New Diagnostics
Whole genome sequencing

Derrick Crook

Public Health England

University of Oxford

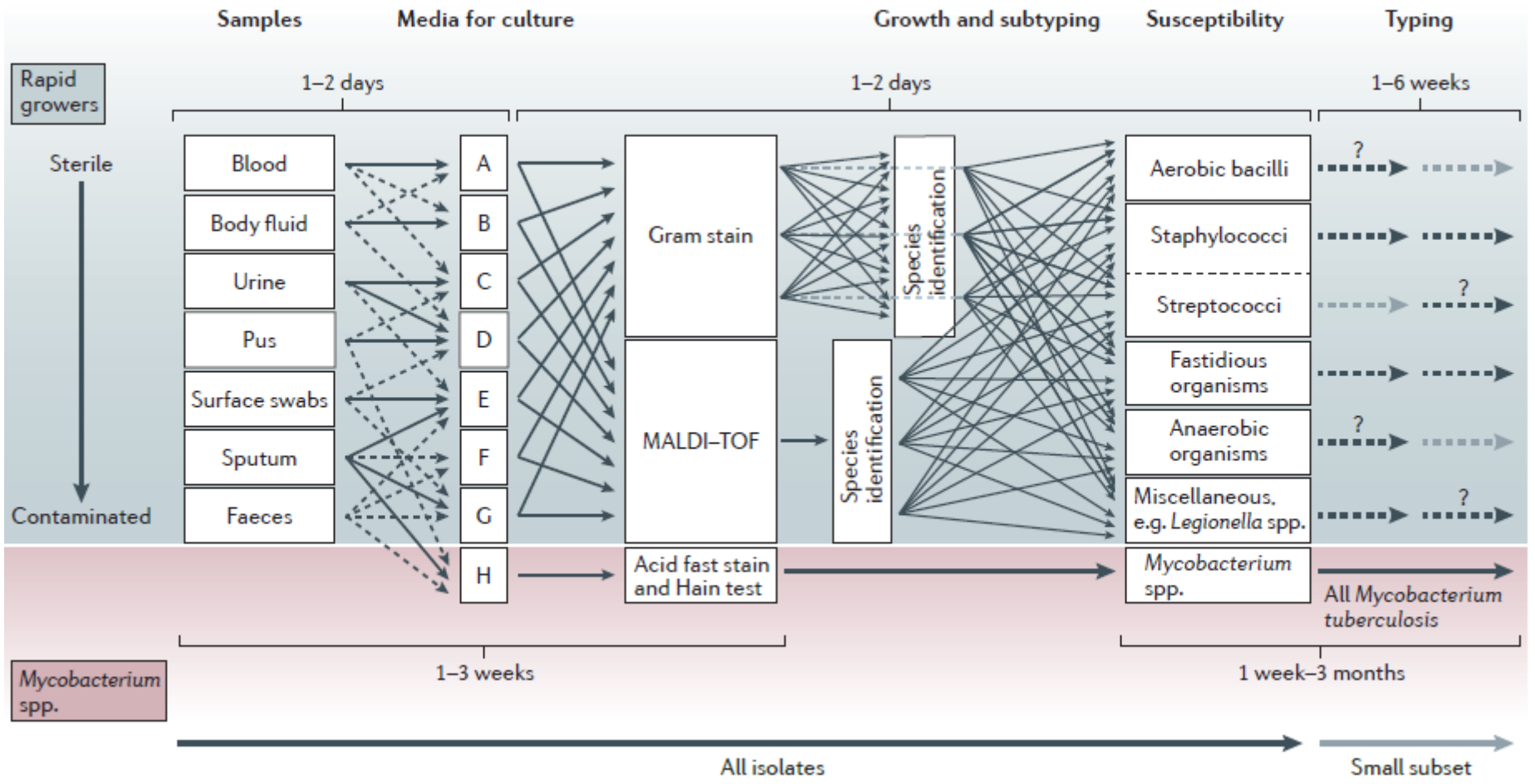
Oxford University Hospitals FT Trust

UNIVERSITY OF OXFORD

Public Health England

Background to WGS of pathogens

Current processes for bacteria



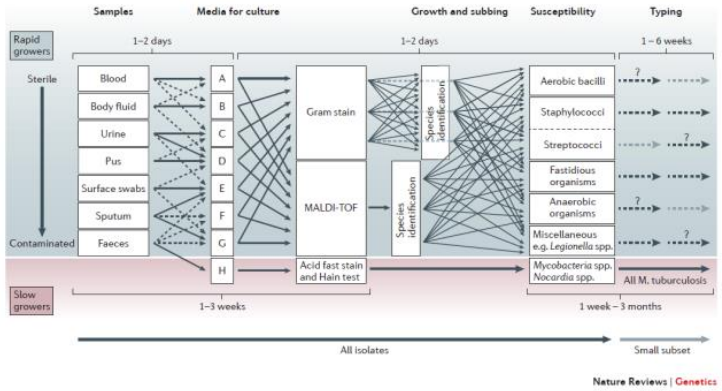
Nature Reviews Genetics 13, 601-612 (September 2012)

NATURE REVIEWS | GENETICS

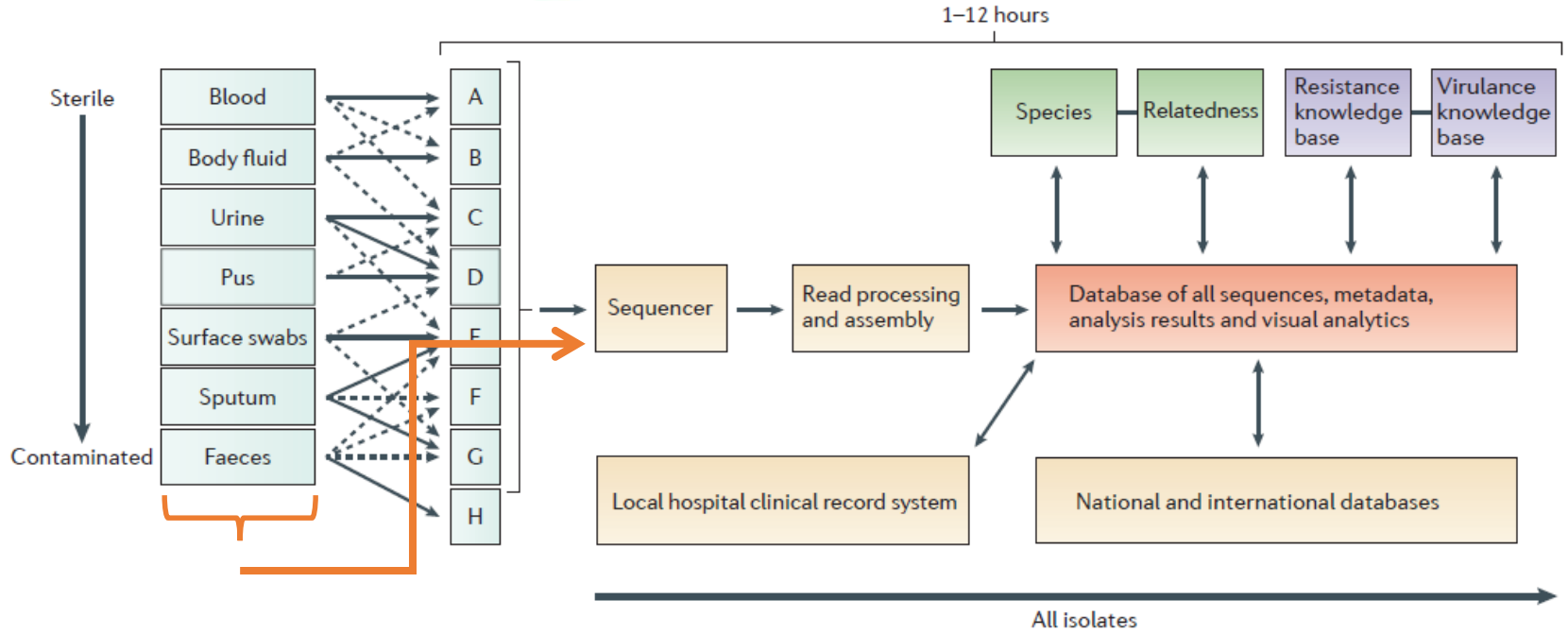
UNIVERSITY OF OXFORD

Public Health England

Future practice



In one step generate the complete diagnostic, typing and surveillance information

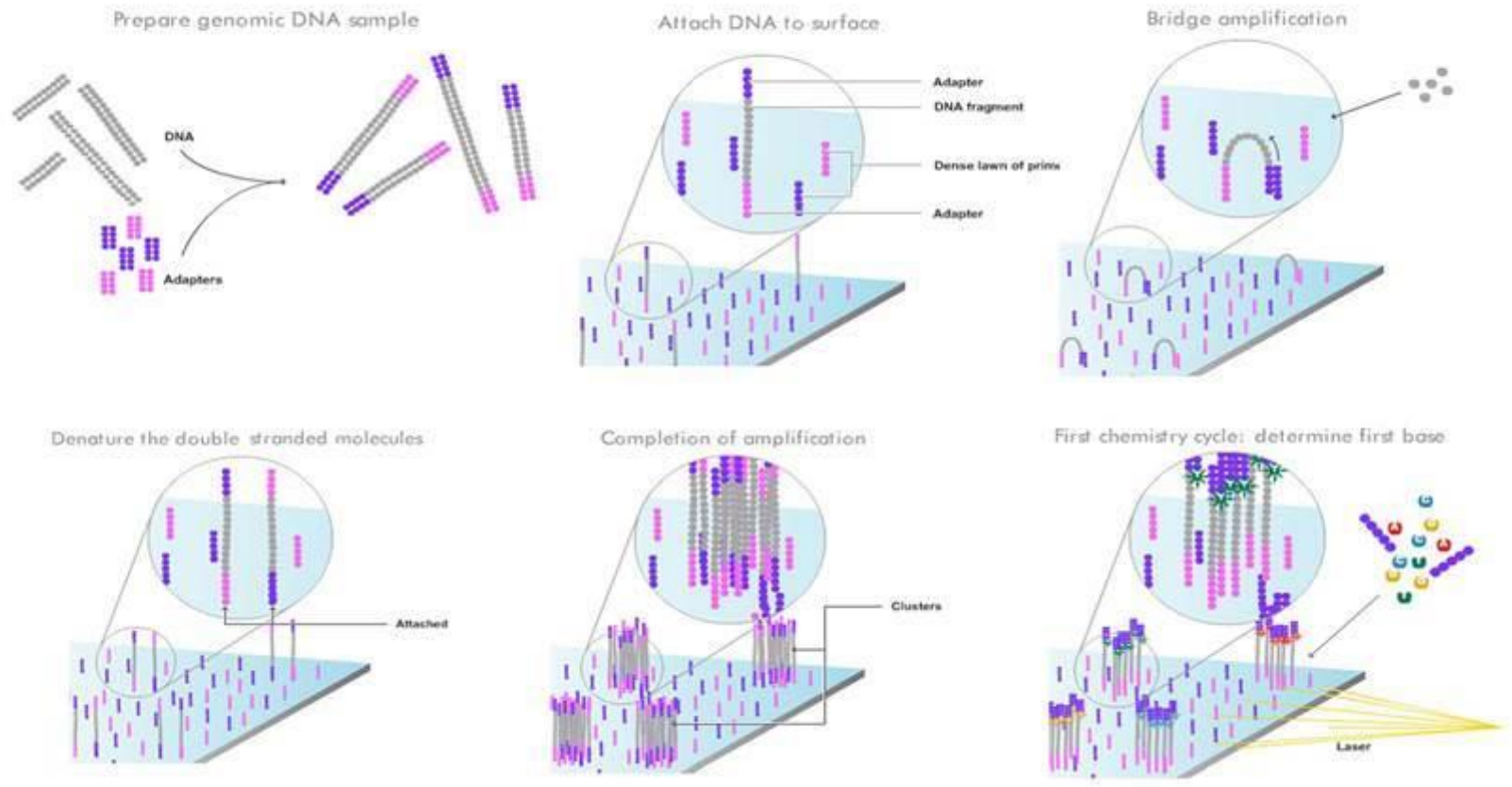


Nature Reviews Genetics 13, 601-612 (September 2012)

UNIVERSITY OF OXFORD

Public Health England

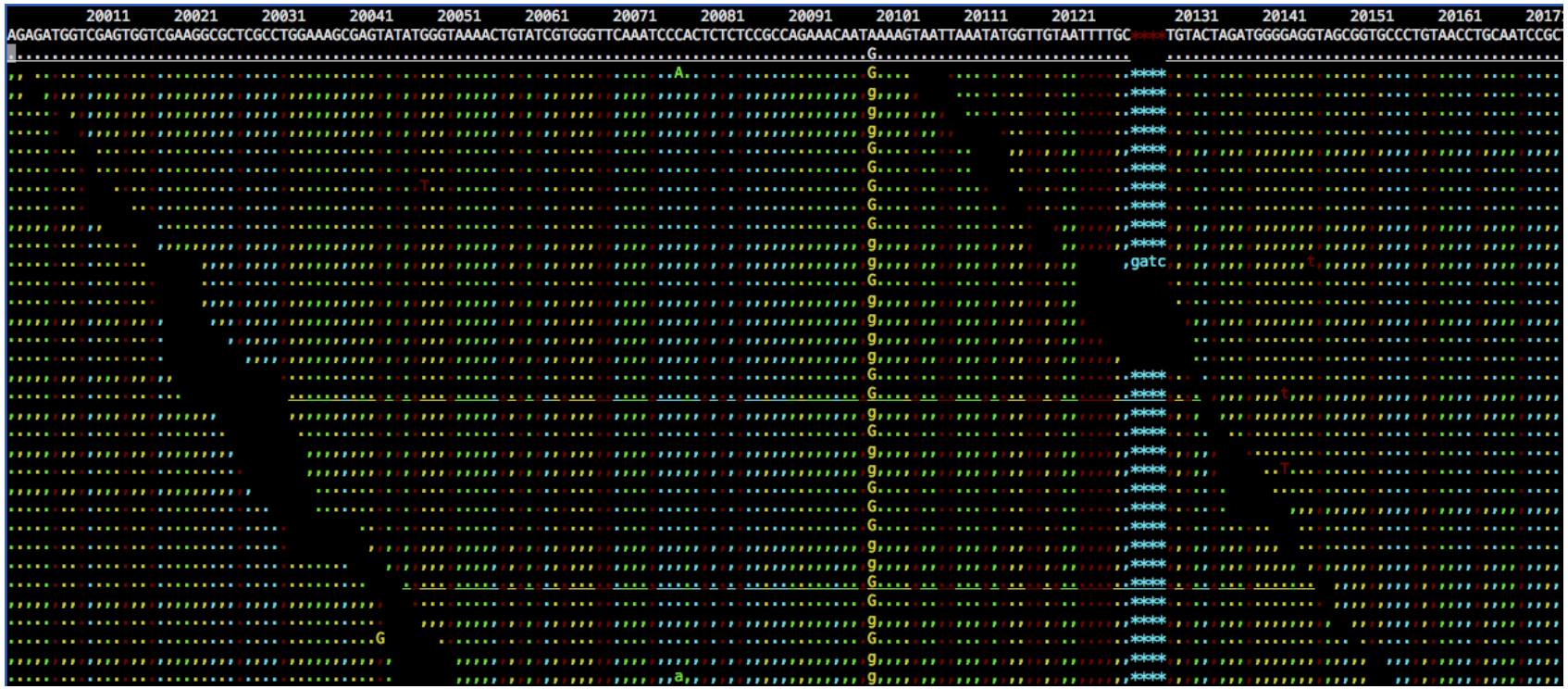
Parallel sequencing by synthesis with reversible fluorescent terminator nucleotides



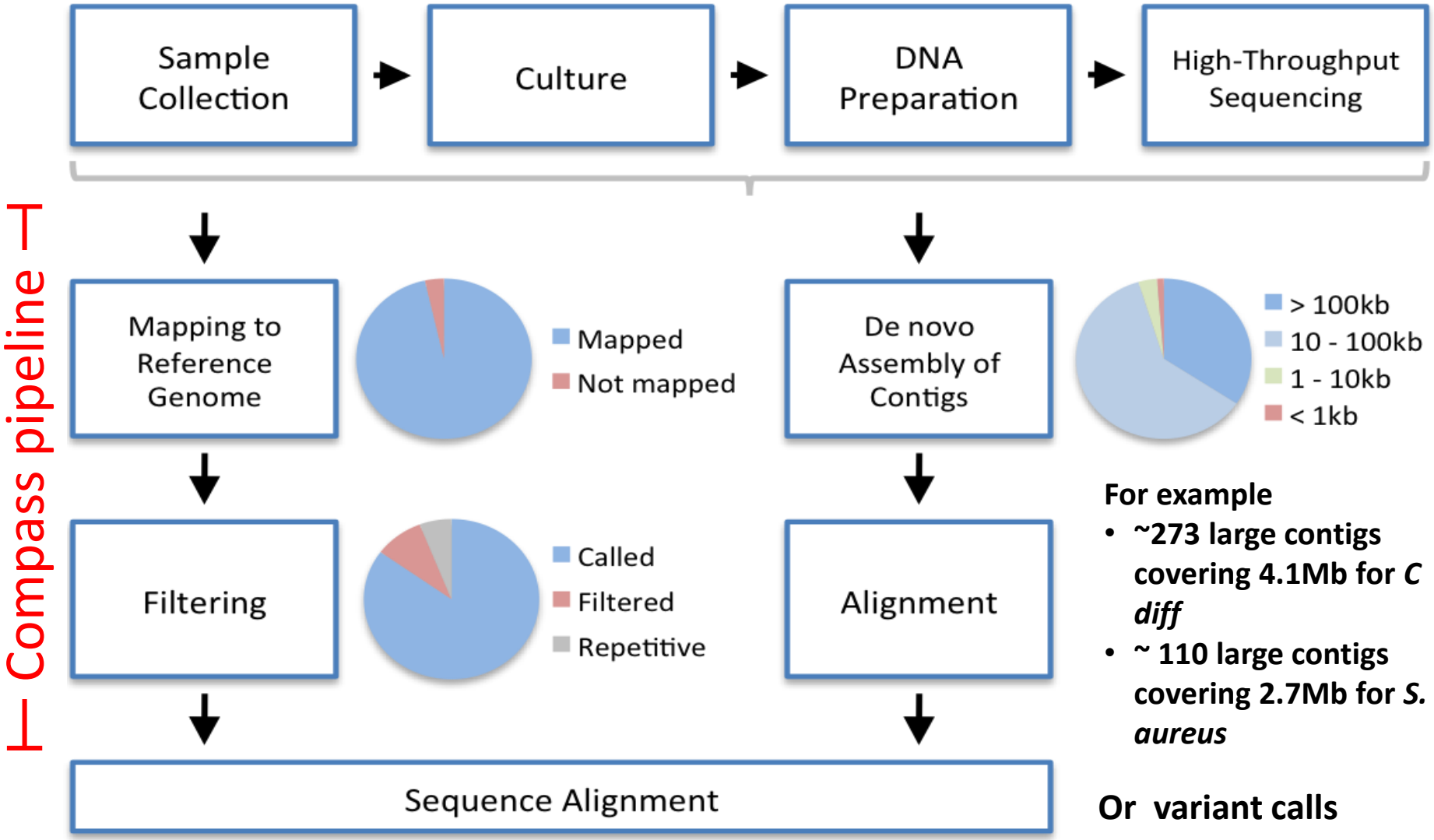
Google Images

Processing sequencing data

- Next generation sequencing produces millions of short fragments of DNA sequence (reads) from which genomes must be reconstructed



Whole genome sequencing – short reads

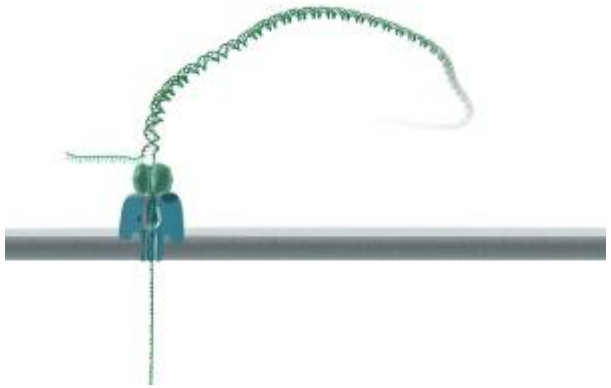


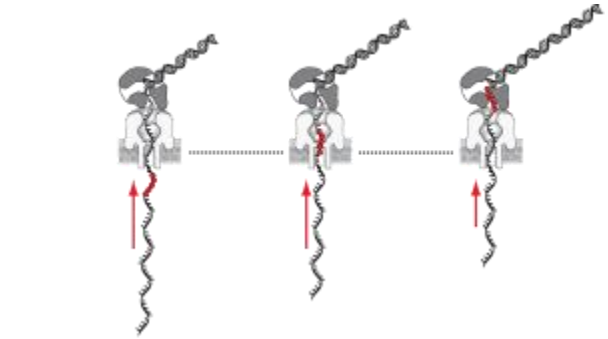
For example

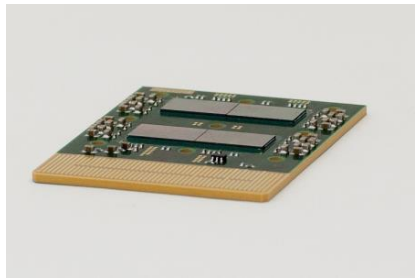
- **~273 large contigs covering 4.1Mb for *C diff***
- **~ 110 large contigs covering 2.7Mb for *S. aureus***

Or variant calls

Strand sequencer





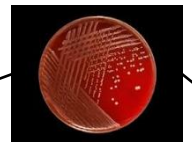


Oxford NANOPORE Technologies

UNIVERSITY OF OXFORD

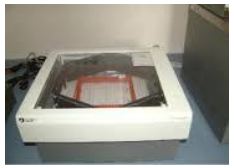
Public Health England

Vision for the future



Culture/sample

Now



Species Identification,
resistance prediction &
Typing:
strain-typing
patho-typing

Speed
Accuracy
Cost

Future

UNIVERSITY OF OXFORD

Public Health England

What is needed to get a solution in place?

