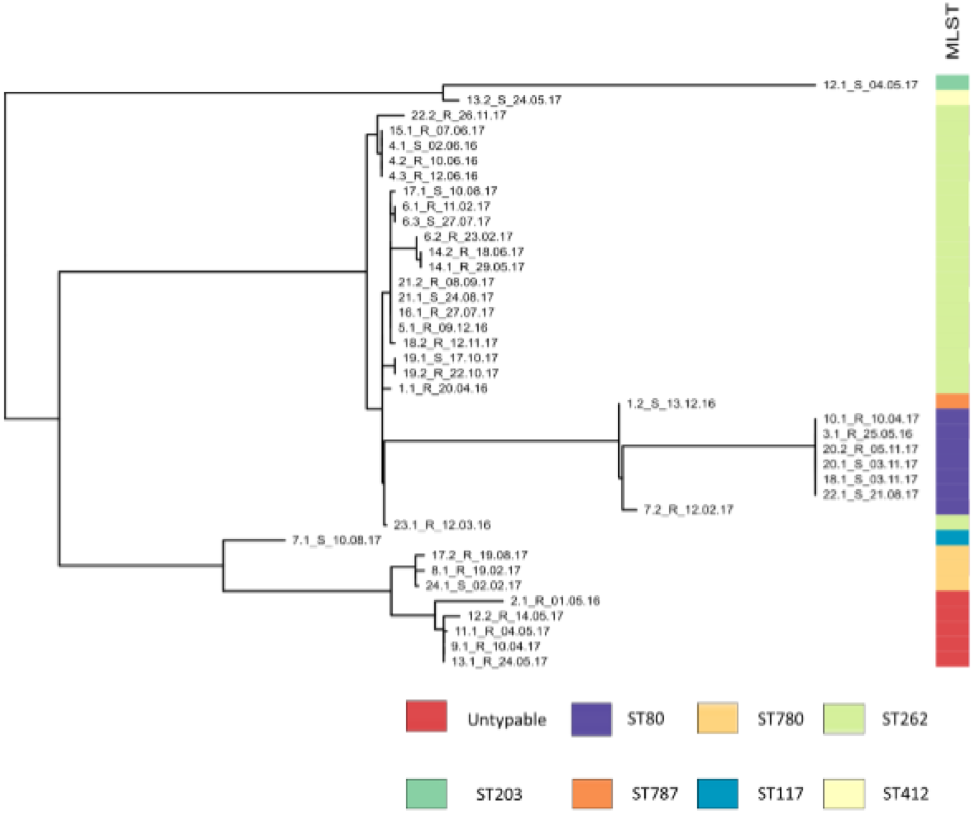
VRE Outbreak Heartlands Hospital

Specimen	Resistant	Sensitive	Total
Blood Culture	21	13	34
Rectal sample	5	0	5
Total	26	13	39

26 isolates VRE, 13 isolates VSE

Sequencing on Illumina platform, 2 x 150 nt

VRE Outbreak Heartlands Hospital



Identify SNPs against reference (*E. faecium* E745) using Snippy 3.2

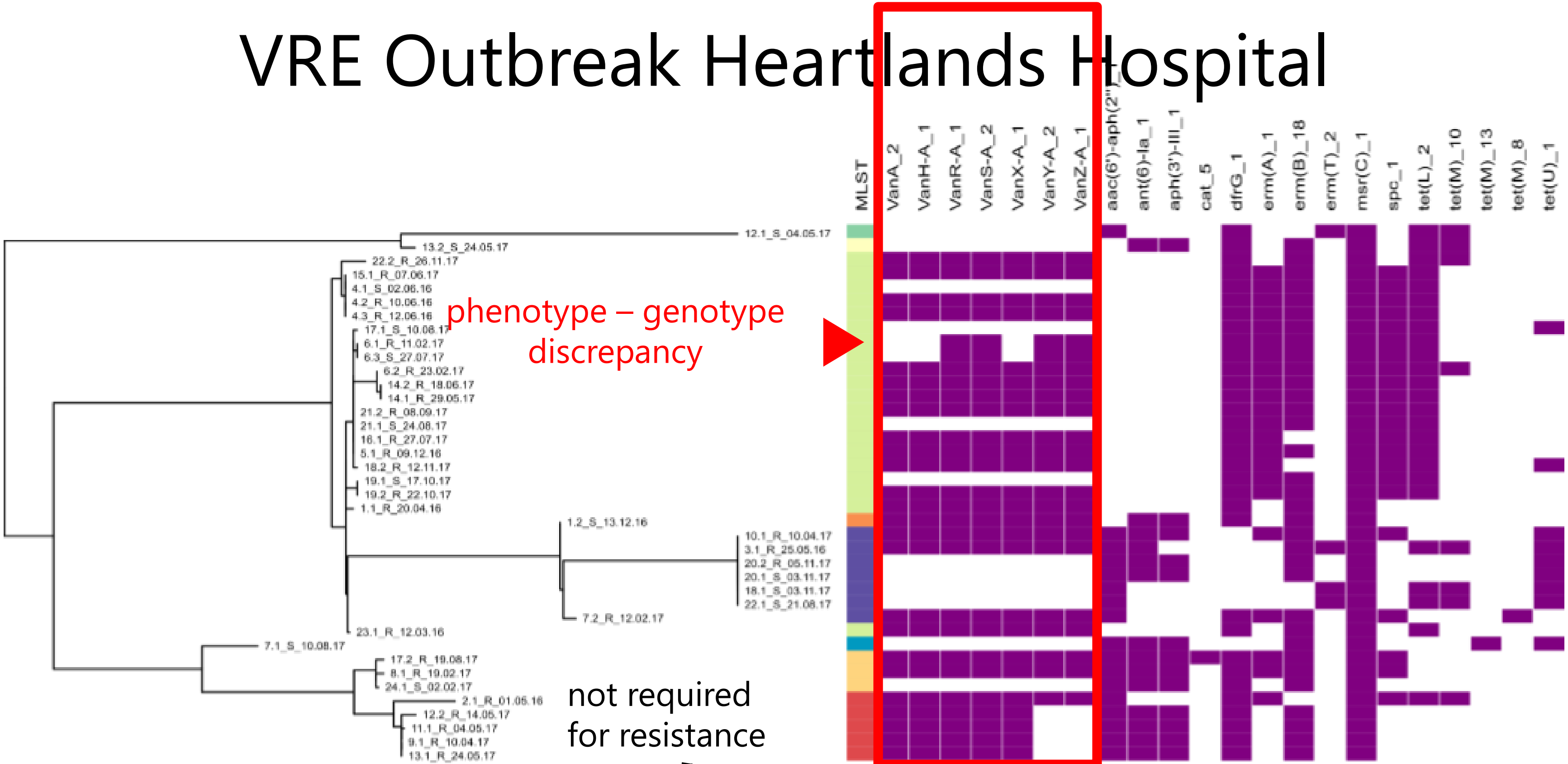
Recombination filtering using Gubbins

Maximum likelihood tree

Sequence types (MLST) determined using mlst 2.8; 5 strains untypable due to missing *pstS* allele