- For outbreak detection
- Diagnostic utility
- Develop software to automate
  - Validate/test

Transmission: we need to be able to interpret variation

A group of patient isolates with a number of variants (single nucleotide polymorphisms (SNPs) or variants (SNVs), indels, recombinations) between them

- some pairs with 0 variants
- some pairs with 1, 2, 3,.... "some" variants
- some pairs with "many" variants

Decision: what does having "some" variation between pairs of isolates mean in terms of transmission?

Transmission: Five pillars of wisdom
Outbreak detection

Interpret observed differences between isolates in the context of diversity in

Five pillars of wisdom for assessing probability of transmission: specific to each species

- Key aspects of diversity for interpreting data:
  - diversity in re-sequencing and re-calling same sample (~0)
  - diversity within an individual at one point in time (“cloud”)
  - diversity within an individual over time (evolutionary clock)
  - diversity within a point source outbreak (eg TB household)
  - diversity in the community – background diversity

- Estimate prediction intervals for expected variants to be observed between highly related isolates (i.e. transmissions) based on clinically relevant timescales
  - thresholds are organism and assembly pipeline dependent
  - derived from the evolutionary clock (this can vary)

What is needed for having confidence in recovering diagnostic information?

~~Five~~ Seven pillars of wisdom

UNIVERSITY OF OXFORD

Public Health England

How are we doing with measuring resistance

- We need the following:
  - software for processing the DNA sequence
  - Knowledgebase of variation conferring resistance to interrogate
  - Continuous updating of these knowledgebases; how do we find new variation

- How are we doing:
  - *Staphylococcus aureus*
  - *Escherichia coli*
  - *Mycobacterium tuberculosis*

Resistance prediction from WGS

Iterative method of development

- A derivation set: compare genotypic prediction vs a gold-standard phenotypic susceptibility test
- Refine the catalogue and software
- A replication set: re-evaluate resistance prediction vs phenotype recording **very major** and **major errors**
- Analyse discrepant and improve the software, knowledge base and (if necessary) phenotypic methodology
- Test the revised algorithm with a fresh set of samples

S. aureus: Resistance prediction algorithm

- Derivation set of 501 samples
- Algorithm was refined after the derivation set.
- Many of the discrepant results were found to be phentypic errors in the routine laboratory.
- Other discrepants were resolved by improvements in the bio-informatics software
- The improved algorithm was tested against a further 487 isolates (the ‘validation’ set).

Gordon et al J Clin Microbiol. 2014 Feb 5

Blinded validation study of resistance prediction from WGS *Staphylococcus aureus* (478)

	Phenotype: resistant		Phenotype: susceptible		Error Rates	
	Genotype		Genotype		ME	VME
Antimicrobial	Susceptible	Resistant	Susceptible	Resistant	(%)	(%)
Penicillin	2	398	84	3	3.4	0.5
Methicillin	0	55	432	0	0.0	0.0
Ciprofloxacin	2	64	421	0	0.0	3.0
Erythromycin	1	80	404	2	0.5	1.2
Clindamycin	1	76	2	0	0.0	1.3
Tetracycline	0	18	467	2	0.4	0.0
Vancomycin	0	0	491	0	0.0	n/a
Fusidic acid	1	39	445	0	0.0	2.6
Trimethoprim	0	2	200	1	0.5	0.0
Gentamicin	1	2	484	0	0.0	33.3
Mupirocin	0	2	485	0	0.0	0.0
Rifampicin	0	5	482	0	0.0	0.0
Total	8	741	4397	8	0.2	1.1

Gordon et al J Clin Microbiol. 2014 Feb 5

UNIVERSITY OF OXFORD

Public Health England

Previous phenotyping studies

Study	Comparison	no of isolates	Categorical agreement (%)	ME rate (%)	VME rate (%)
Ligozzi 2002	Vitek 2 vs agar dilution	100	94-100	0	0
Fahr 2003	BD Phoenix vs broth dilution plus mecA PCR	116	97.6	1.2	1.7
Nonhoff 2005	Vitek 2 vs agar dilution	273	-	1.5	0.7
Carroll 2006	BD Phoenix vs agar dilution	232	98.2	0.3	0.4
Giani 2012	BD Phoenix vs broth dilution	95	98	1.3	2.1
Bobenchik 2014	Vitek 2 vs broth dilution	134	98.9	0.1	1.4
This study	WGS vs combined disc diffusion / BD Phoenix	491	98.8	0.2	1.1

Escherichia coli

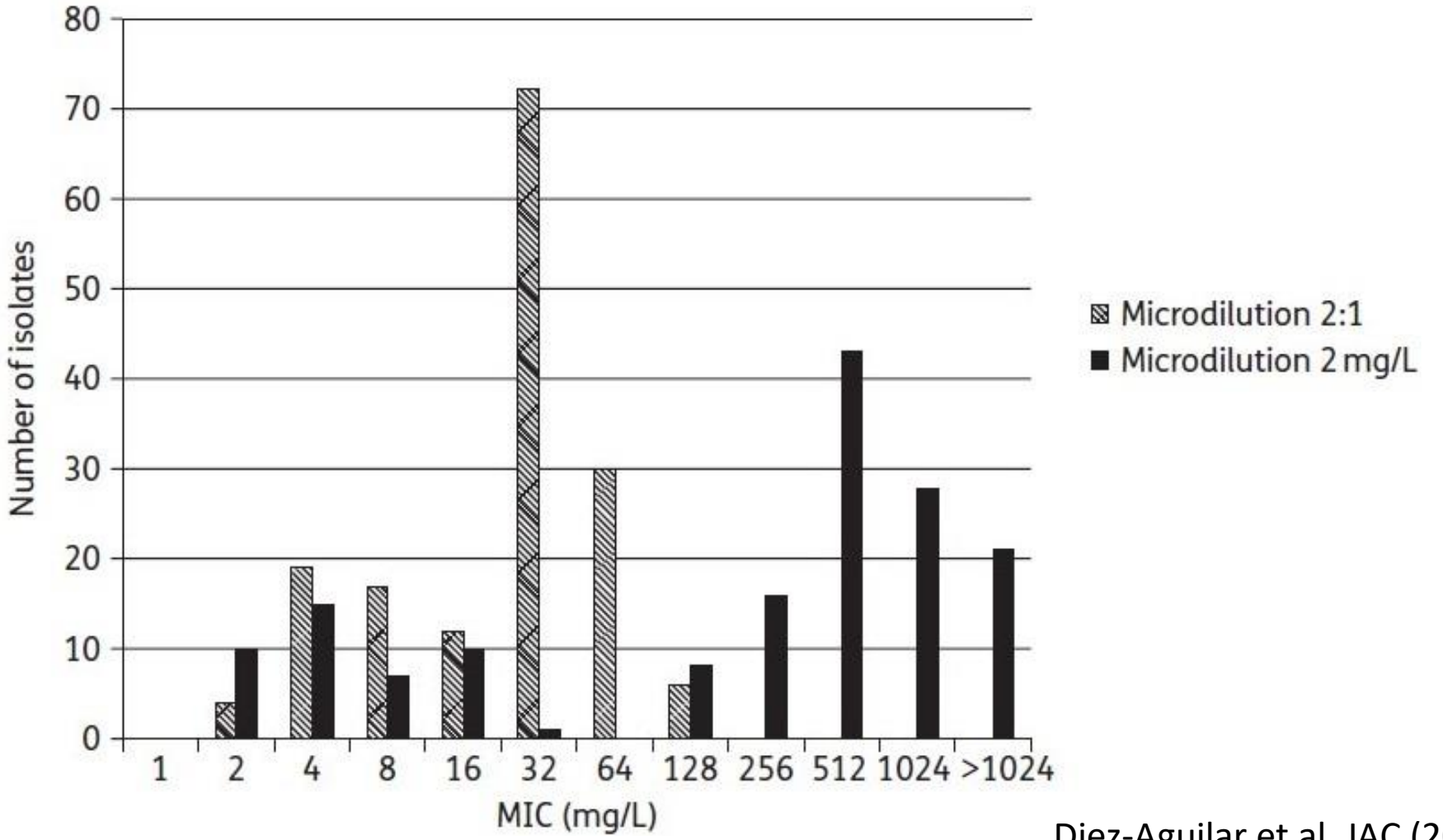
Sensitivity and specificity of genotypic resistance predictions versus gold standard "reference" phenotype results for 74 *Escherichia coli* bloodstream isolates

Table 3. Sensitivity and specificity of genotypic resistance predictions versus comparison with standard phenotype results for 74 *E. coli* bloodstream isolates.

	Susceptible by comparison standard phenotype		Resistant by comparison standard phenotype			
Antibiotic	susceptible by genotype (row %)	resistant by genotype (row %; major error)	susceptible by genotype (row %; very major error)	resistant by genotype (row %)	Sensitivity (95% CI)	Specificity (95% CI)
Amoxicillin	23 (31)	1 (1)	0 (0)	50 (68)	1.00 (0.91–1.00)	0.96 (0.77–1.00)
Co-amoxiclav	46 (62)	0 (0)	0 (0)	28 (38)	1.00 (0.85–1.00)	1.00 (0.90–1.00)
Gentamicin	60 (81)	0 (0)	0 (0)	14 (19)	1.00 (0.73–1.00)	1.00 (0.93–1.00)
Ciprofloxacin	48 (65)	0 (0)	0 (0)	26 (35)	1.00 (0.84–1.00)	1.00 (0.91–1.00)
Ceftriaxone	43 (58)	1 (1)	1 (1)	29 (39)	0.97 (0.81–1.00)	0.98 (0.87–1.00)
Ceftazidime	43 (58)	11 (15)	1 (1)	19 (26)	0.95 (0.73–1.00)	0.80 (0.66–0.89)
Meropenem	74 (100)	0 (0)	0 (0)	0 (0)	—	1.00 (0.94–1.00)
Total	337 (65)	13 (3)	2 (0.3)	166 (32)	0.99 (0.95–1.00)	0.96 (0.94–0.98)

J. Antimicrob. Chemother. (2013)

Disparity in Coamoxiclav phenotype



Diez-Aguilar et al. JAC (2015)

Discordance in amoxicillin-clavulanate phenotyping methods

- random stratified sample of 250 *E. coli* isolates causing bacteraemia in Oxfordshire in 2013-14 for AMC resistance by repeated phenotyping and whole genome sequencing (WGS) to determine the resistance mechanisms

- MICs were determined by agar dilution, in triplicate for AMC (taking the maximum for analysis) and amoxicillin; and in duplicate for clavulanate

Results: CLSI vs EUCAST – Amoxicillin-clavulanate susceptibility testing

- Non-susceptible 156 (62%) – EUCAST , 112 (45%) – CLSI
- 61% of isolates had higher EUCAST MIC than CLSI MIC
- MIC mean difference across replicates, EUCAST 0.8 doubling dilutions less than CLSI
- Standard deviation of EUCAST MICs greater than CLSI MICs (2.0 vs 1.1)
- 56% of isolates had at least 1 replicate with an MIC of ≤ 16mg/L for clavulanate alone

Significant discordance in amoxicillin-clavulanate genotypic attribution of phenotypic “resistance”

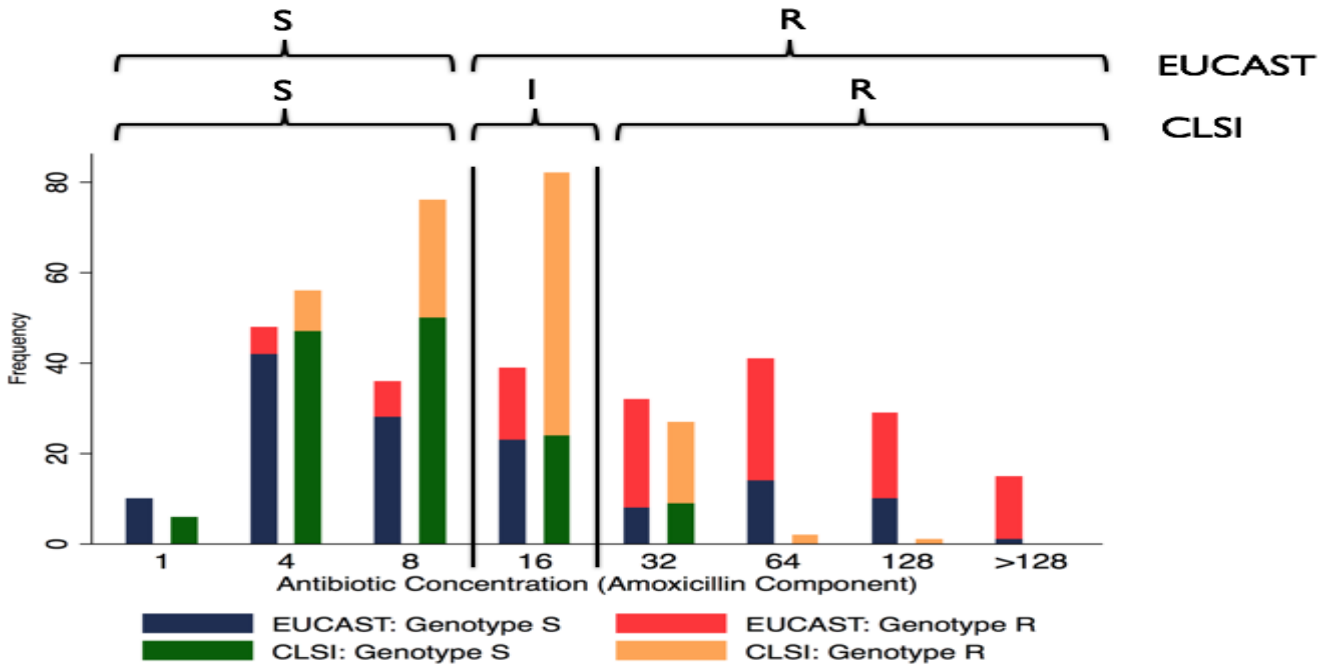


Fig 1: Overall (maximum of replicates) MICs of samples, by method and presence/absence of genetic mechanism

Method	All repeats sensitive		Both phenotypes observed		All repeats resistant	
	WGS-S	WGS-R	WGS-S	WGS-R	WGS-S	WGS-R
CLSI	103 (75%)	35 (25%)	6 (25%)	18 (75%)	27 (31%)	61 (69%)
EUCAST	80 (85%)	14 (15%)	23 (62%)	14 (38%)	33 (28%)	86 (72%)
Overall*	80 (85%)	14 (15%)	29 (43%)	39 (57%)	27 (31%)	61 (69%)

: Phenotypes seen in all repeats by both methods. Note showing N(row%)

27 resistant by both methods had no identifiable mechanism for resistance

UNIVERSITY OF OXFORD

Public Health England

Mycobacterium tuberculosis

Anti-tuberculosis drug resistance prediction

- Arguably 15 drugs are available for treating TB with more new drugs in development
- Is genomic variation which confers resistance limited to somewhere between 20 to 30 genes?
- Current knowledge indicates molecular prediction of INH, rifampicin resistant or pan-susceptible isolates is ~ 95% accurate
- The knowledge base of variation conferring resistance to ‘all drugs’ is incomplete

Can we discover explanatory variation in TB?


Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance: a retrospective cohort study

Timothy M Walker*, Thomas A Kohl*, Shaheed V Omar*, Jessica Hedge*, Carlos Del Ojo Elias, Phelim Bradley, Zamin Iqbal, Silke Feuerriegel, Katherine E Niehaus, Daniel J Wilson, David A Clifton, Georgia Kapatai, Camilla L C Ip, Rory Bowden, Francis A Drobniewski, Caroline Allix-Béguec, Cyril Gaudin, Julian Parkhill, Roland Diel, Philip Supply, Derrick W Crook, E Grace Smith, A Sarah Walker, Nazir Ismail†, Stefan Niemann†, Tim E A Peto†, and the Modernizing Medical Microbiology (MMM) Informatics Group

- Investigation of 3651 isolates :
  - Using a heuristic method of predicting resistance
- divided into
  - a 2099 derivation set
  - a 1552 validation set
- Resistance is conferred by genomic variation:
  - Non-synonymous mutations , deletions and insertions in relevant genes – 23 genes
  - Arises mostly de-novo in a non-recombining genome leading to homoplasy

Lancet Infect Dis 2015; 15: 1193–1202

TB drug resistance prediction in a validation set

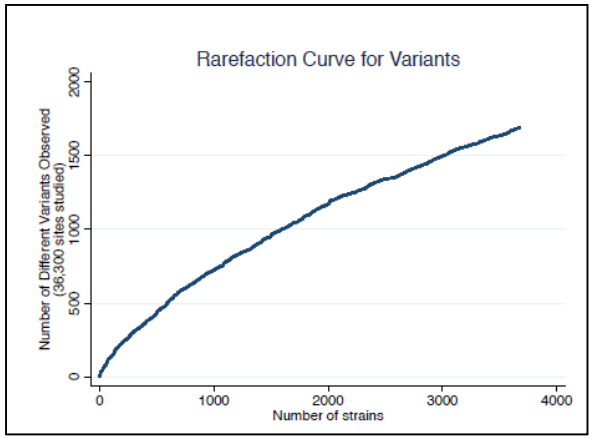
	Phenotypically Resistant Genotype					Phenotypically Sensitive Genotype					All		Excluding Unclassfied		
	R	S_0	S_s	U	Total	R	S_0	S_s	U	Total	Sensitivity	Specificity	Sensitivity	Specificity	% Unclassifed
Isoniazid	310	18	1	35	364	19	1,065	52	52	1188	85.2	98.4	94.2	98.3	5.6
Rifampicin	275	8	1	16	300	10	1,200	4	38	1252	91.7	99.2	96.8	99.2	3.5
Ethambutol	158	7	1	26	192	67	1003	79	210	1359	82.3	95.1	95.2	94.2	15.2
Pyrazinamide	43	27	5	104	179	2	1,218	67	83	1370	24.0	99.9	57.3	99.8	12.1
Streptomycin	284	6	9	49	348	11	970	34	189	1204	81.6	99.1	95.0	98.9	15.3
Ofloxacin	5	4	2	0	11	0	489	134	38	661	45.5	100.0	45.5	100.0	5.7
Amikacin	52	5	0	2	59	3	427	38	140	608	88.1	99.5	91.2	99.4	21.3
Total	1127	75	19	232	1453	112	6372	408	750	7642	77.6	98.5	92.3	98.4	10.8

Table 1: Genotypic predictions in the validation-set based on: R (resistance-determinant); S0 (zero non-synonymous variants/SNPs present); Ss (only sensitive variants present); U (unclassified variants present). Weighted mean sensitivity and specificity given for all phenotypes, and with the 10.8% of phenotypes associated with previously unclassified variation (U) excluded.

Lancet Infect Dis 2015; 15: 1193–1202

Filling the resistance gap

Comprehensive Resistance Prediction for Tuberculosis: an International Consortium (CRyPTIC) and ResSeq-TB



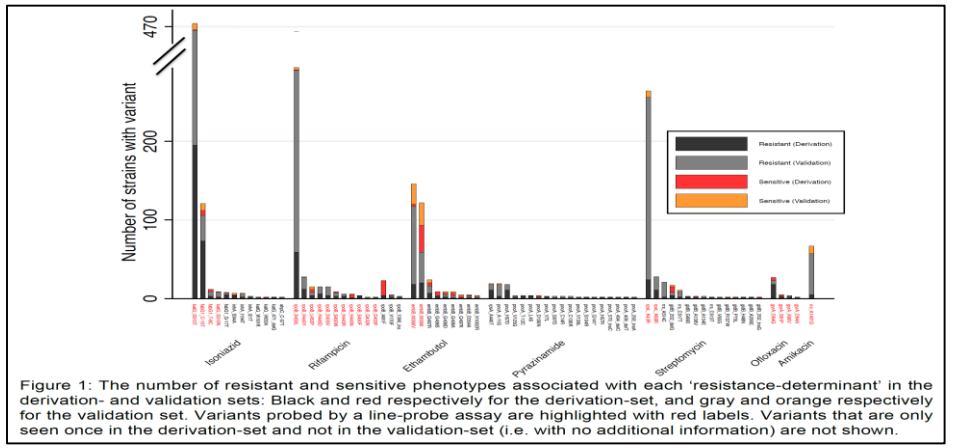
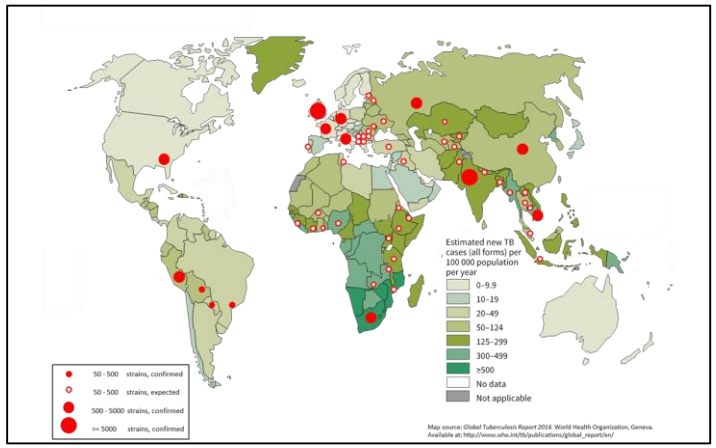


Figure 1: The number of resistant and sensitive phenotypes associated with each 'resistance-determinant' in the derivation- and validation sets: Black and red respectively for the derivation-set, and gray and orange respectively for the validation set. Variants probed by a line-probe assay are highlighted with red labels. Variants that are only seen once in the derivation-set and not in the validation-set (i.e. with no additional information) are not shown.



Phenotyping

Genotypic characterisation

- 100,000 WGS TB pledged
- ~ 40,000 with extensive DST
- Analysis:
  - Heuristic approach
  - GWAS
  - Machine Learning
  - Thermodynamic modelling of proteins
  - Molecular genetic characterisation

UNIVERSITY OF OXFORD

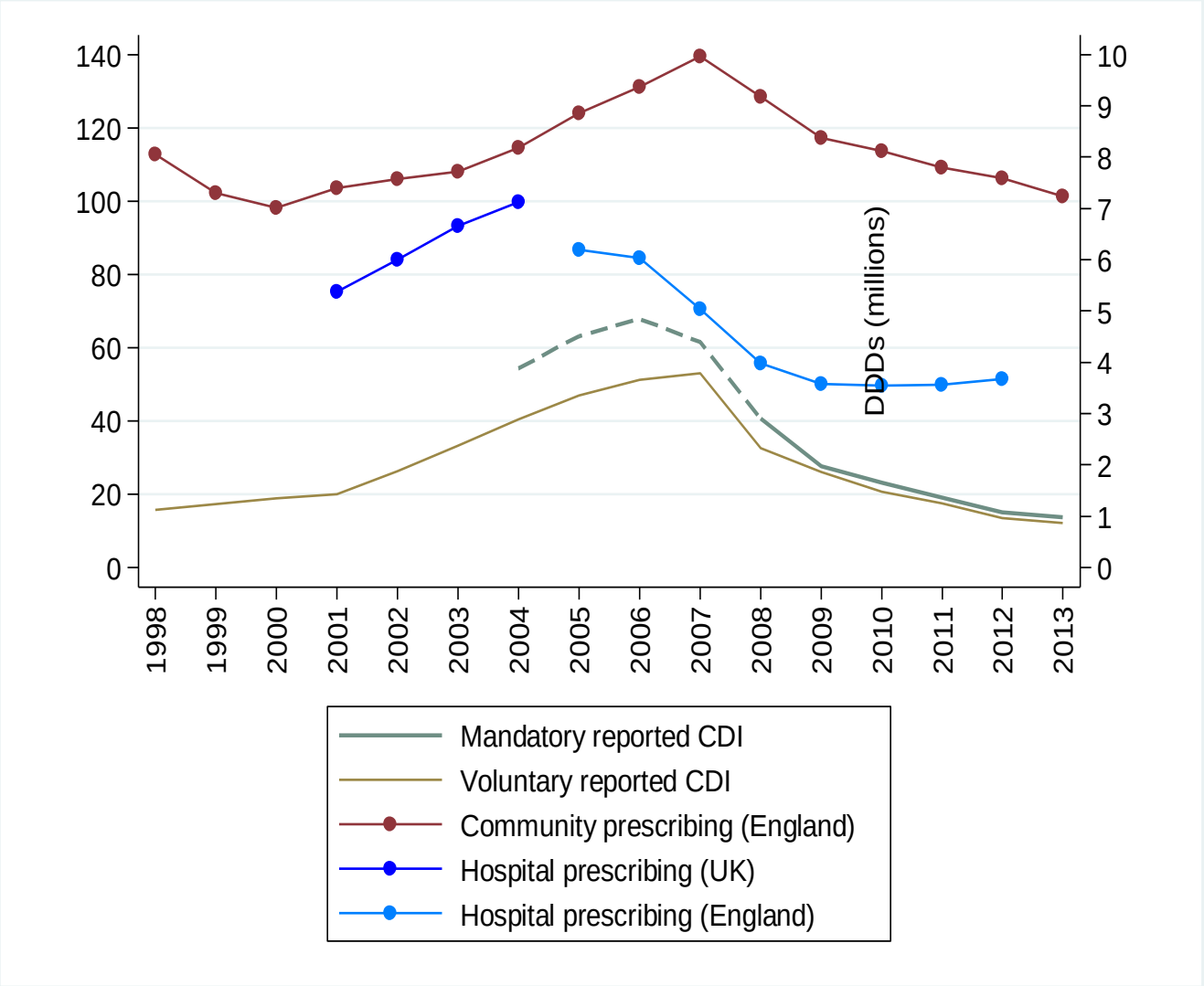
BILL & MELINDA GATES foundation

wellcome trust

Public Health England

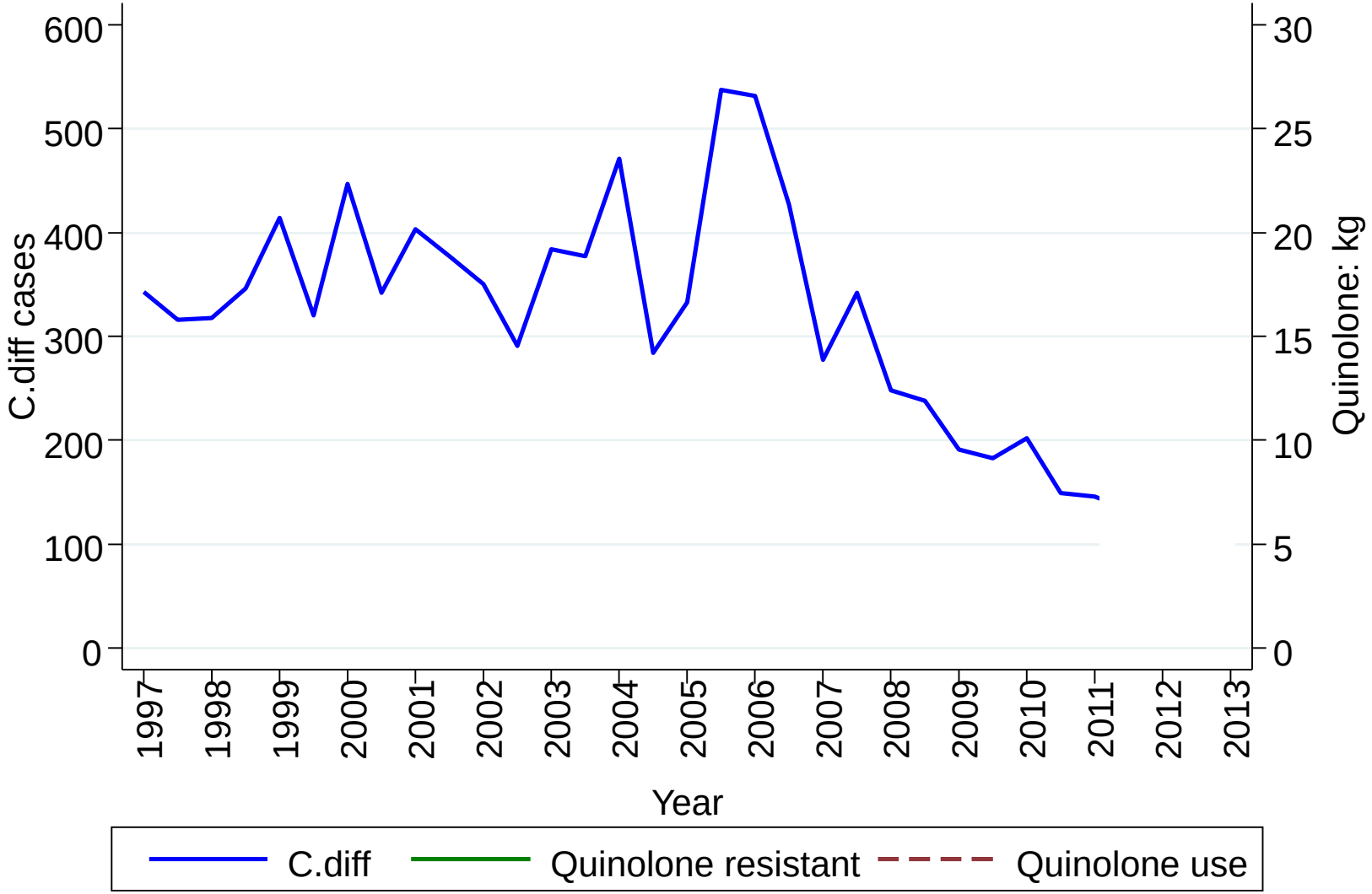
Selection, dispersal and control of C. difficile