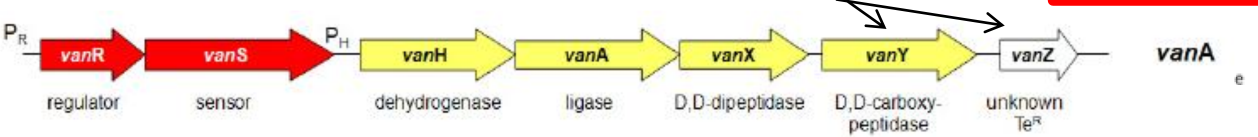
Outbreak is multi-clonal

VRE Outbreak Heartlands Hospital



phenotype – genotype discrepancy

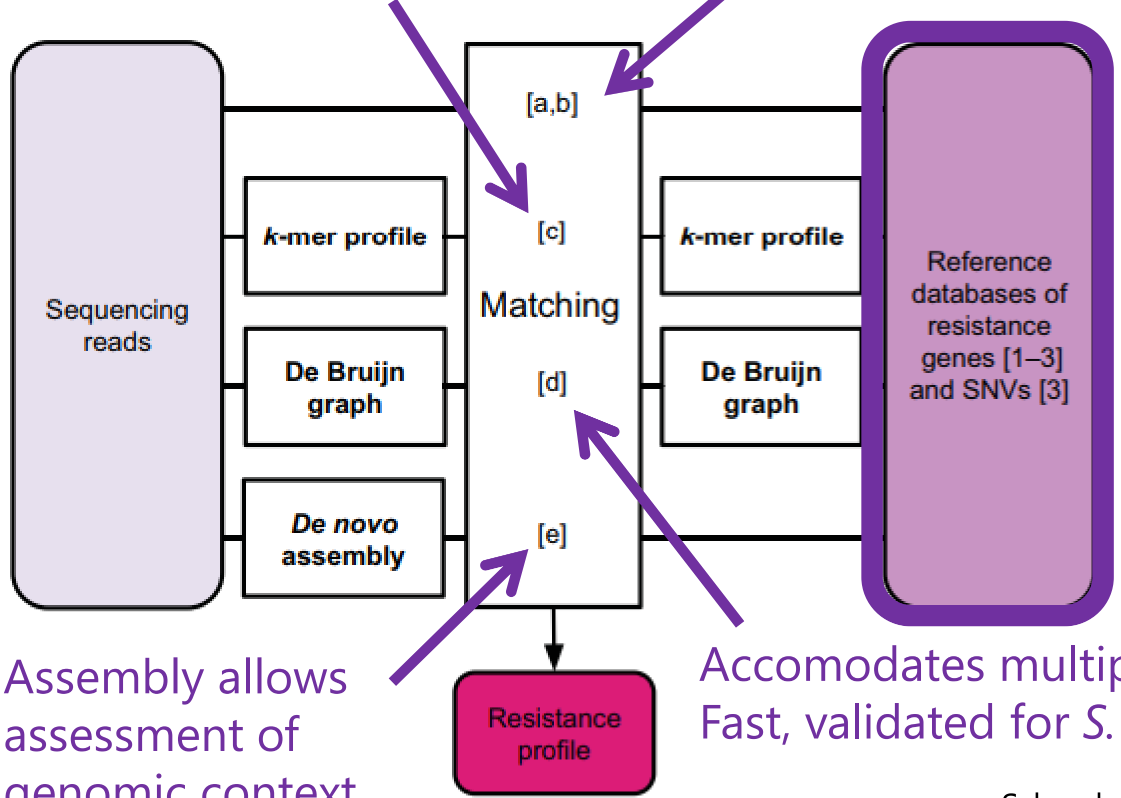
not required for resistance



Presence of antibiotic resistance genes identified using Abricate 0.7, using ResFinder database

Fastest option

Direct mapping of reads: no loss of data



[a] BWA implemented in SEAR
[b] Bowtie2 implemented in SRST2
[c] *k*-merResistance
[d] Mykrobe predictor
[e] BLAST (several tools, including Abricate)

Assembly allows assessment of genomic context

Accomodates multiple bacteria in sample
Fast, validated for *S. aureus* and *M. tuberculosis*

Schurch & Van Schaik, 2017. Ann. N Y Acad Sci 1388:108-120

What is an antibiotic resistance gene?

"a gene that confers resistance to an antibiotic in an otherwise susceptible microbial host"

✓ Horizo ally acquired resistance genes: e.g. *bla*KPC, *ermB, vanA*

associated with mobile genetic elements

Mutations in housekeeping genes that confer ✓ ance

e.g. target modification, porin mutation

Conse ✗ ve genes that make a species intrinsically resistant

e.g. efflux pumps

Antibiotic Resistance Gene Databases

ARDB: no updates, many intrinsic resistance genes: do not use

CARD: frequently updated, based on ARDB. Contains some intrinsic resistance genes; database for 'resistance SNPs' but can give false positives

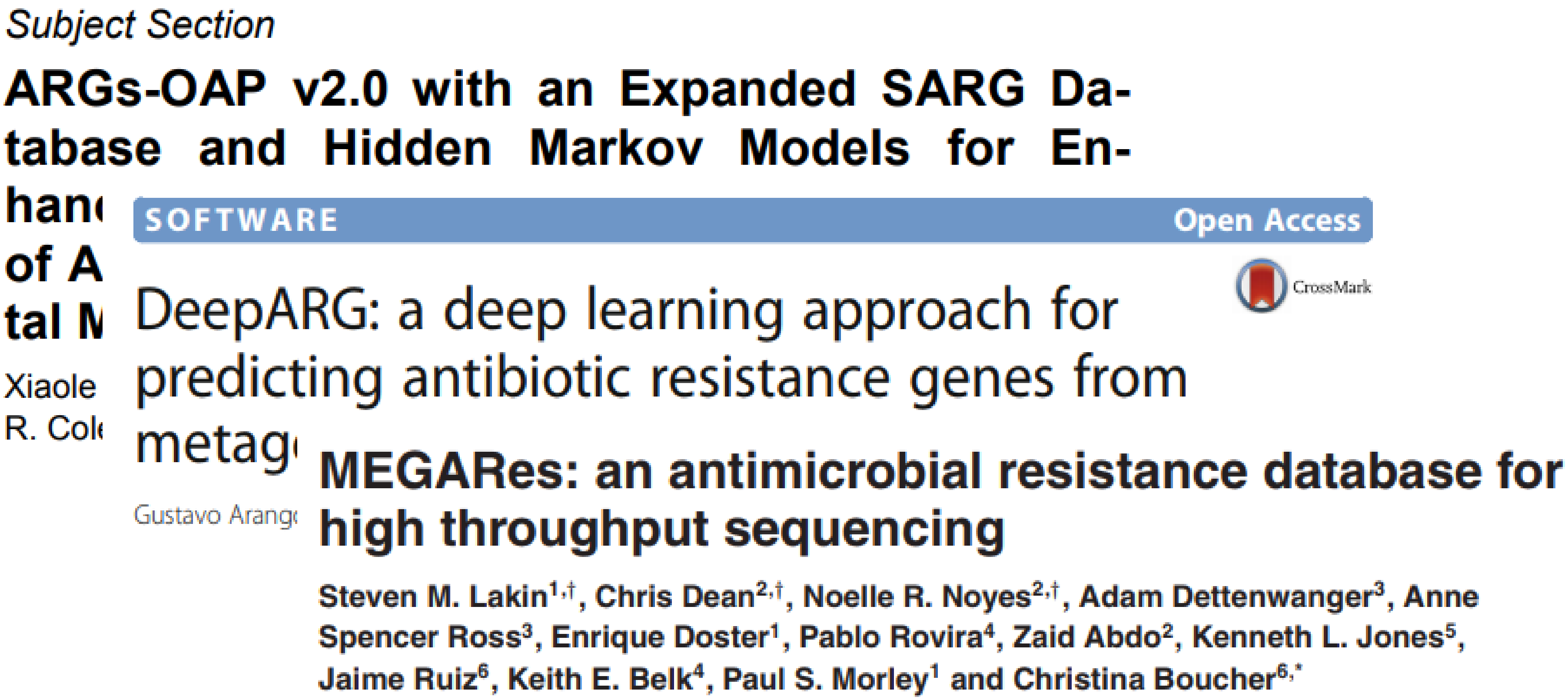
ResFinder

ARG-ANNOT

} Frequently updated. Acquired resistance genes only, small number of intrinsic resistance genes.

ResFams: prediction of resistance genes from metagenomes, many false positives or unvalidated genes

Antibiotic Resistance Gene Databases



Subject Section

ARGs-OAP v2.0 with an Expanded SARG Database and Hidden Markov Models for En-
han
of A
tal M

Xiaole
R. Col

SOFTWARE Open Access

DeepARG: a deep learning approach for predicting antibiotic resistance genes from metag

Gustavo Arang

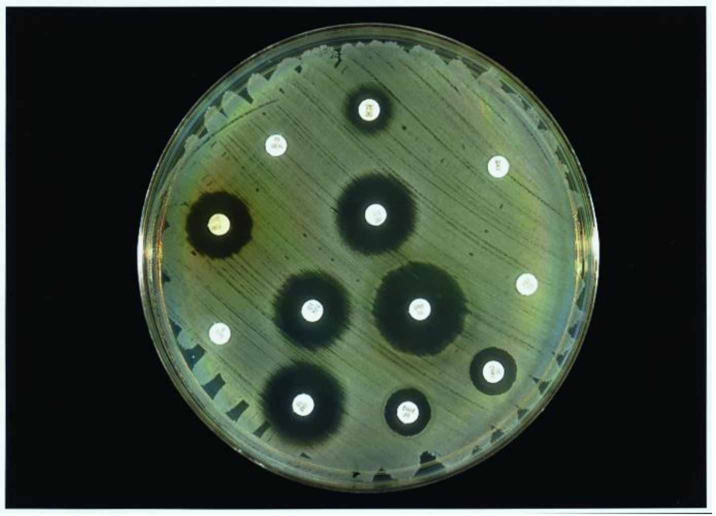
MEGARes: an antimicrobial resistance database for high throughput sequencing

Steven M. Lakin[1,†], Chris Dean[2,†], Noelle R. Noyes[2,†], Adam Dettenwanger[3], Anne Spencer Ross[3], Enrique Doster[1], Pablo Rovira[4], Zaid Abdo[2], Kenneth L. Jones[5], Jaime Ruiz[6], Keith E. Belk[4], Paul S. Morley[1] and Christina Boucher[6,*]

Antibiotic Resistance Genes

When reporting antibiotic resistance genes….

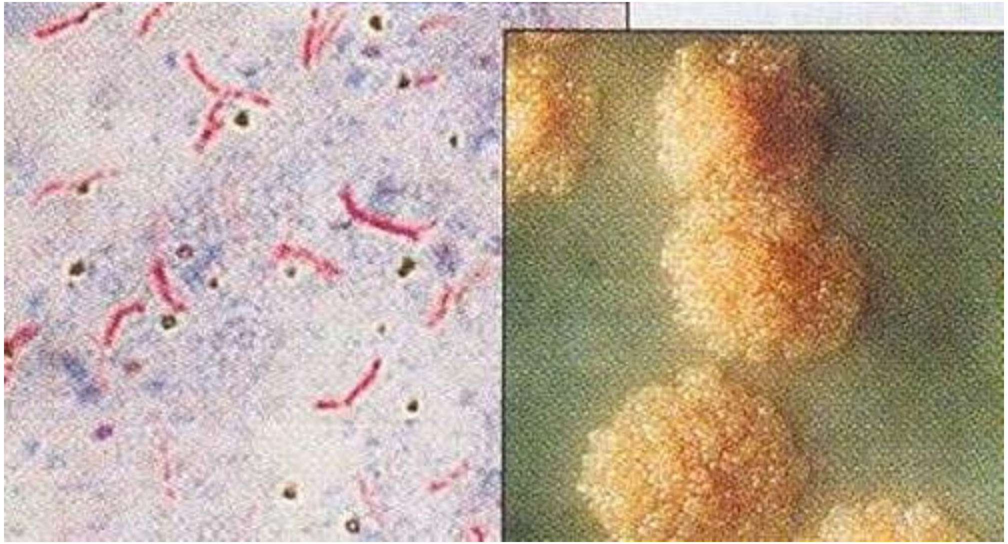
Sequence-based diagnostics of antibiotic resistance genes?



?

Sequence-based diagnostics of antibiotic resistance genes?

Mycobacterium tuberculosis



Very slow growth

No horizontal gene transfer

Large collection of strains + resistance phenotypes

WGS accurately predicts susceptibility 93.6% – 99.0% to four key antibiotics

CRyPTIC Consortium and the 100,000 Genomes Project. New Engl J Med. 2018 379:1403

Sequence-based diagnostics of antibiotic resistance genes?