Change in incidence and quinolone usage nationally



Under review

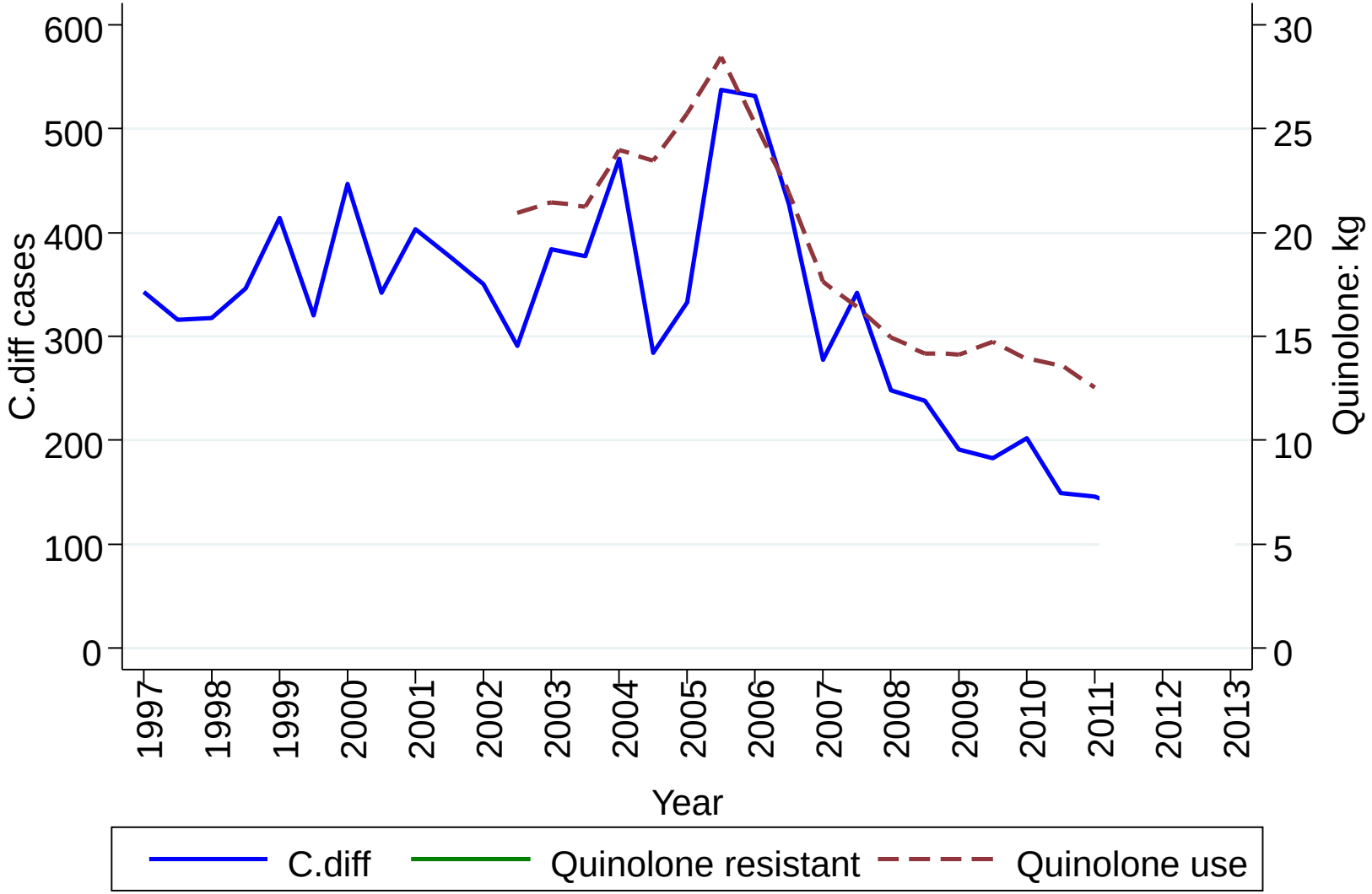
Oxfordshire *C. difficile* cases



UNIVERSITY OF OXFORD

Public Health England

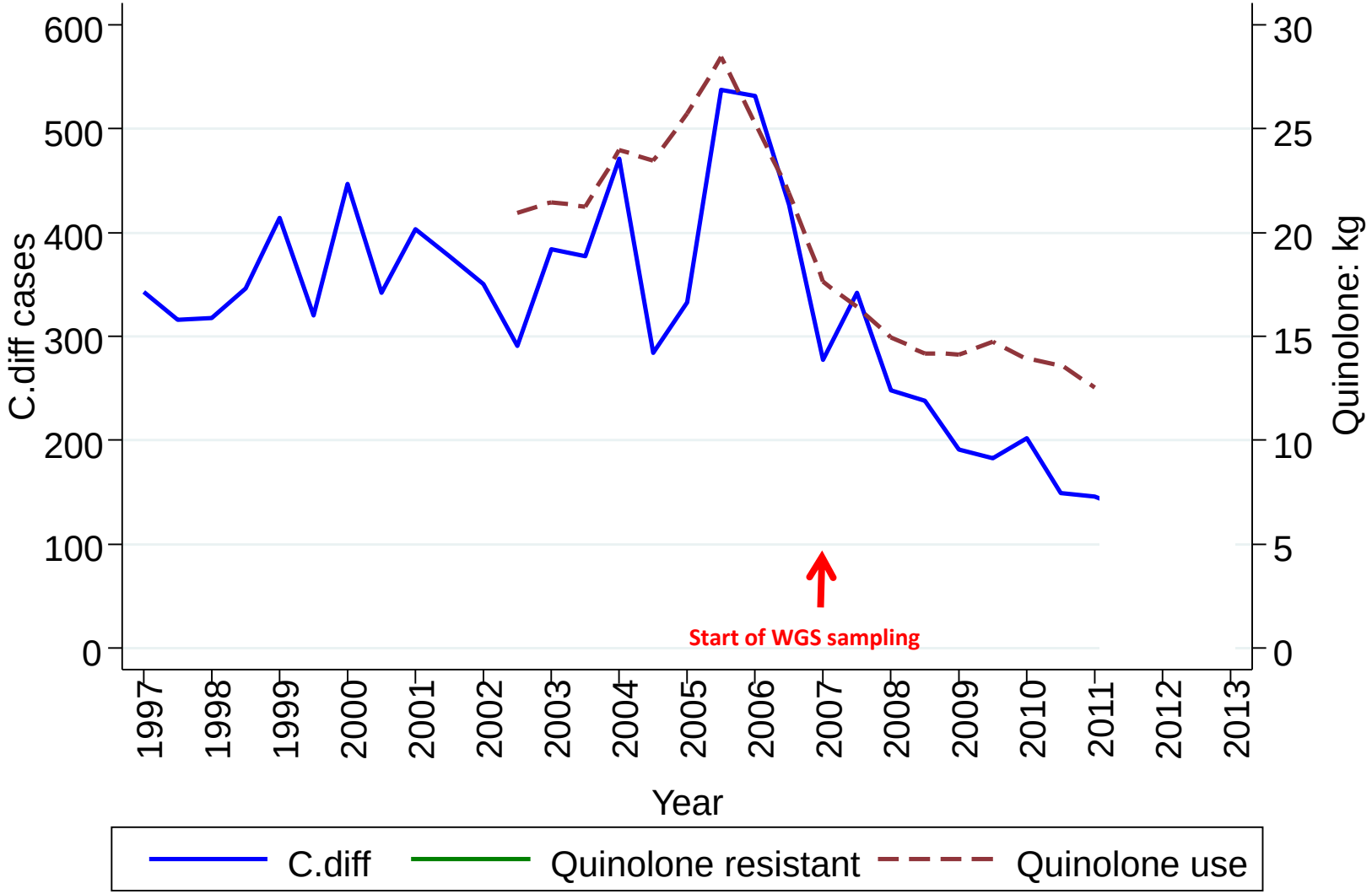
Oxfordshire *C. difficile* cases



UNIVERSITY OF OXFORD

Public Health England

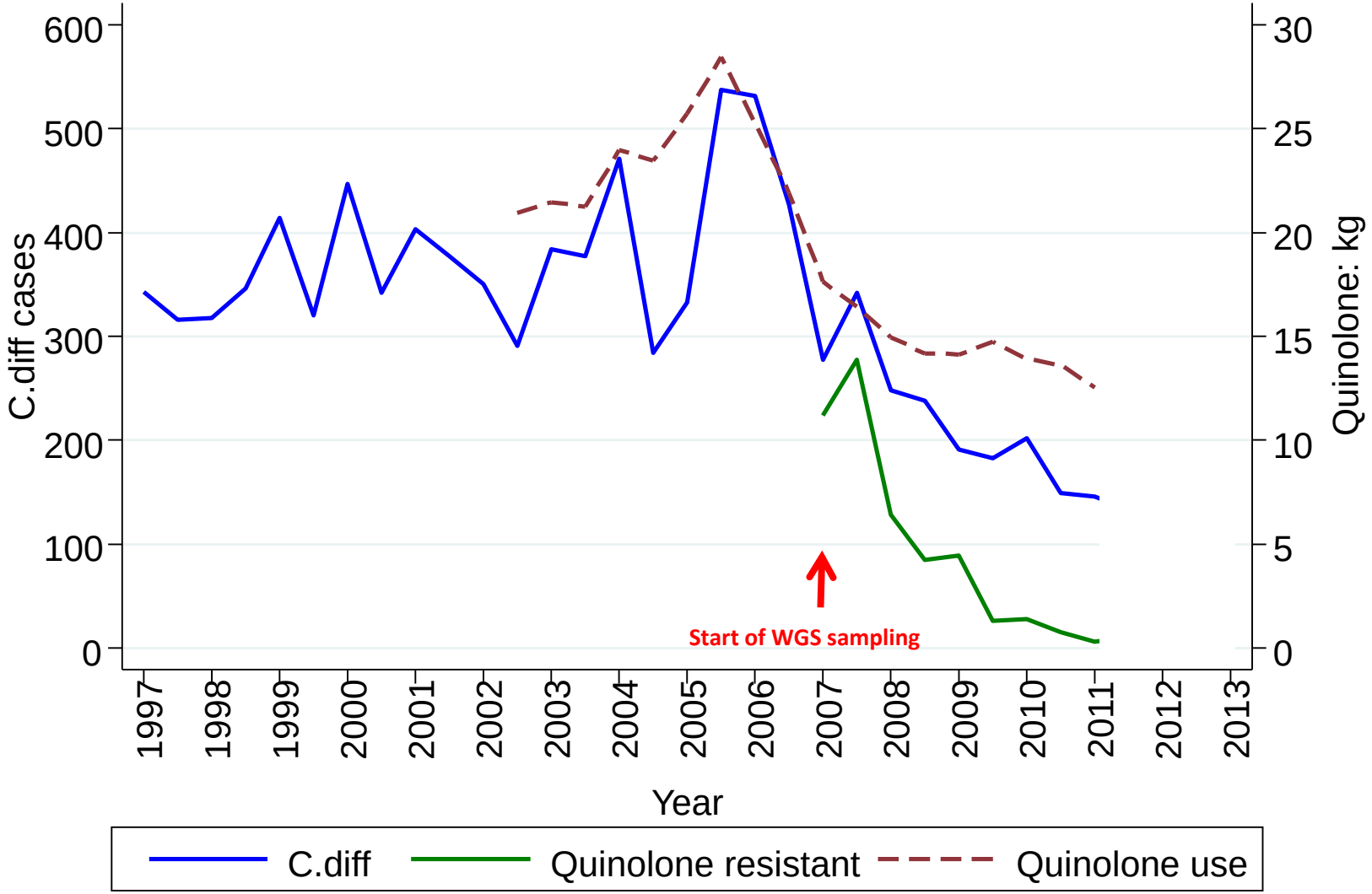
Oxfordshire *C. difficile* cases



UNIVERSITY OF OXFORD

Public Health England

Oxfordshire *C. difficile* cases

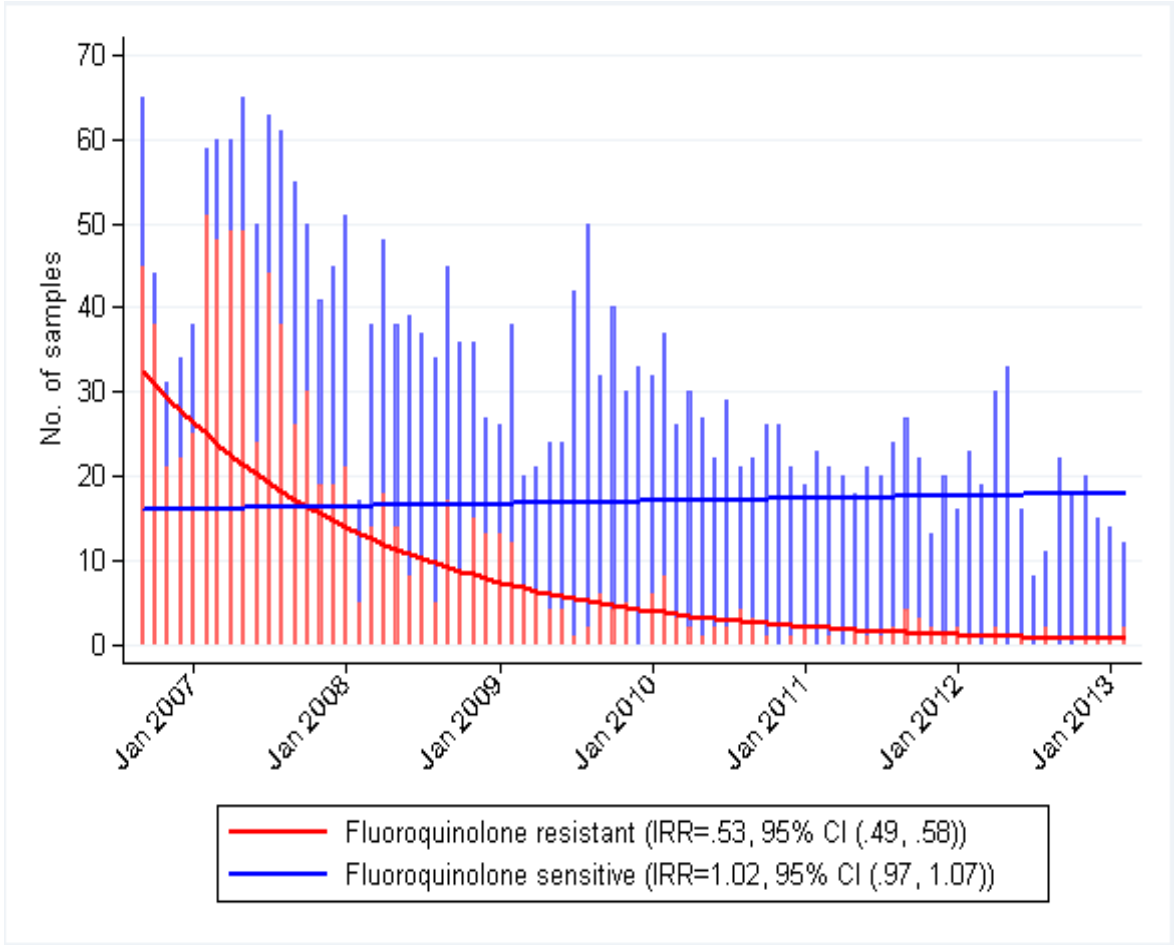


UNIVERSITY OF OXFORD

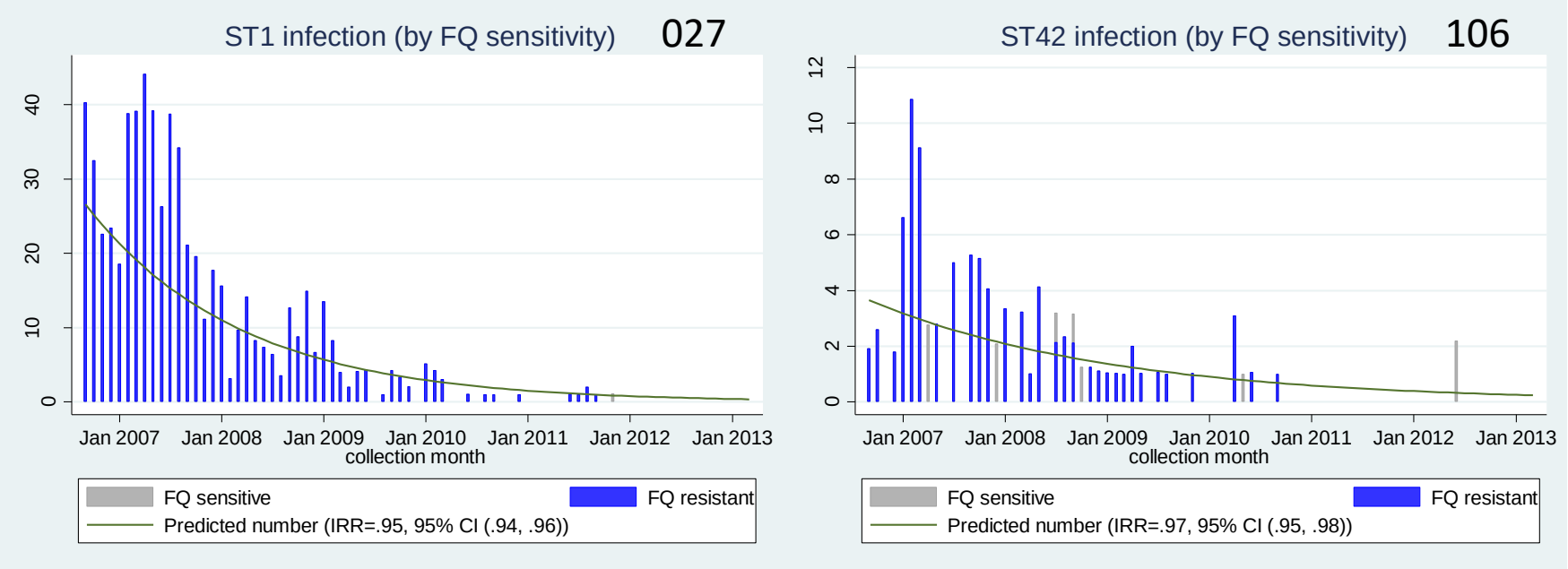
Public Health England

Fluoroquinolone resistant

Declining CDI in Oxford

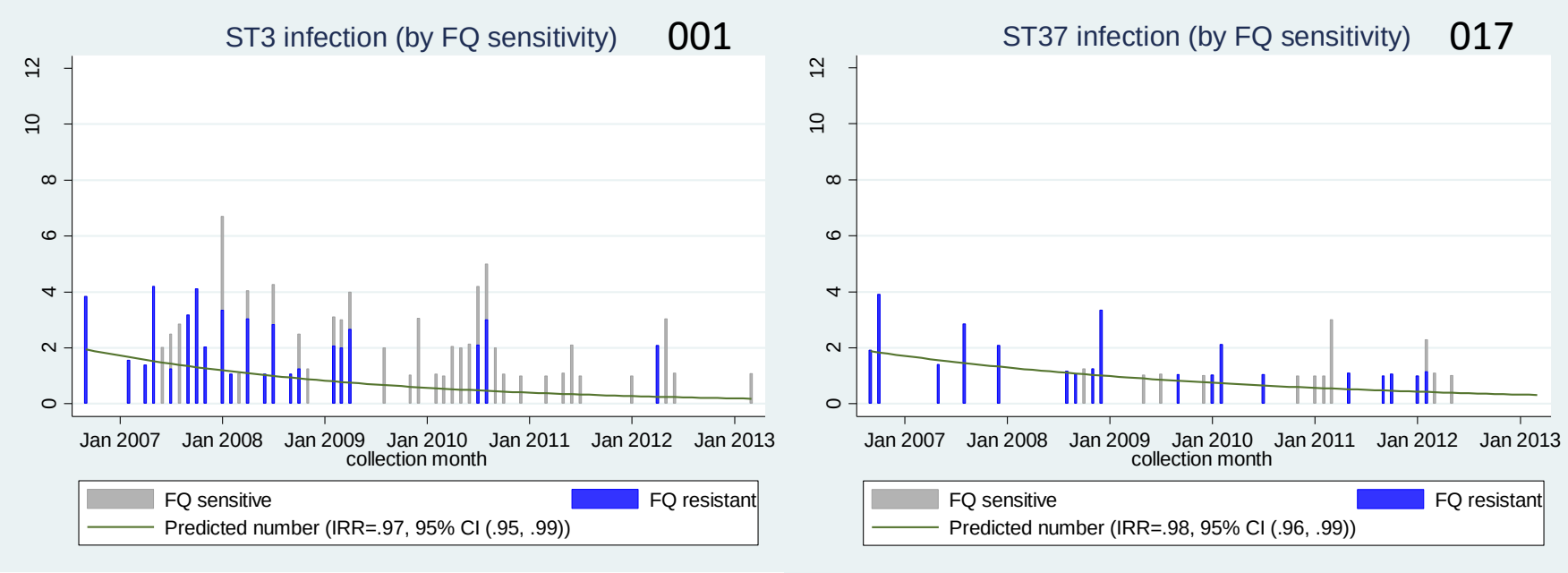


Incidence of FQ resistant genotypes has declined (1)



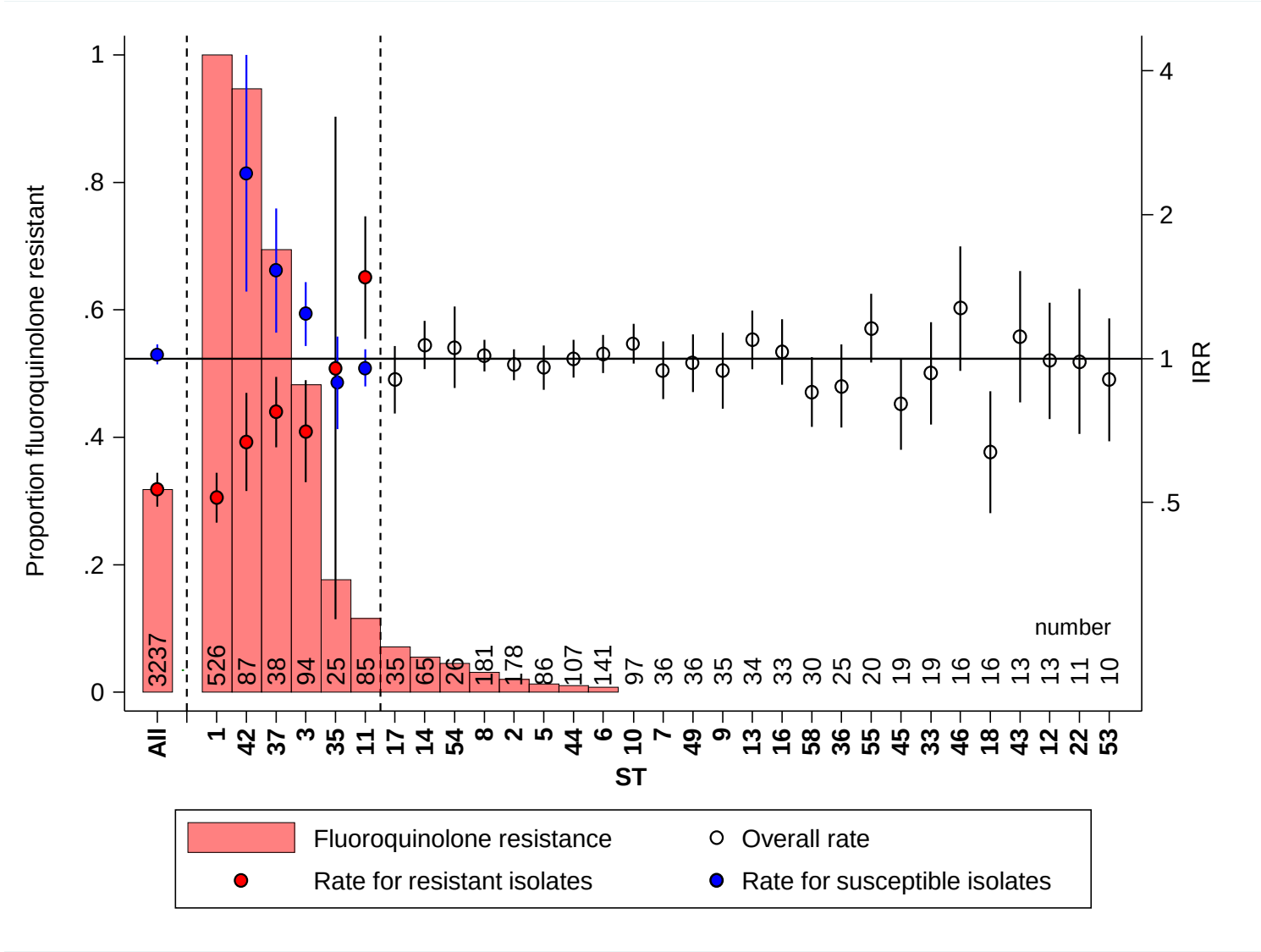
Green line: number of cases (per month) predicted by a Poisson model, (with time as the only covariate), modelling FQ resistant cases (blue) to illustrate declining incidence.

Incidence of FQ resistant genotypes has declined (2)



Green line: number of cases (per month) predicted by a Poisson model, (with time as the only covariate), modelling FQ resistant cases (blue) to illustrate declining incidence.