Common causes of bloodstream infections

Escherichia coli
Staphylococcus aureus
Enterococcus


beLibrary © Ho and Chew

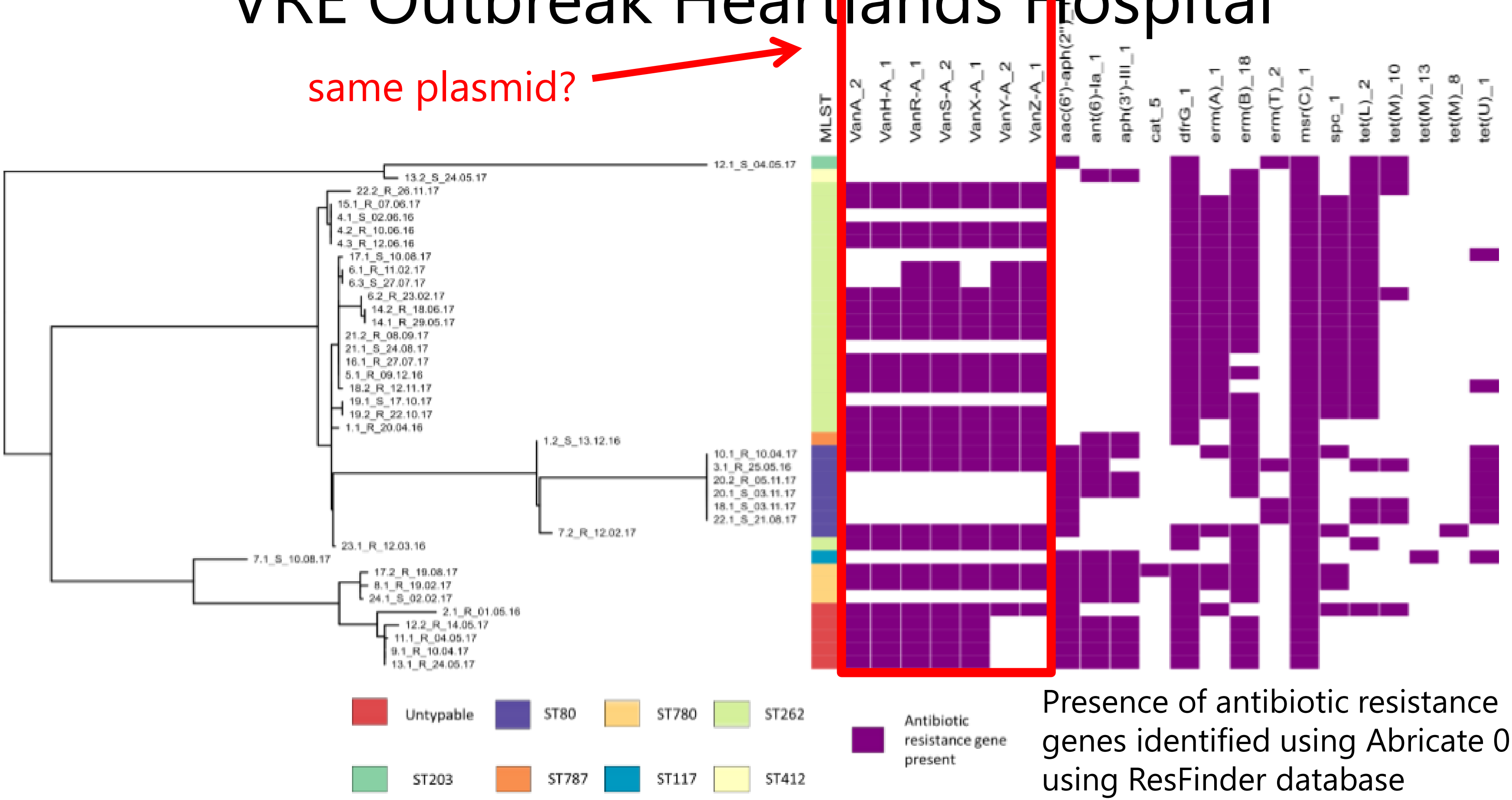
Fast growth

Horizontal gene transfer

Emergence of new resistance determinants (e.g. *mcr*)

Fast WGS (e.g. of bloodstream cultures) for screening while susceptibility testing is ongoing

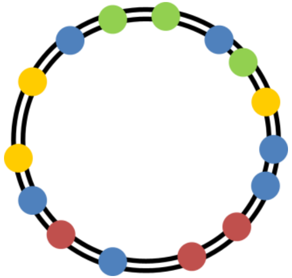
VRE Outbreak Heartlands Hospital



same plasmid?

Presence of antibiotic resistance genes identified using Abricate 0.7, using ResFinder database

Prediction of Plasmids from Illumina Data



Sequence repeats in
plasmid sequences
complicates assembly

42 completely assembled genomes
12 genera
148 plasmids (1.5 – 339 kbp)

?

Illumina + PacBio sequences
available

Arredondo-Alonso *et al.*, 2017. Microb. Genomics 3:128

Prediction of Plasmids from Illumina Data

or

<u>Comparison of completely automated approaches</u>

PlasmidFinder: web-based tool to detect replicon sequences in contigs *Carattoli et al., Antimicrob Agents Chemother. 2014 58:3895*

cBar: detects differences in k-mer composition between plasmids and chromosomes *Zhou & Xu Bioinformatics 2010 26:2051*

Recycler: extracts circular sequences from De Bruijn assembly graphs *Rozov et al. Bioinformatics 2017 33:475*

PlasmidSPAdes: plasmids have different coverage (higher copy number) than chromosome *Antipov et al. Bioinformatics 2016 32:3380*

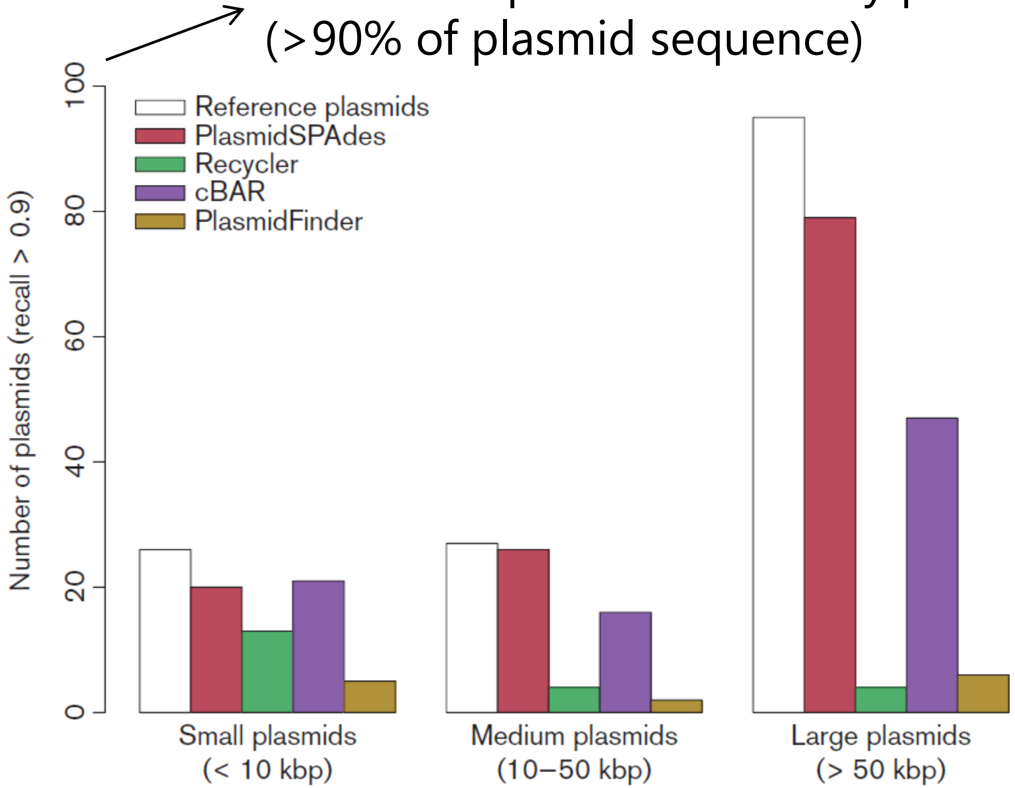

On the (im)possibility of reconstructing plasmids from whole-genome short-read sequencing data

Sergio Arredondo-Alonso,[1] Rob J. Willems,[1] Willem van Schaik[1,2] and Anita C. Schürch[1,*]

Arredondo-Alonso *et al.*, 2017. Microb. Genomics 3:128

Prediction of Plasmids from Illumina Data

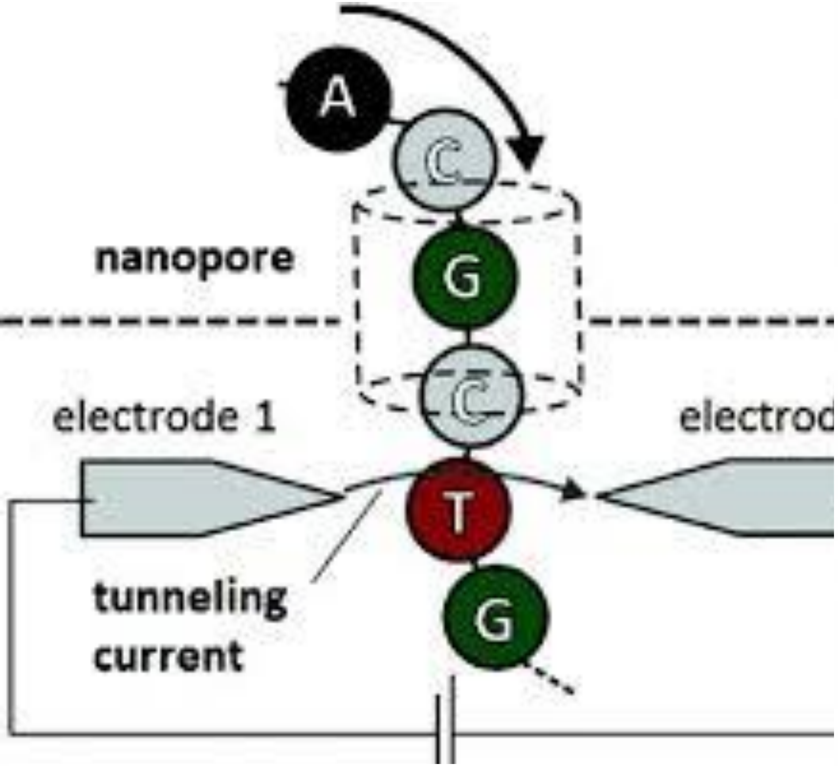
Number of plasmids correctly predicted
(>90% of plasmid sequence)



Results are variable

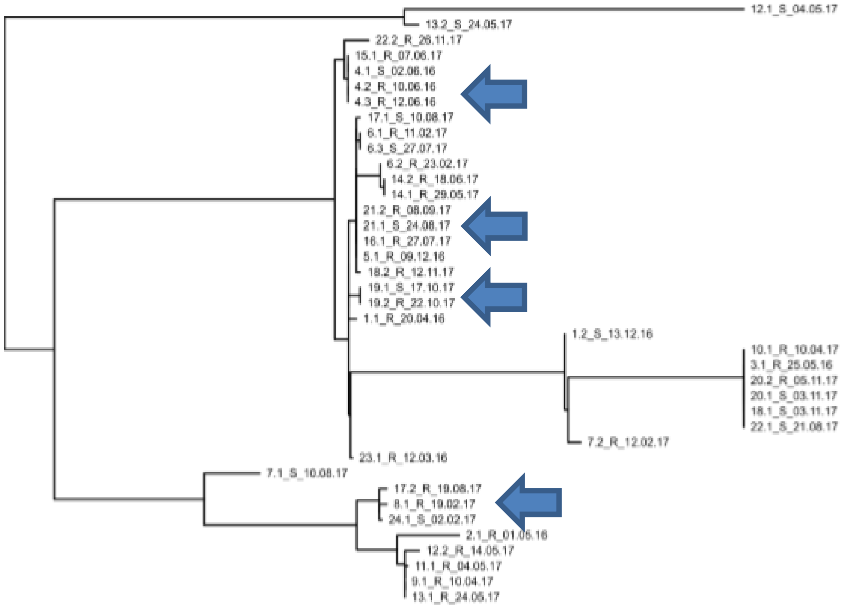
Good performance PlasmidSPAdes, but all plasmids in single bin