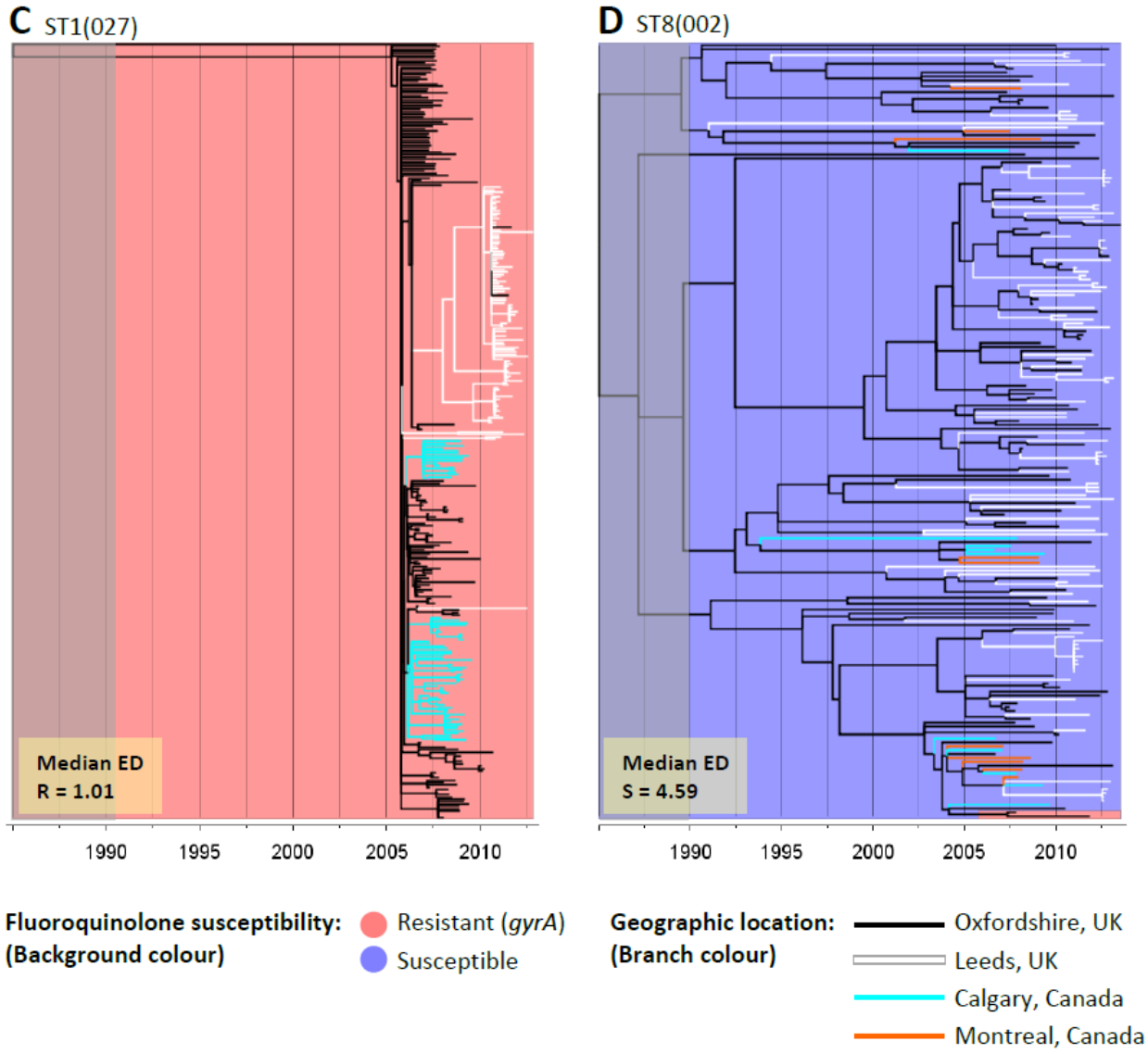
Changes in quinolone resistance over time



UNIVERSITY OF OXFORD

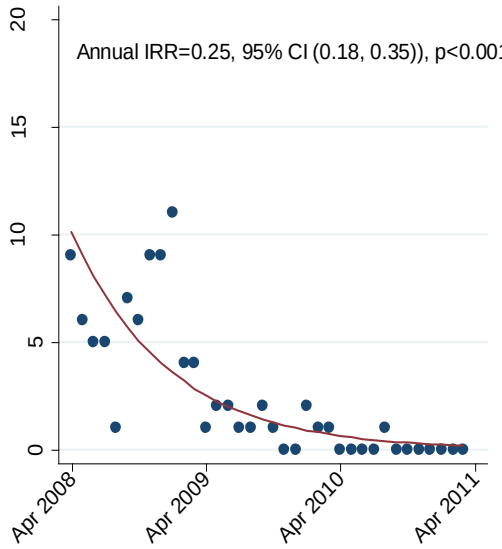
Public Health England

Phylogenetic patterns of quinolone resistant vs susceptible

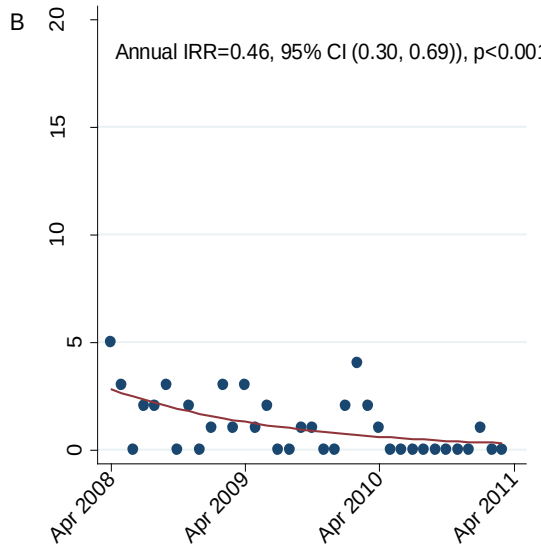


Declines in incidence varying by category

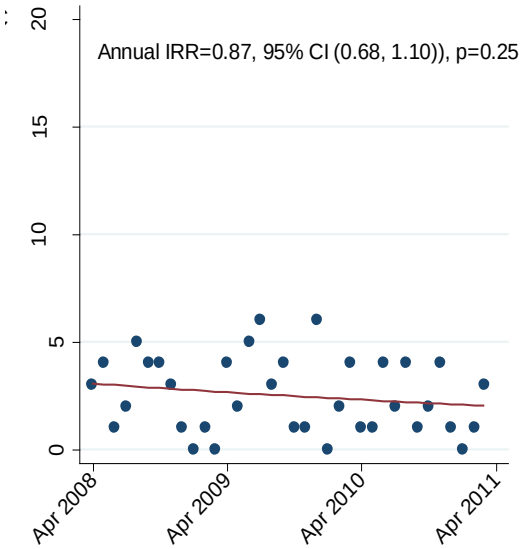
Fluoroquinolone resistant cases with a hospital link



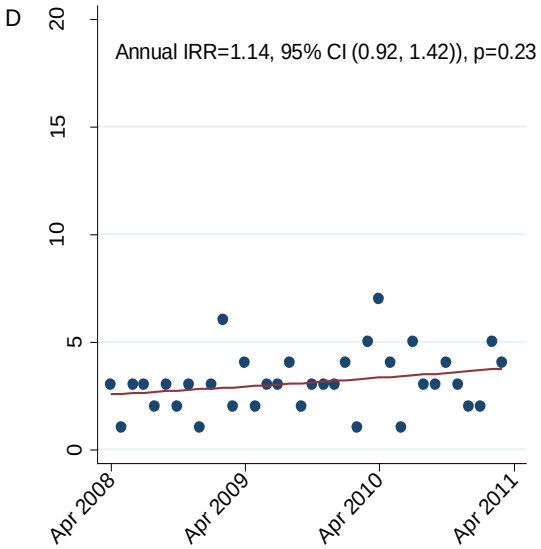
Fluoroquinolone resistant cases with no hospital link



Fluoroquinolone susceptible cases with a hospital link



Fluoroquinolone susceptible cases with no hospital link



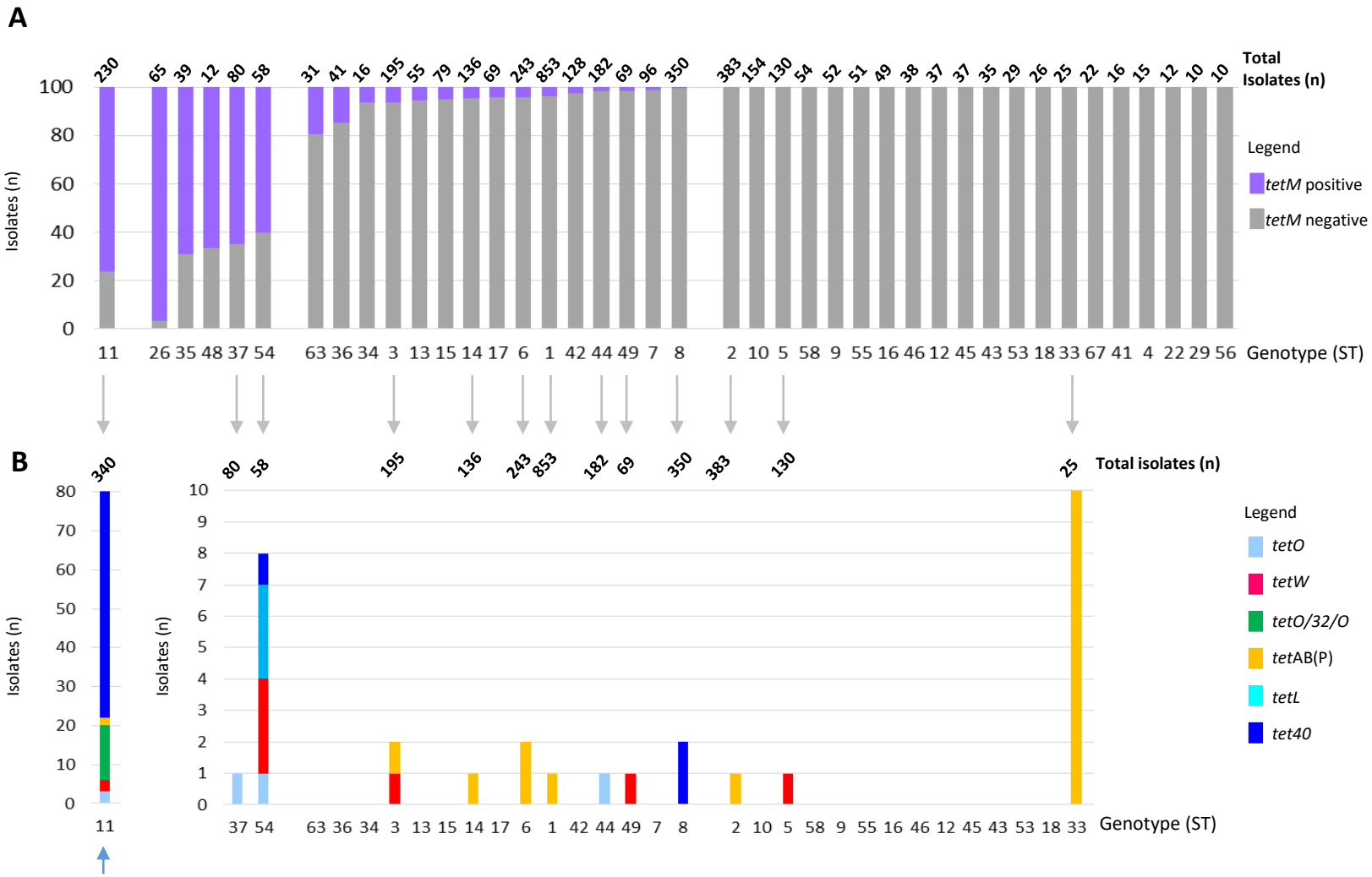
UNIVERSITY OF OXFORD

Public Health England

The decline of *C. difficile* in England

- It has declined by close to 70% since 2006
- Quinolone use declined by ~ 50% preceding the decline in CDI
- The decline is attributable to the simultaneous disappearance of 4 quinolone resistant lineages. The remaining 69 lineages are largely unchanged in incidence
- Resistant lineages had undergone rapid clonal expansion and were geographically structured
- A quinolone effect is a likely explanation for the decline in CDI

Tetracycline resistance in *C. difficile*

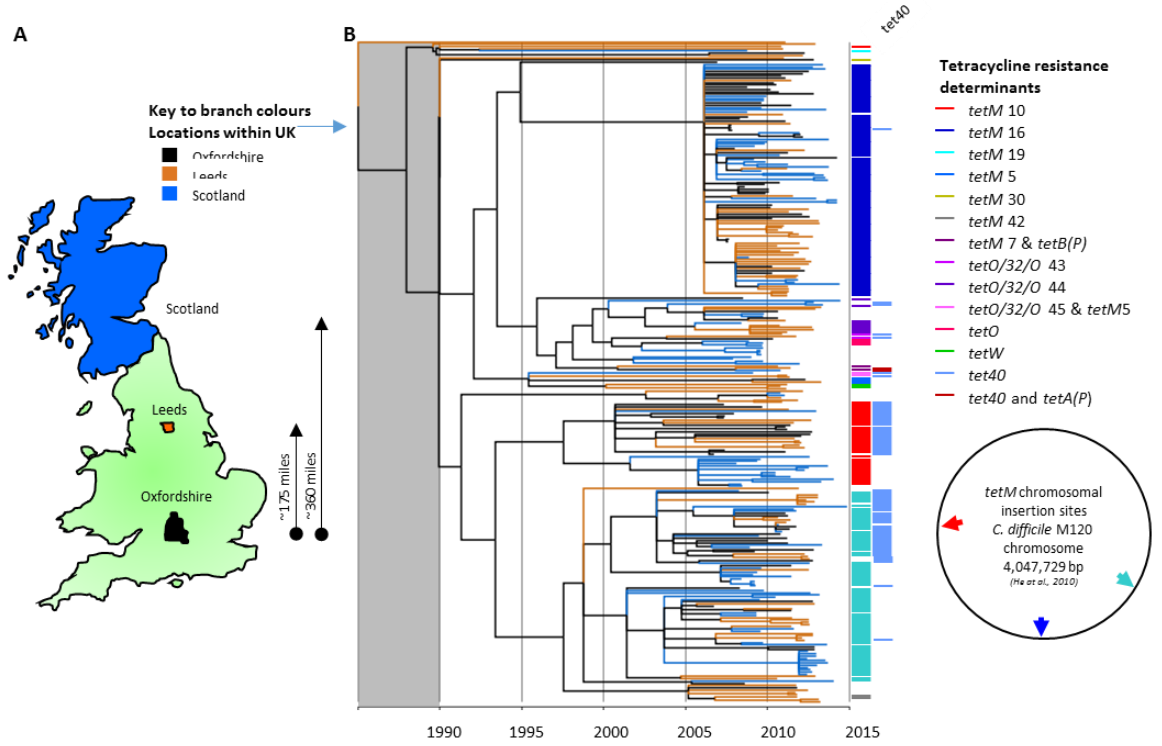


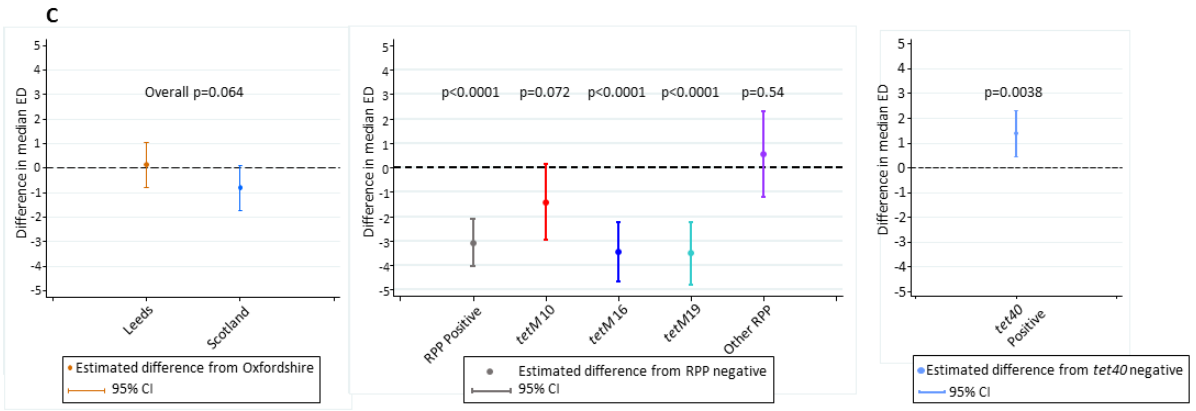
ST11 isolates examined for (A) (n=230; Oxford EIA positive and negative, Oxford infant and farm, Leeds, Optimer European and N. American)), plus additional Scottish ST11s (n=110)) to illustrate the overall prevalence of 'non-*tetM*' tetracycline resistance determinants among 340 genomes.

UNIVERSITY OF OXFORD

Public Health England

Tetracycline resistance in *C. difficile*





UNIVERSITY OF OXFORD

Public Health England

Tetracycline resistance in *C. difficile*

- Tetracycline particularly TetM alleles are characterised by:
  - Tn916 is rarely acquired
  - Strong selection and dispersal of clades across geographies

- Quinolone resistance mutation occurs frequently and is not under selection within the ribotype 078 lineage

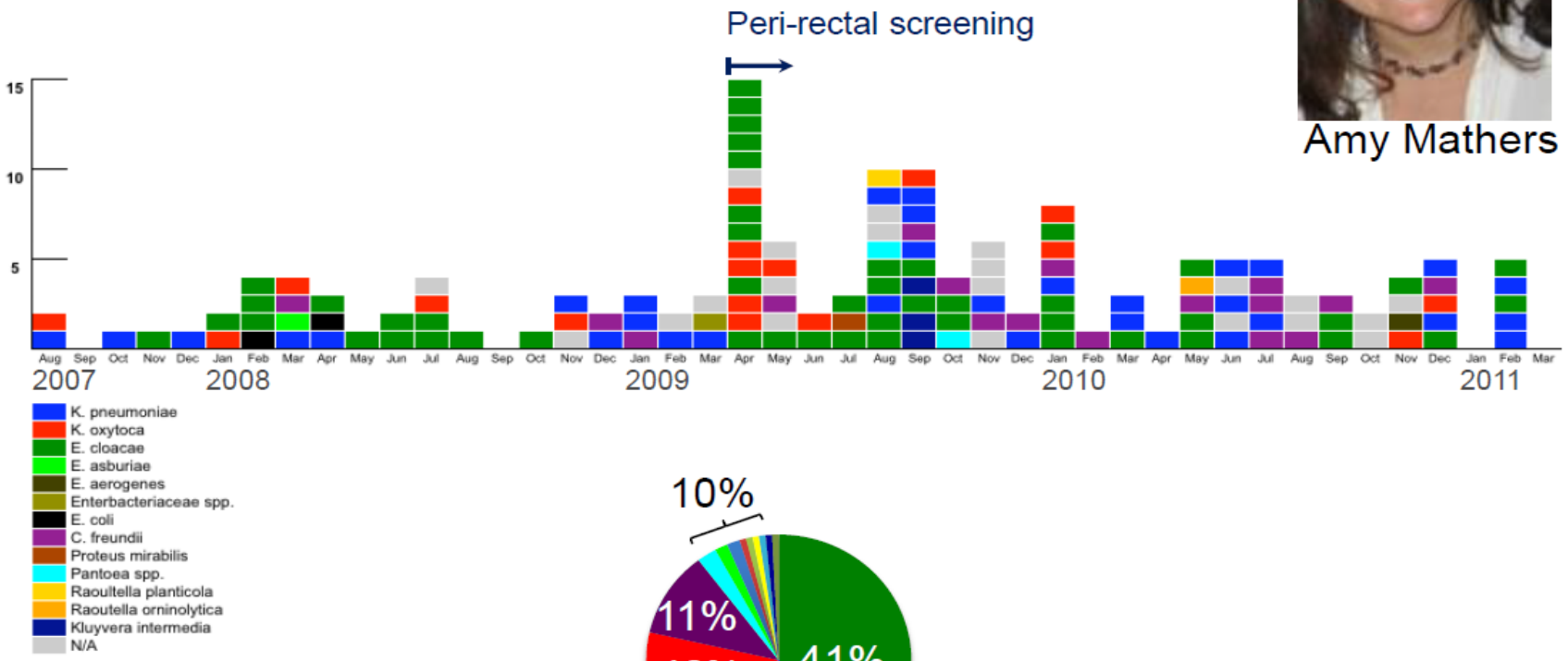
- Reservoir likely farm animals e.g. pigs (reported in the Netherlands) and other meat producing animals