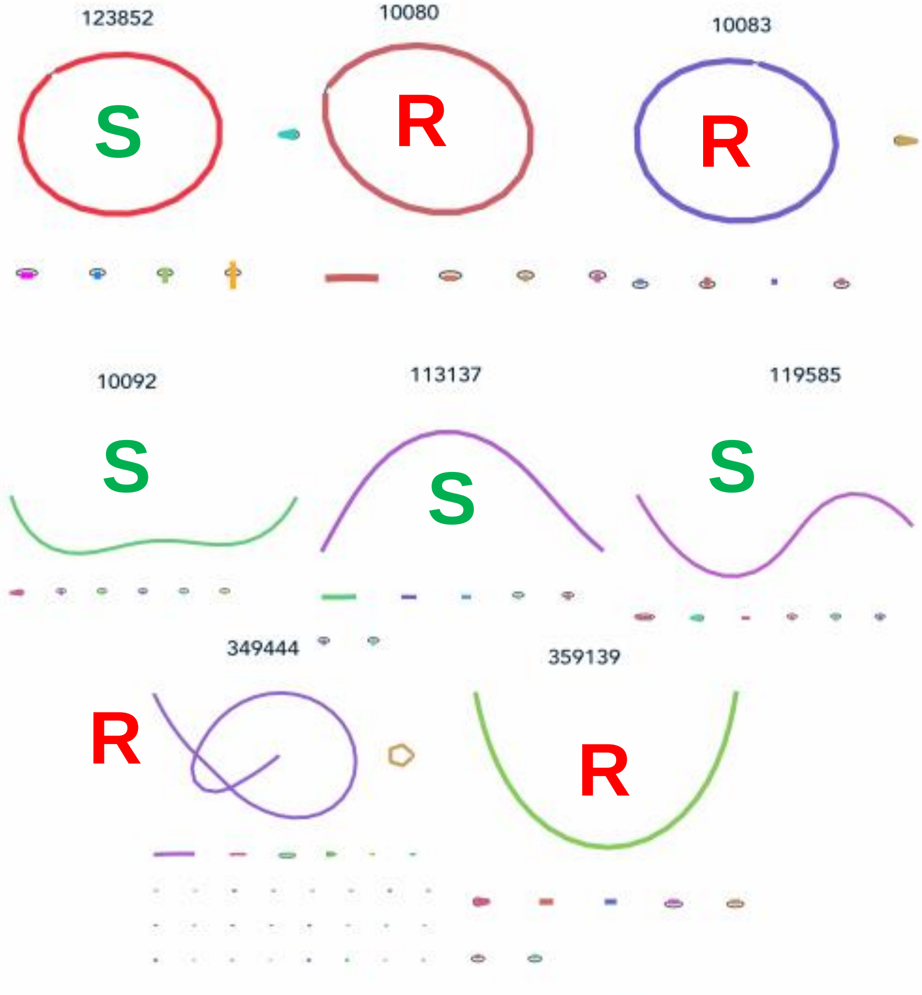
Arredondo-Alonso *et al.*, 2017. Microb. Genomics 3:128



Oxford NANOPORE Technologies

Long reads can cover repeats



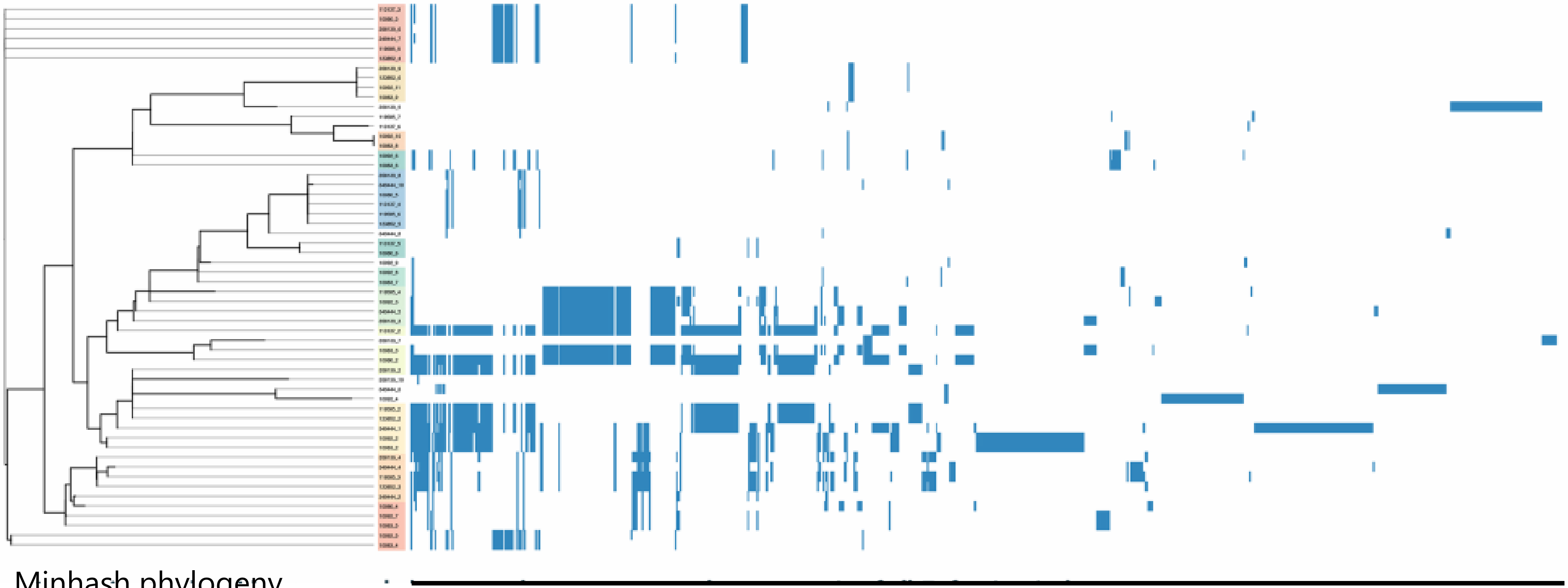


Four susceptible/resistant strain-pairs
<5 SNPs core genome
Acquisition of vancomycin resistance genes?

Steven Dunn

Hybrid assemblies (Illumina + MinION): Unicycler

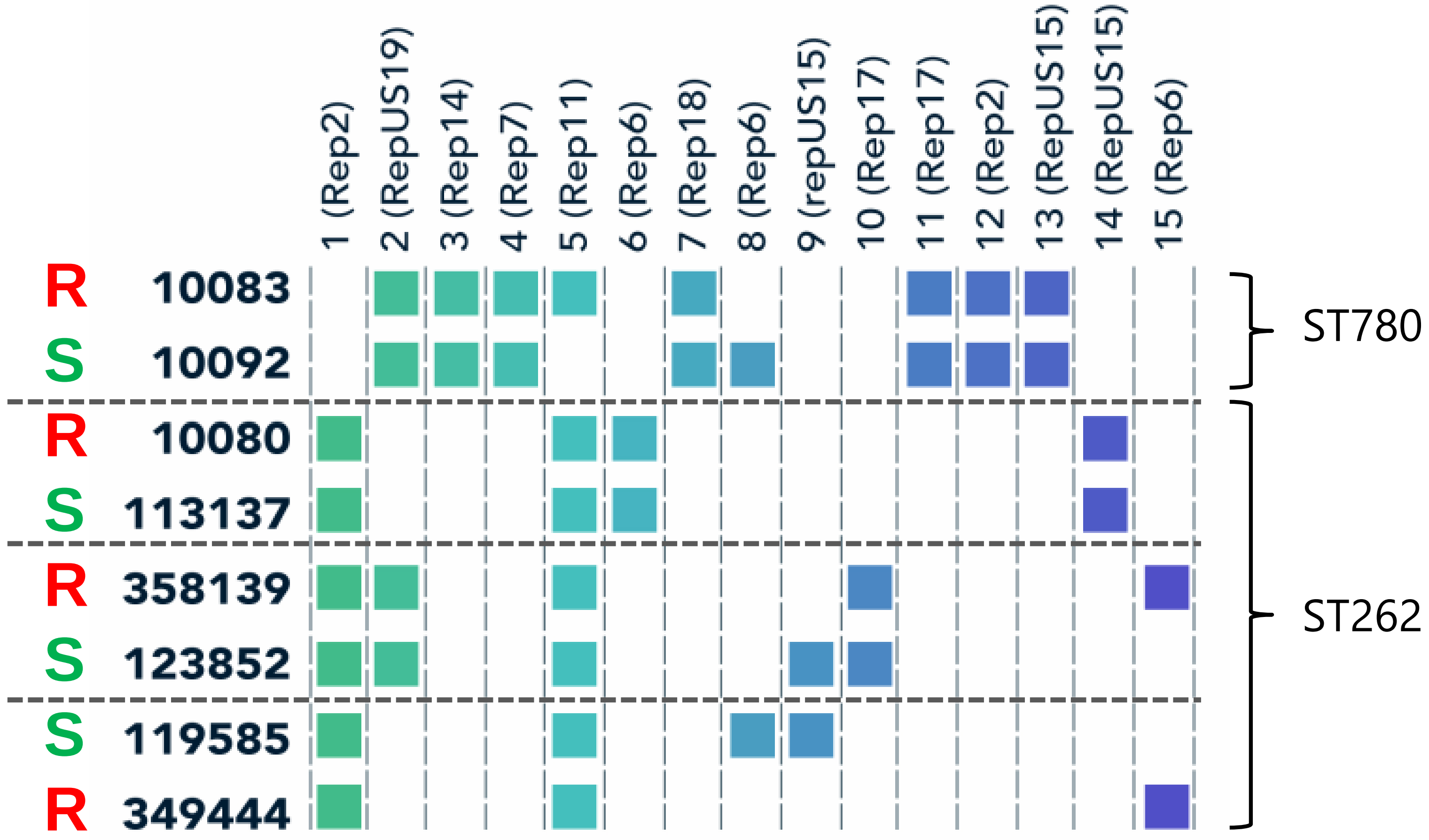
Phylogeny of non-chromosomal contigs



Minhash phylogeny
visualised with Mashtree

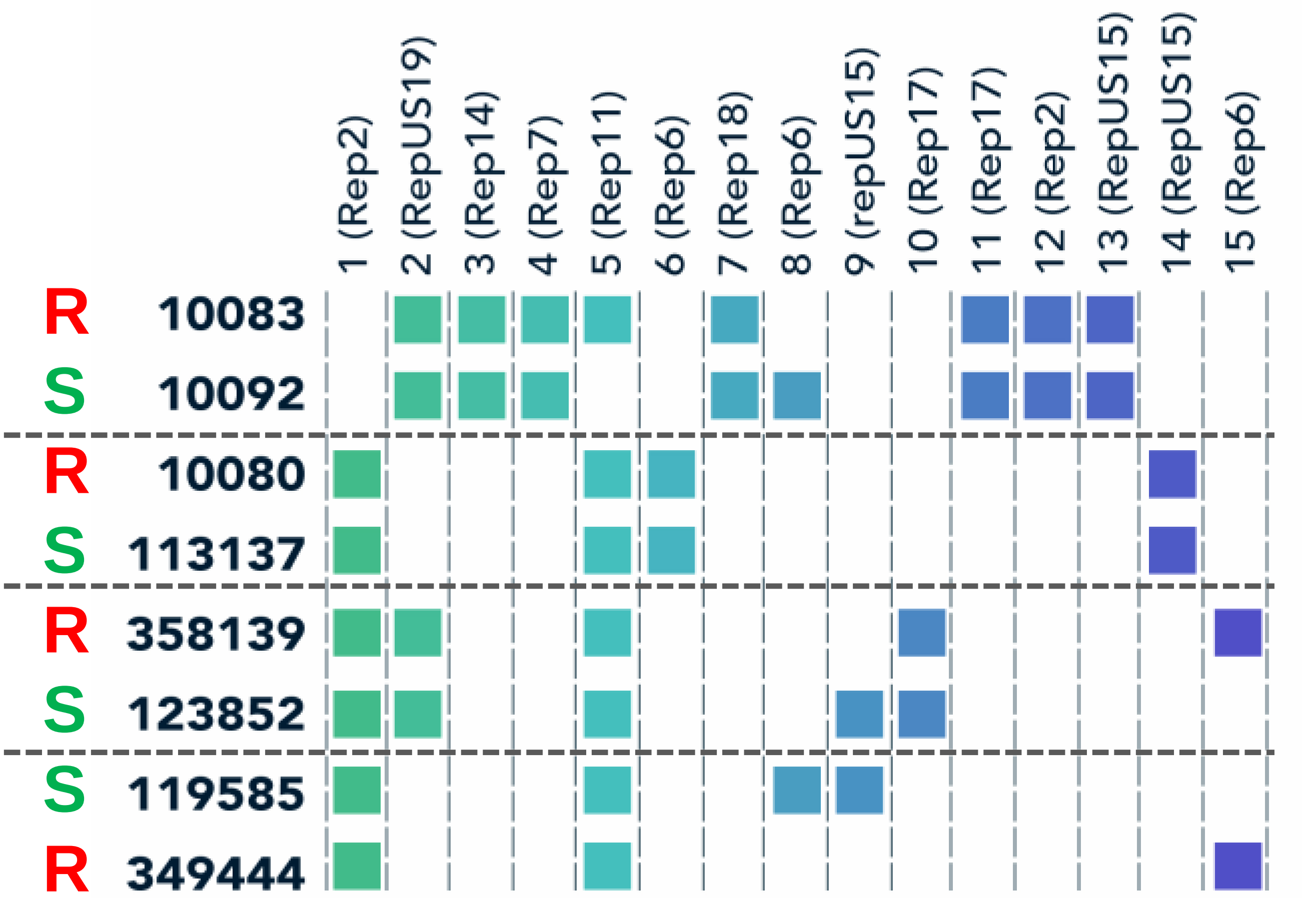
Gene presence/absence determined by Roary
Visualisation by Phandango

Steven Dunn



		1 (Rep2)	2 (RepUS19)	3 (Rep14)	4 (Rep7)	5 (Rep11)	6 (Rep6)	7 (Rep18)	8 (Rep6)	9 (repUS15)	10 (Rep17)	11 (Rep17)	12 (Rep2)	13 (RepUS15)	14 (RepUS15)	15 (Rep6)	
R	**10083**		■	■	■	■		■				■	■	■			ST780
S	**10092**		■	■	■			■	■			■	■	■			
R	**10080**	■				■	■								■		ST262
S	**113137**	■				■	■								■		
R	**358139**	■	■			■					■					■	
S	**123852**	■	■			■				■	■						
S	**119585**	■				■			■	■							
R	**349444**	■				■										■	