Biological processes impacting on KPC AMR

A single hospital

Virginia KPC Outbreak

Amy Mathers



Antimicrob. Agents Chemother; April 2016

UNIVERSITY OF OXFORD

Public Health England

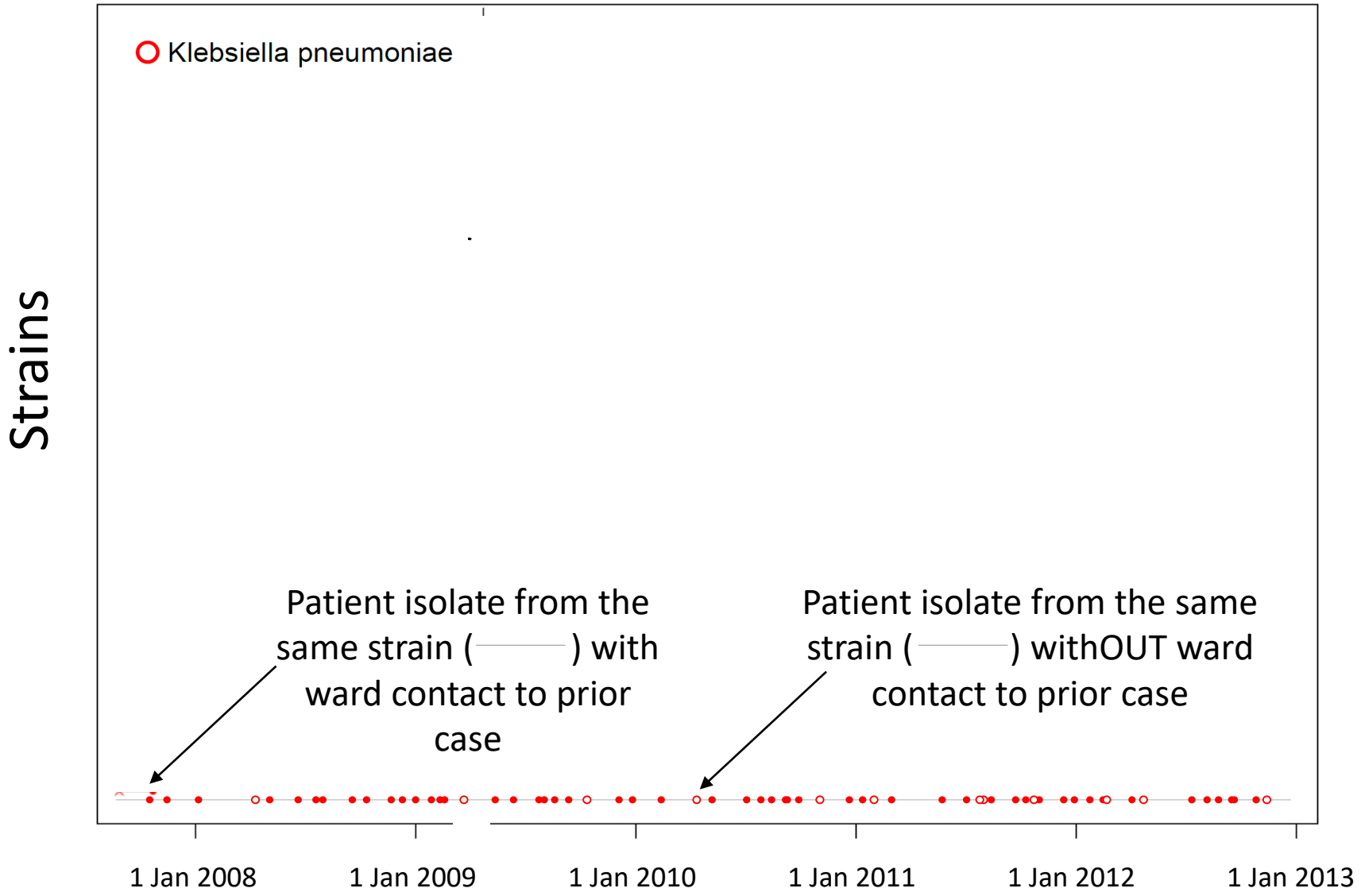
bla$_{KPC}$ in Virginia

- Virginia “outbreak” – ongoing since August 2007

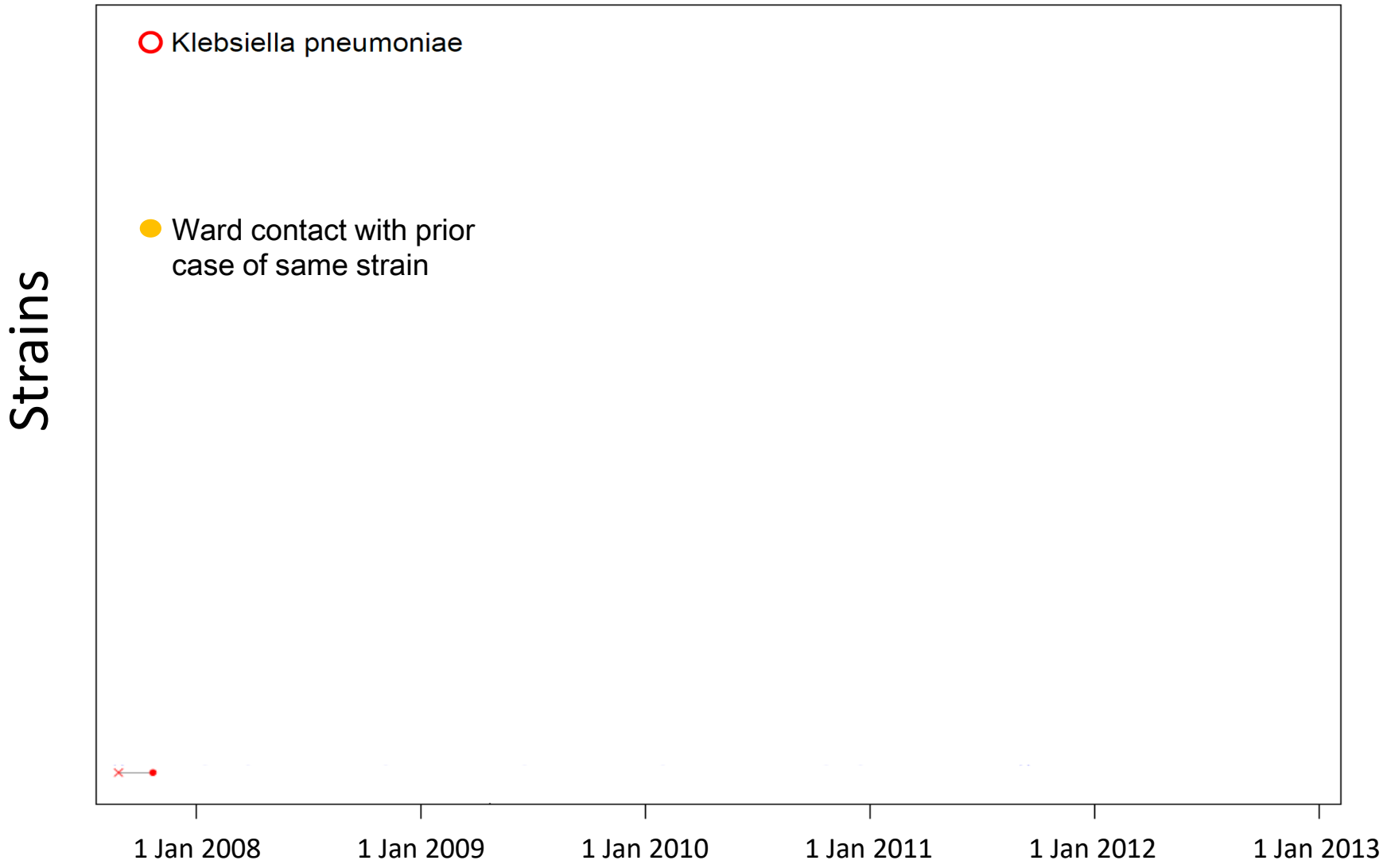
- 281 *bla*$_{KPC}$-positive Enterobacteriaceae
  - Isolated August 2007 – December 2012
  - From 182 patients
  - All Illumina sequenced

- Multiple species of *bla*$_{KPC}$-positive Enterobacteriaceae
  - 9 different genera
  - 13 different species
  - 62 different “strains”
    (defined conservatively as ~500 SNPs variation in “core”)

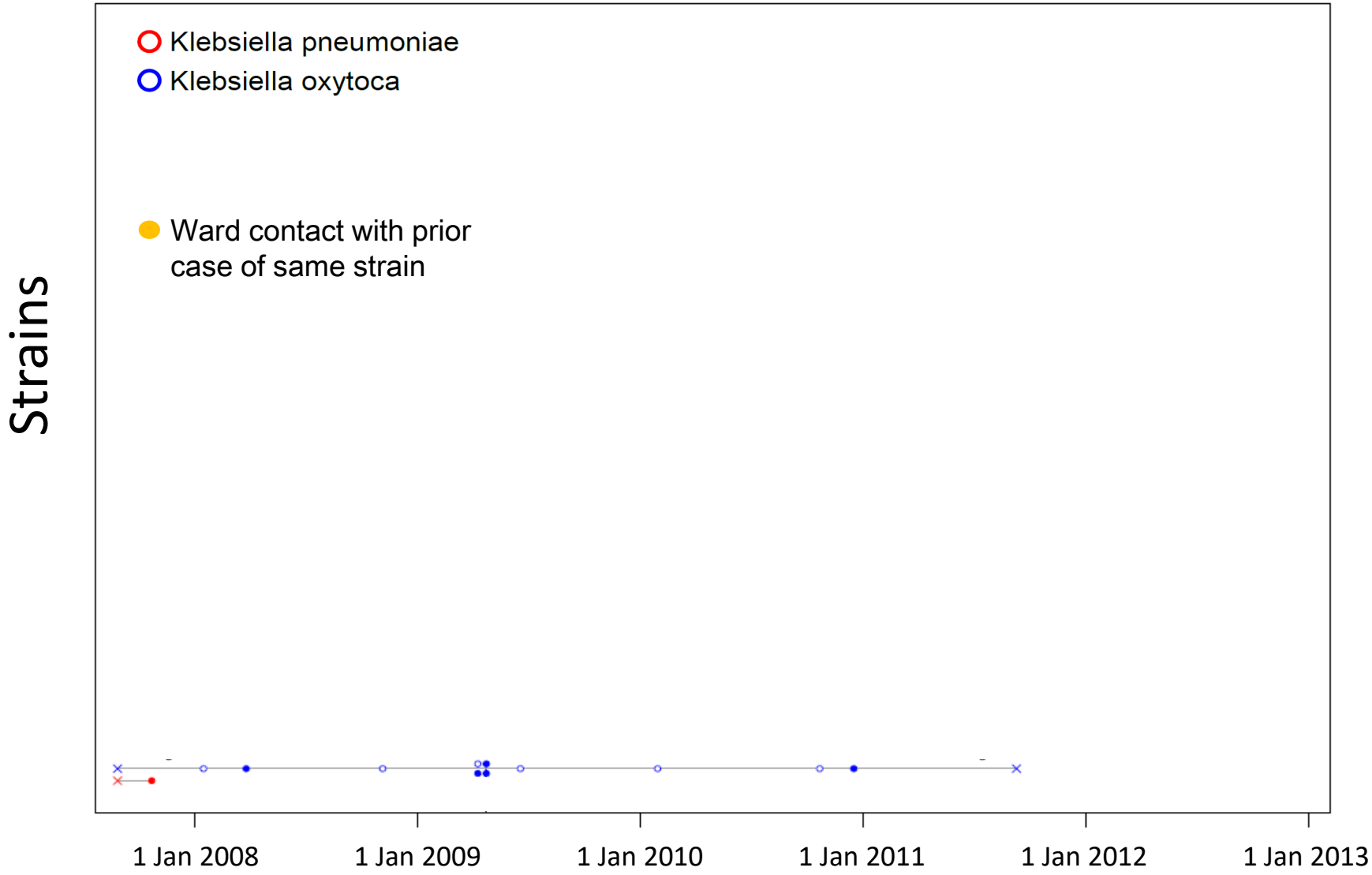
Idealised outbreak timeline



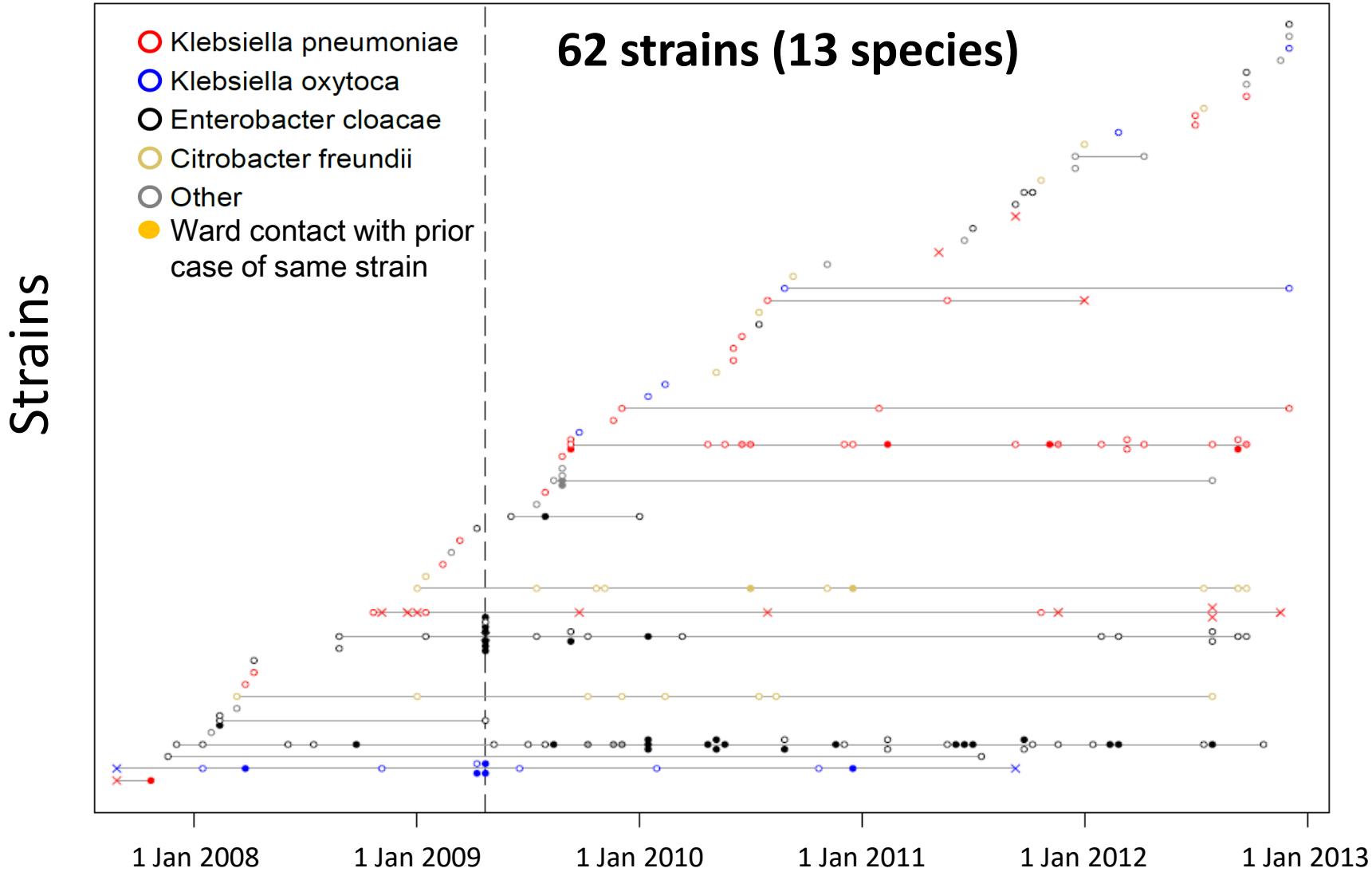
Actual outbreak timeline



Actual outbreak timeline



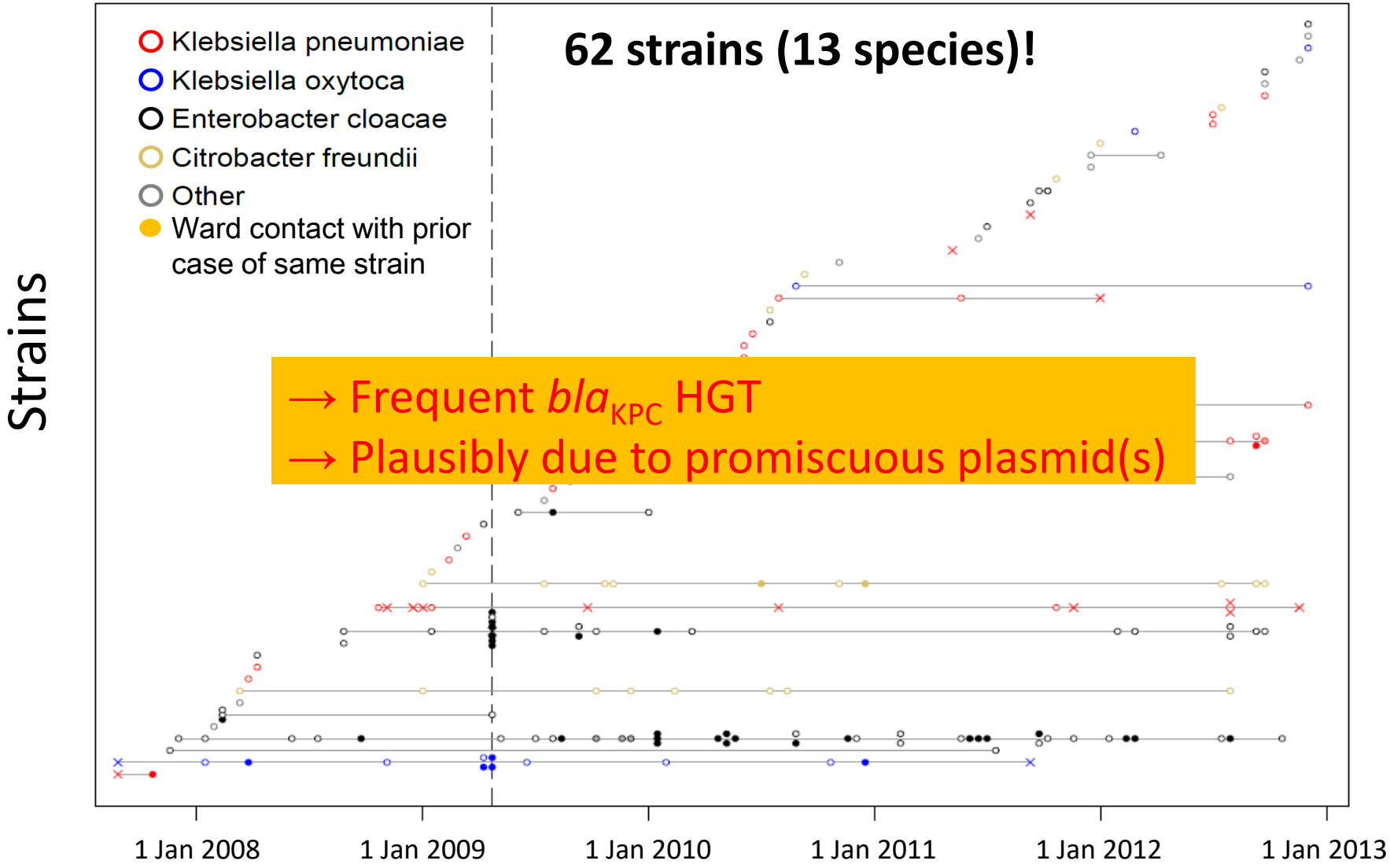
Enormous host strain diversity



UNIVERSITY OF OXFORD

Public Health England

Enormous host strain diversity



Plasmid-mediated outbreak?

- Hypothesis: outbreak is driven by one or a few promiscuous plasmids carrying *bla*$_{KPC}$
- Assumption: plasmid structures relatively stable within outbreak
- Approach:
  - Generate outbreak-specific plasmid references (index patient)
  - Use these to assess plasmid presence across outbreak isolates
  - **Definition: ≥99% sequence identity over ≥80% reference length**
    - Assessed via BLASTn (reference plasmid vs isolate's *de novo* assembly)
    - Stringent identity threshold: expect few SNP changes
    - Lenient length threshold: single events can affect large regions

    - Note: does not assess structural continuity (since this is impossible in many isolates due to repeat structures)

Spread of index plasmids

- Two *bla*$_{KPC}$ conjugative plasmids from index patient
  - pKPC_UVA01 (43,621 bp) and pKPC_UVA02 (113,105 bp)

Spread of index plasmids

- Two *bla*$_{KPC}$ conjugative plasmids from index patient
  - pKPC_UVA01 (43,621 bp) and pKPC_UVA02 (113,105 bp)

Species	Isolates
Citrobacter amalonaticus	2
Citrobacter freundii	30
Enterobacter aerogenes	4
Enterobacter asburiae	1
Enterobacter cloacae	96
Escherichia coli	2
Klebsiella oxytoca	35
Klebsiella pneumoniae	94
Kluyvera intermedia	7
Proteus mirabilis	1
Raoultella ornothinolytica	1
Serratia marcescens	5
Other (unknown)	3
Total	**281**

Spread of index plasmids

- Two *bla*$_{KPC}$ conjugative plasmids from index patient
  - pKPC_UVA01 (43,621 bp) and pKPC_UVA02 (113,105 bp)

Species	Isolates	pKPC_UVA01
Citrobacter amalonaticus	2	1
Citrobacter freundii	30	29
Enterobacter aerogenes	4	2
Enterobacter asburiae	1	0
Enterobacter cloacae	96	84
Escherichia coli	2	1
Klebsiella oxytoca	35	9
Klebsiella pneumoniae	94	31
Kluyvera intermedia	7	7
Proteus mirabilis	1	1
Raoultella ornothinolytica	1	1
Serratia marcescens	5	0
Other (unknown)	3	0
Total	**281**	**166 (59%)**

Spread of index plasmids

- Two *bla*$_{KPC}$ conjugative plasmids from index patient
  - pKPC_UVA01 (43,621 bp) and pKPC_UVA02 (113,105 bp)

Species	Isolates	pKPC_UVA01	pKPC_UVA02
Citrobacter amalonaticus	2	1	0
Citrobacter freundii	30	29	7
Enterobacter aerogenes	4	2	0
Enterobacter asburiae	1	0	0
Enterobacter cloacae	96	84	2
Escherichia coli	2	1	0
Klebsiella oxytoca	35	9	25
Klebsiella pneumoniae	94	31	18
Kluyvera intermedia	7	7	0
Proteus mirabilis	1	1	0
Raoultella ornothinolytica	1	1	0
Serratia marcescens	5	0	0
Other (unknown)	3	0	0
Total	**281**	**166 (59%)**	**52 (19%)**

Spread of index plasmids

- Two *bla*$_{KPC}$ conjugative plasmids from index patient
  - pKPC_UVA01 (43,621 bp) and pKPC_UVA02 (113,105 bp)

Species	Isolates	pKPC_UVA01	pKPC_UVA02	Neither
Citrobacter amalonaticus	2	1	0	1
Citrobacter freundii	30	29	7	1 (3%)
Enterobacter aerogenes	4	2	0	2
Enterobacter asburiae	1	0	0	1
Enterobacter cloacae	96	84	2	10 (10%)
Escherichia coli	2	1	0	1
Klebsiella oxytoca	35	9	25	1 (3%)
Klebsiella pneumoniae	94	31	18	45 (48%)
Kluyvera intermedia	7	7	0	0
Proteus mirabilis	1	1	0	0
Raoultella ornothinolytica	1	1	0	0
Serratia marcescens	5	0	0	5
Other (unknown)	3	0	0	3
Total	**281**	**166 (59%)**	**52 (19%)**	**70 (25%)**

← mostly known endemic clone previously described with other plasmids

→ Consistent with local plasmid-mediated outbreak, plus occasional imports from other healthcare institutions

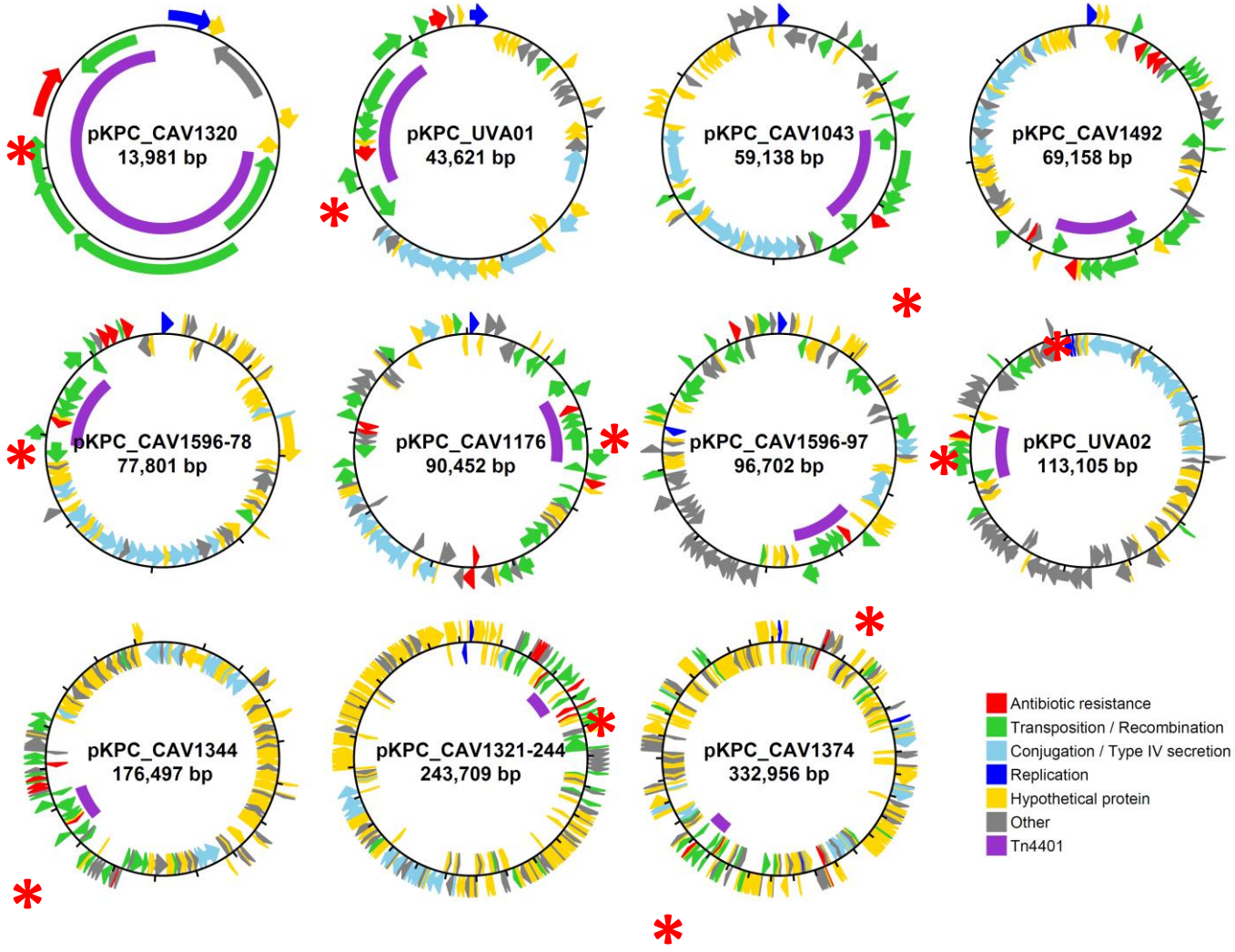
Long-read sequencing

- Needed to validate conclusions, given structural uncertainties of short-read WGS

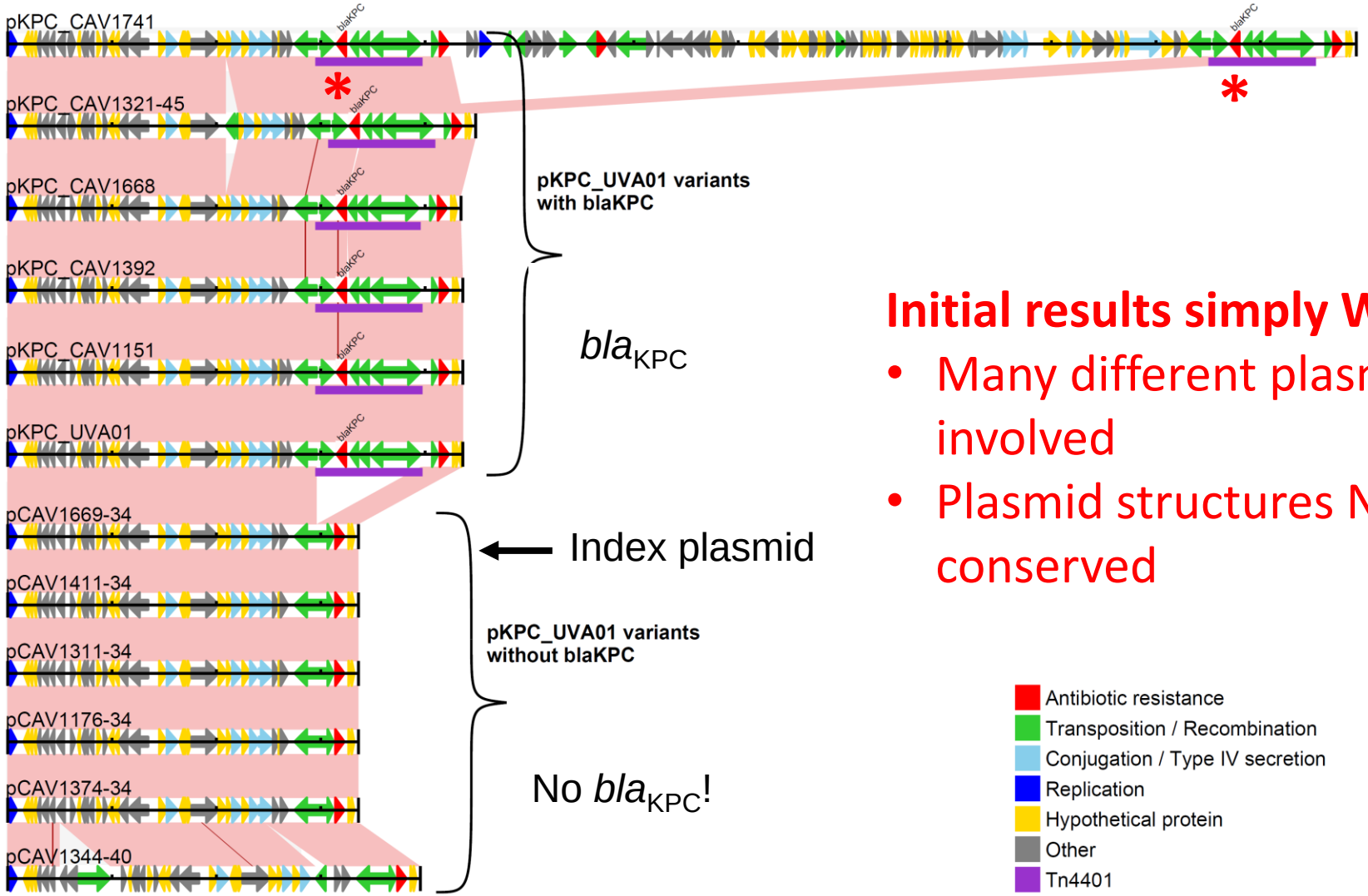
- PacBio sequencing
  - 17 **randomly chosen** isolates
  - Fully closed plasmid structures

11 different *bla*$_{KPC}$ (*) plasmids!

14kb
to
330kb



Structural diversity of pKPC_UVA01



Initial results simply WRONG:

- Many different plasmids involved
- Plasmid structures NOT conserved

UNIVERSITY OF OXFORD

Public Health England

A highly dynamic dispersal of KPC within the clinical ecosystem

- KPC dispersing at 3 scales:
  - Isolates spreading KPC between patients
  - Frequent transfer of *bla*$_{KPC}$ plasmids between strains/species
  - Frequent transfer of *bla*$_{KPC}$ transposon Tn4401 between plasmids

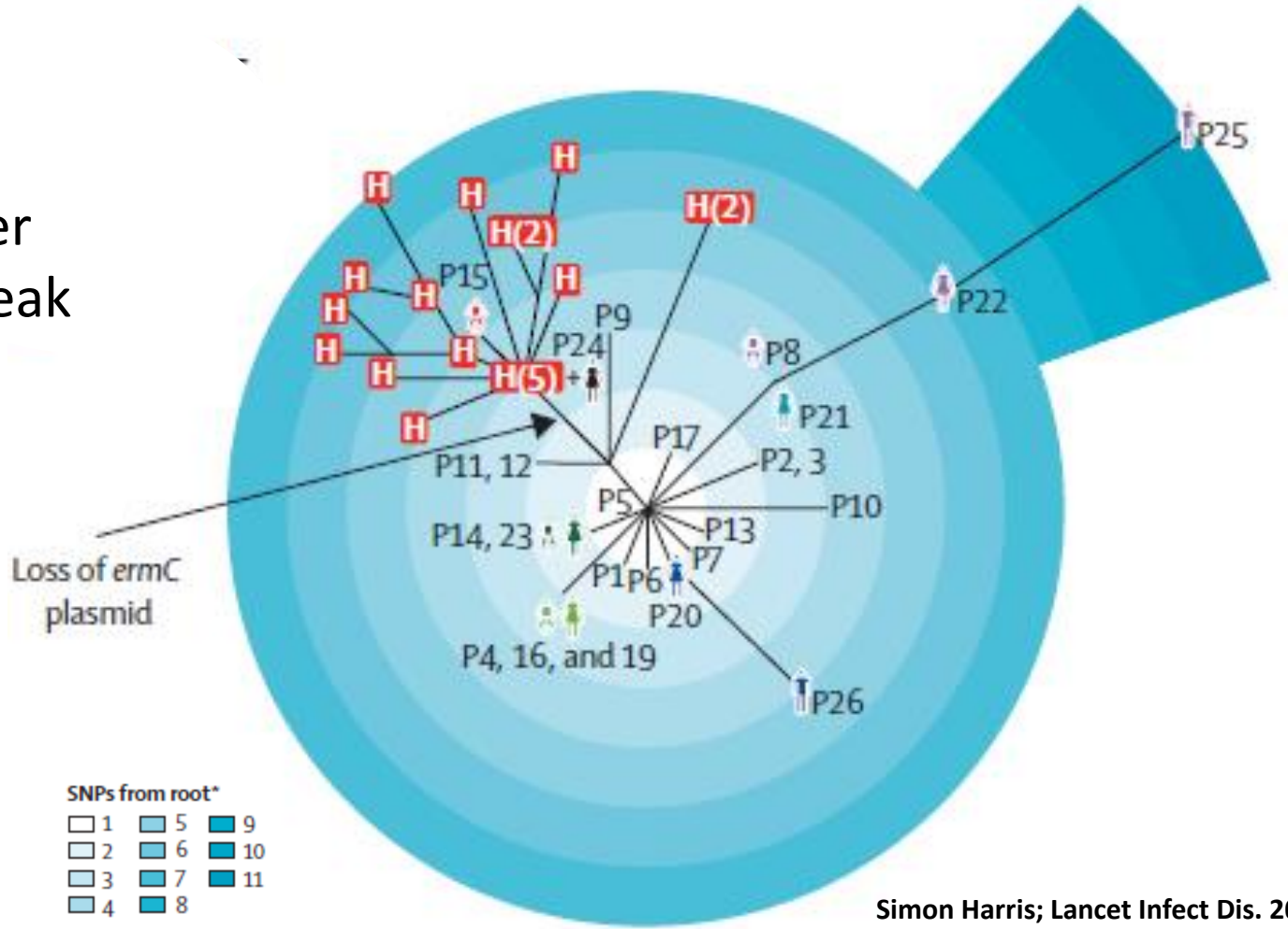
Example pathogens

- All are relevant to reference services, but could become part of routine service

Staphylococcus aureus

Transmission on a neonatal intensive care unit

Healthcare worker associated outbreak



Simon Harris; Lancet Infect Dis. 2013 Feb;13(2):130-6

UNIVERSITY OF OXFORD

Public Health England

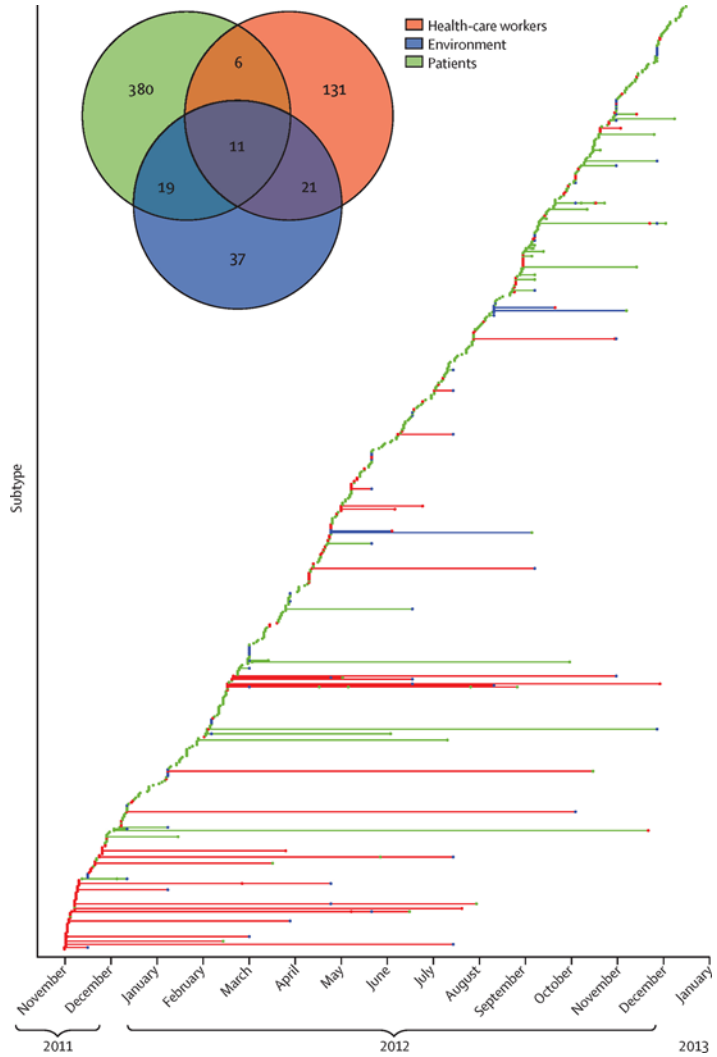
Transmission on an adult general ICU

Identification of *Staphylococcus aureus* in health-care workers and patients admitted to the intensive care and high-dependency unit

	Nurses (n=149)	Doctors (n=40)	Physiotherapists (n=9)	Total health-care workers (n=198)	Patient admissions (n=1933)
Age, years					
16–29	40 (27%)	9 (23%)	8 (89%)	57 (29%)	154 (8%)
30–39	71 (48%)	17 (43%)	1 (11%)	89 (45%)	150 (8%)
40–49	28 (19%)	11 (28%)	0	39 (20%)	208 (11%)
50–59	9 (6%)	3 (8%)	0	12 (6%)	264 (14%)
≥60	1 (1%)	0	0	1 (1%)	1157 (60%)
Male sex	25 (17%)	24 (60%)	2 (22%)	51 (26%)	1164 (60%)
Nasal carriage at enrolment	54 (36%)	16 (40%)	3 (33%)	73 (37%)	386* (21%)
MRSA	8 (5%)	0	0	8 (4%)	39 (2%)
Total acquisitions of *S aureus* during study	60	5	4	69	97
MRSA	3 (5%)	0	1 (25%)	4 (6%)	19 (20%)
Culture negative to positive acquisitions during study	31	5	4	40	68
MRSA	0	0	1 (25%)	1 (3%)	14 (21%)
Culture positive to new subtype acquisitions during study	29	0	0	29	29
MRSA	3 (10%)	0	0	3 (10%)	5 (17%)
Acquisition isolates available for whole-genome sequencing	59 (98%)	5 (100%)	4 (100%)	68 (99%)	86 (89%)

Data are n (%) or n. MRSA=meticillin-resistant *S aureus*.