Steven Dunn



		1 (Rep2)	2 (RepUS19)	3 (Rep14)	4 (Rep7)	5 (Rep11)	6 (Rep6)	7 (Rep18)	8 (Rep6)	9 (repUS15)	10 (Rep17)	11 (Rep17)	12 (Rep2)	13 (RepUS15)	14 (RepUS15)	15 (Rep6)
R	**10083**		■	■	■	■		■				■	■	■		
S	**10092**		■	■	■			■	■			■	■	■		
R	**10080**	■				■	■								■	
S	**113137**	■				■	■								■	
R	**358139**	■	■			■					■					■
S	**123852**	■	■			■				■	■					
S	**119585**	■				■			■	■						
R	**349444**	■				■										■

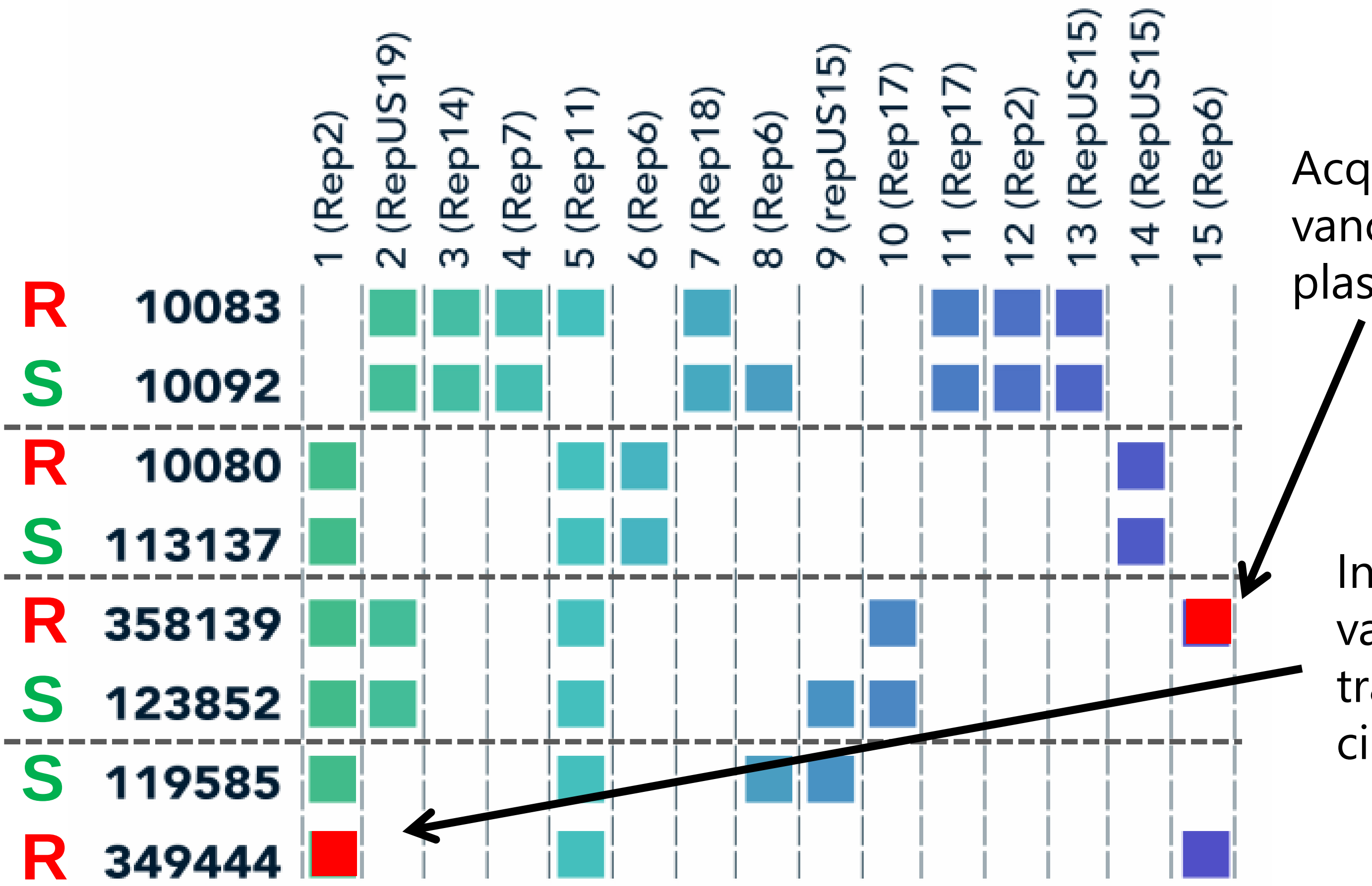
?

?

Vancomycin resistance transposon not resolved

Multiple insertion sites? Intra-strain transposition?

Steven Dunn



		1 (Rep2)	2 (RepUS19)	3 (Rep14)	4 (Rep7)	5 (Rep11)	6 (Rep6)	7 (Rep18)	8 (Rep6)	9 (repUS15)	10 (Rep17)	11 (Rep17)	12 (Rep2)	13 (RepUS15)	14 (RepUS15)	15 (Rep6)
R	10083		■	■	■	■		■				■	■	■		
S	10092		■	■	■			■	■			■	■	■		
R	10080	■				■	■								■	
S	113137	■				■	■								■	
R	358139	■	■			■					■					🟥
S	123852	■	■			■				■	■					
S	119585	■				■			■	■						
R	349444	🟥				■										■

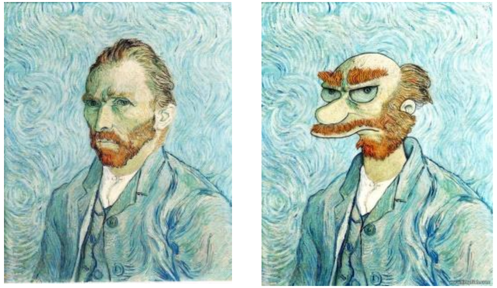
Acquisition of vancomycin resistance plasmid

Insertion of vancomycin resistance transposon into a circulating plasmid

Steven Dunn

Summary

Whole genome sequencing is required for high-resolution typing



UNIVERSITY OF
BIRMINGHAM

Ann Snaith
Maria Papangeli
Steven Dunn
Alan McNally

UMC Utrecht

Sergio Arredondo-Alonso
Rob Willems
Anita Schürch

NHS
University Hospitals
Birmingham
NHS Foundation Trust

Abid Hussain
Katie Hardy

Funding

NWO

THE
ROYAL
SOCIETY

Healthcare
Infection
Society

UNIVERSITY OF
BIRMINGHAM