*Of 1854 patients receiving at least one screen.



7 HCW to patient transmissions over 14 months of observation
2 Environmental samples transmitted to patients
Most acquisitions were between HCW

Lancet Infect Dis. 2017 Feb;17

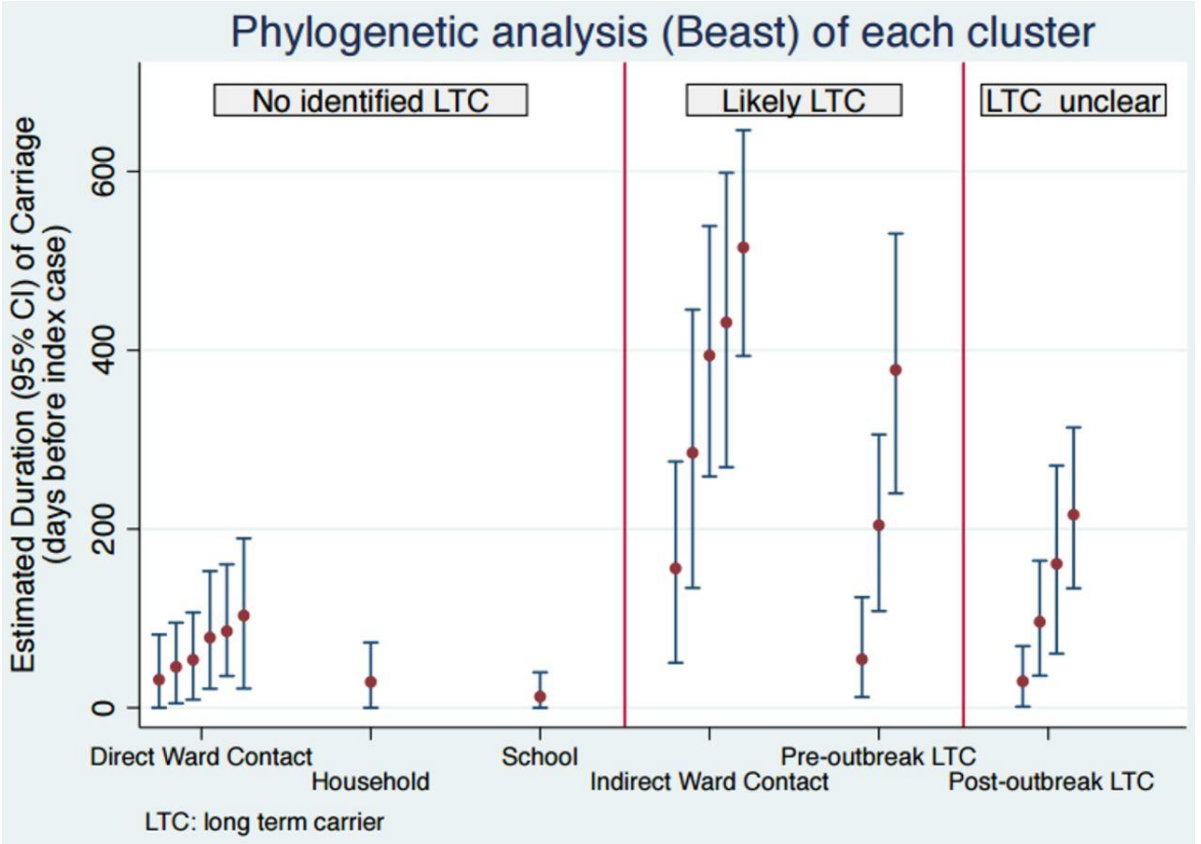
UNIVERSITY OF OXFORD

Public Health England

Duration-adjusted TMRCA for outbreaks with (i) no evidence of a long-term carrier (direct contacts between all cases); (ii) likely LTC (indirect ward contacts or preoutbreak LTC); or (iii) LTC unclear/possible (evidence of a postoutbreak LTC).

Description of 20 outbreaks analyzed by whole-genome sequencing

Outbreak	Epidemiological category	No of cases	Reason for outbreak investigation	MRSA or MSSA	Clonal complex	MLST[a]	*spa*	Dura (day
A	Hospital, general ward	5	MRSA colonization	MRSA	CC22	ST22	t032	367
B	Hospital, general ward	6	*S. aureus* wound infections	MSSA	CC8	ST2021	t008	412
C	Hospital, general ward	7	*S. aureus* wound infections	MRSA	CC8	ST239	t037	98
D	Hospital, general ward	17	MRSA colonization	MRSA	CC8	ST8	t008	88
E	Hospital, surgical unit	8	*S. aureus* wound infections	MRSA	CC22	ST22	t022	18
F	Hospital, multiple wards	50	MRSA colonization	MRSA	CC5	ST228	t041	122
G	Hospital, multiple wards	187	MRSA colonization	MRSA	CC8	ST8	t008	454
H	Hospital, maternity unit	6	PVL–related SSTIs[b]	MRSA	CC1	ST772	t657	70
I	Hospital, maternity unit	9	Scalded skin syndrome	MSSA	CC15	ST2434	t346	70
J	Hospital, neonatal unit	3	MRSA colonization	MRSA	CC59	ST59	t216	8
K	Hospital, neonatal unit	4	MRSA colonization	MRSA	CC22	ST22	t5892	43
L	Hospital, neonatal unit	6	MRSA colonization	MRSA	CC30	ST30	t019	57
M	Hospital, neonatal unit	8	MRSA bacteremia	MRSA	CC88	ST88	t5973	65
N	Hospital, neonatal unit	41	MRSA colonization	MRSA	CC22	ST22	t5892	152
O	Household	3	PVL–related SSTIs	MRSA	CC30	ST30	t019	8
P	Household	4	PVL–related SSTIs	MRSA	CC30	ST30	t019	20
Q	Household	5	PVL–related SSTIs	MRSA	CC30	ST30	t019	195
R	Household	8	PVL–related SSTIs	MRSA	CC30	ST30	t019	44
S	Nursing home	9	PVL–related SSTIs	MRSA	CC30	ST30	t019	298
T	School	5	PVL–related SSTIs	MSSA	CC121	ST121	t645	74



Outbreaks related to long term carriers have greater diversity across cases supported to longer periods to the most recent common ancestor

N. C. Gordon et al. J. Clin. Microbiol. 2017;55:2188-2197

Clostridium difficile

Mail Online health

Deaths from hospital superbug C.diff soar by a THIRD in just one year

'It is absolutely outrageous that year on year older people are dying as a result of failure to ensure high levels of hygiene in our hospitals.

'People go into hospital expecting to be looked after, not to run the risk of secondary infections as a result of poor hygiene.'

risen 28 per cent.

Altogether last year, 8,324 people in England and Wales had clostridium difficile when they died.

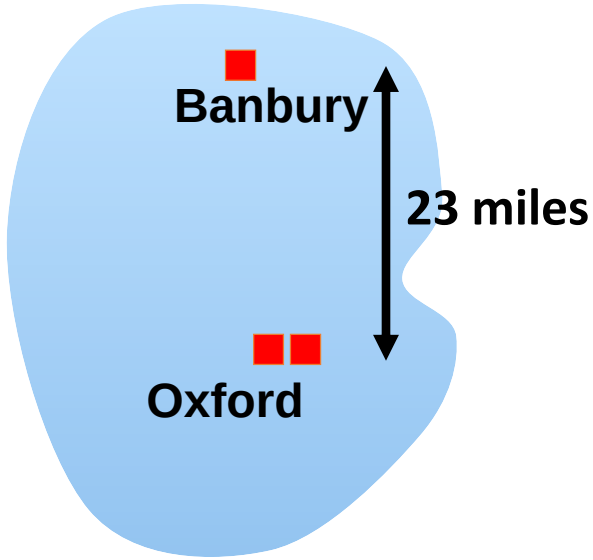
This compares with 6,480 in 2006, with the infection noted as the underlying cause in about half of deaths, according to the Office for National Statistics.

But the number of reported deaths involving the infection has more than doubled since 2005, when there were 3,757.

The ONS said some of the increase may be due to more complete reporting on death certificates after the Government

Role of symptomatic patients in *C. difficile* transmission

- We sequenced 1223 of all 1251 hospital and community CDI cases (98%) in Oxfordshire, September 2007 – March 2011
- Hospital admission and ward movement data, and home postcode district and GP location available for each case



- 3 Hospitals
  - Typical CDI incidence
  - Infection control in line with published guidelines

Eyre: N Engl J Med 2013; 369:1195-1205

Applying sequencing

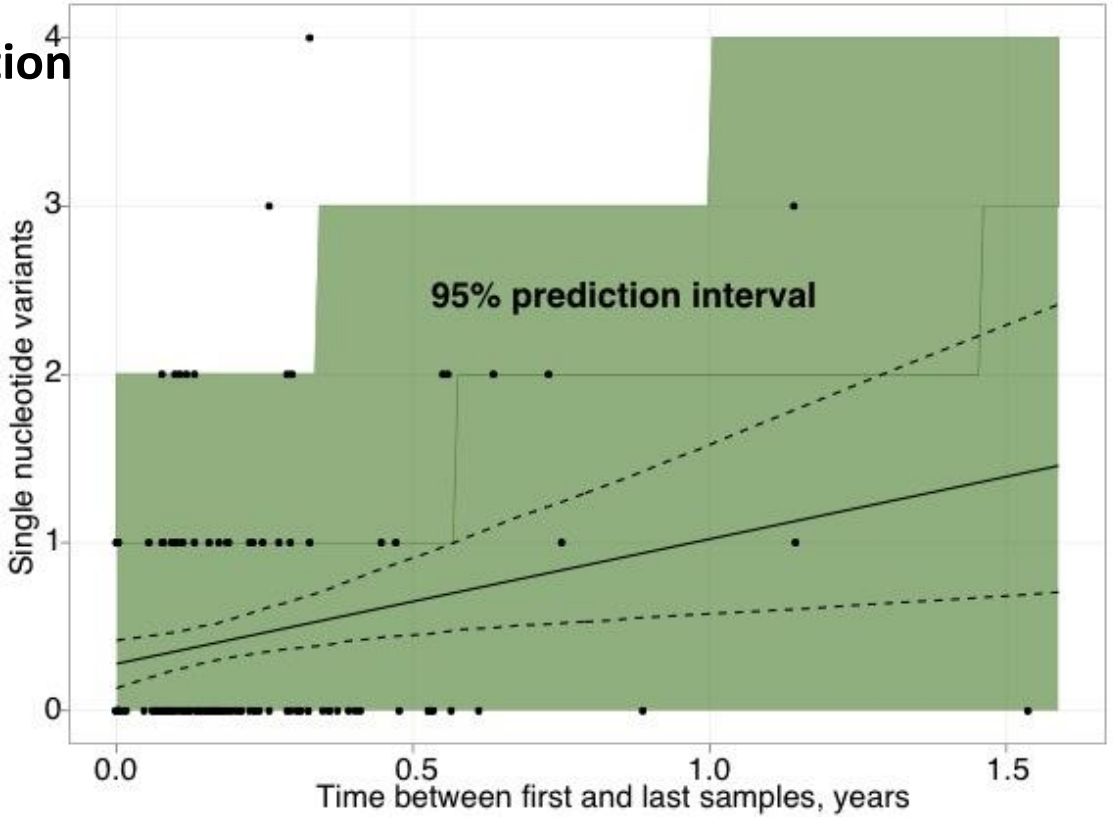
Reproducible sequencing

- 180 genomes sequenced more than once, 1 false SNV per 90 genomes

Within host diversity and evolution

- 0-2 SNVs expected between transmitted isolates up to 123 days apart

- > 10 SNVs likely to be unrelated with a time to most recent common ancestor of ~ 5 years

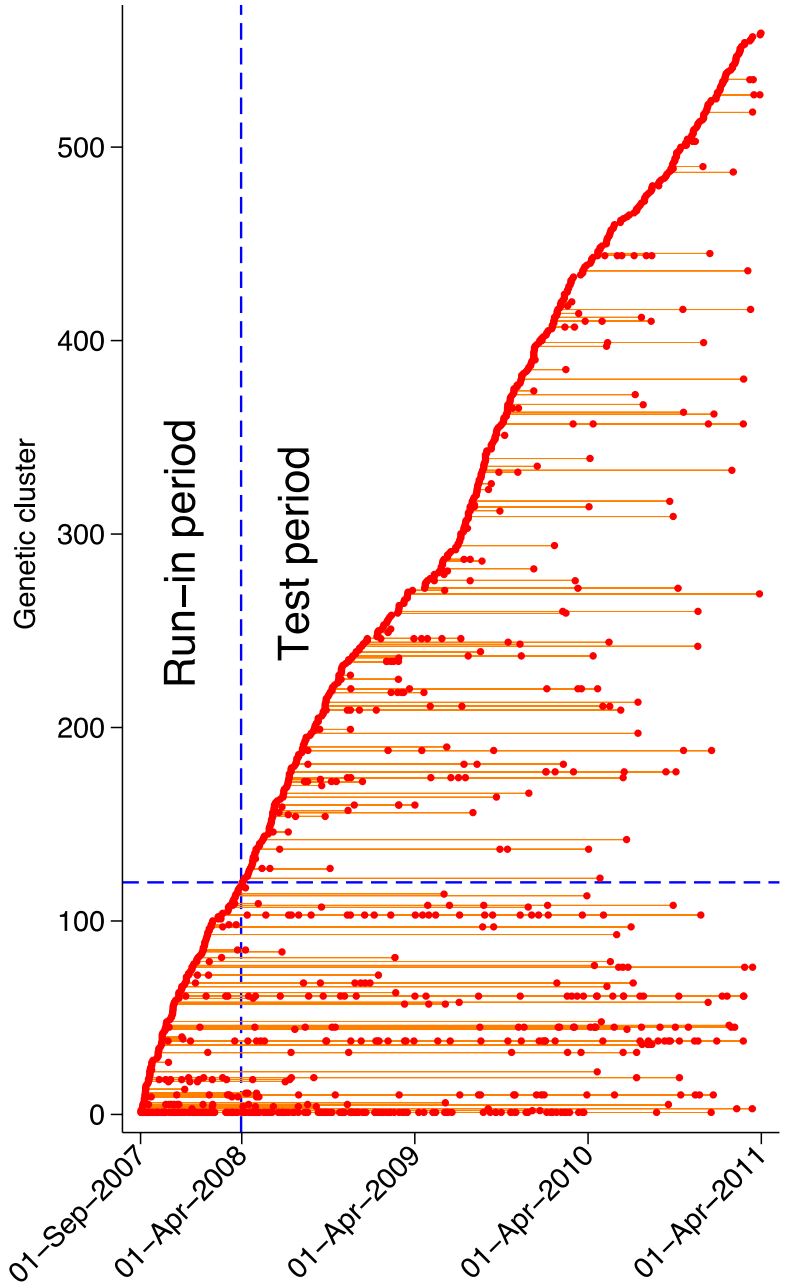


Diversity observed

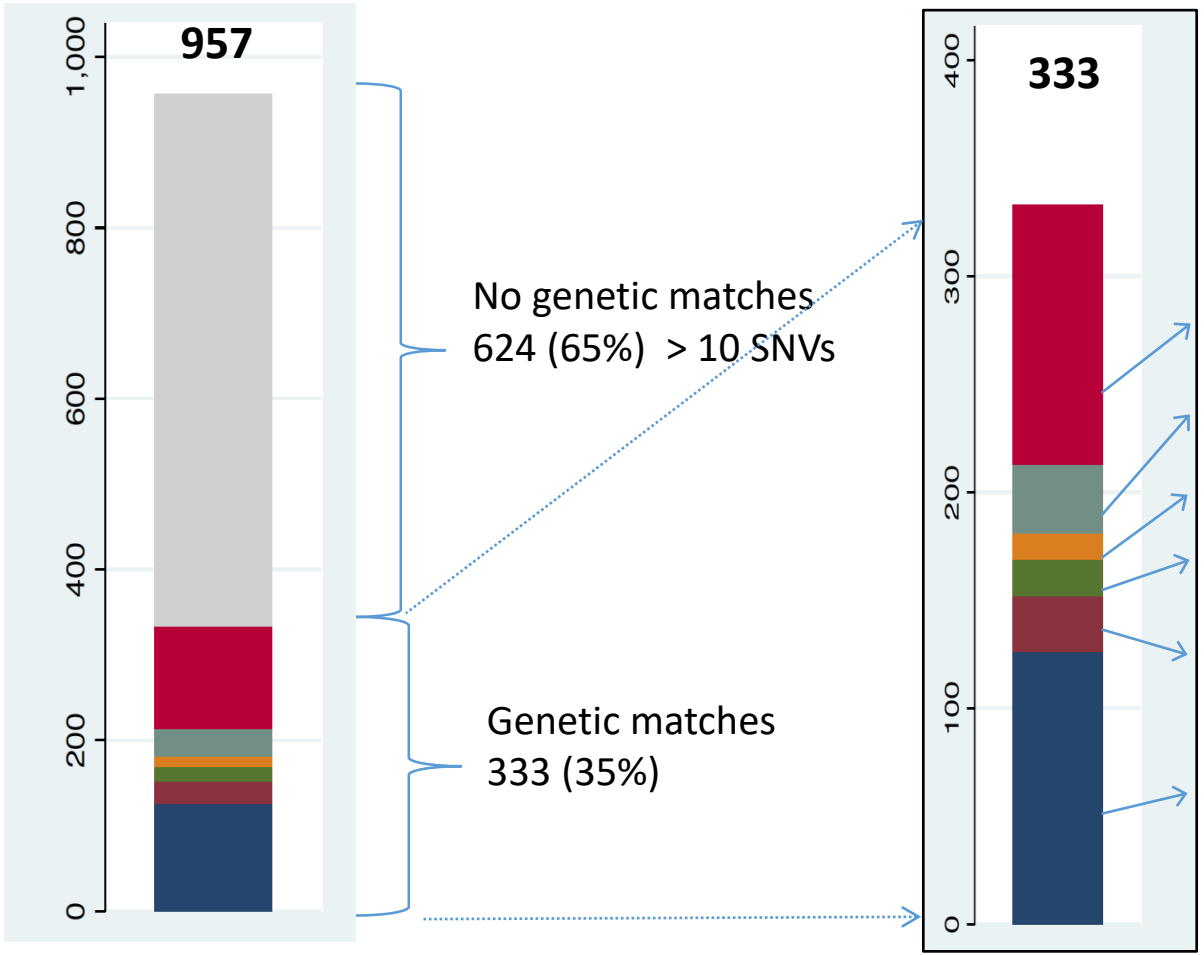
Throughout the study we continued to observe new genetic subtypes, >10 SNVs different to any seen before, at the same rate

⇨The reservoir is **LARGE**

⇨There are multiple exposures to it



Source of new *C. difficile* cases



No known contact	120 (12%)
Community	32 (3%)
Other same Hospital	12 (1%)
Medical specialty	17 (2%)
Indirect ward (‘spore’)	26 (3%)
Direct Ward contact	126 (13%)

All cases

Genetic Matches (0-2 SNV)

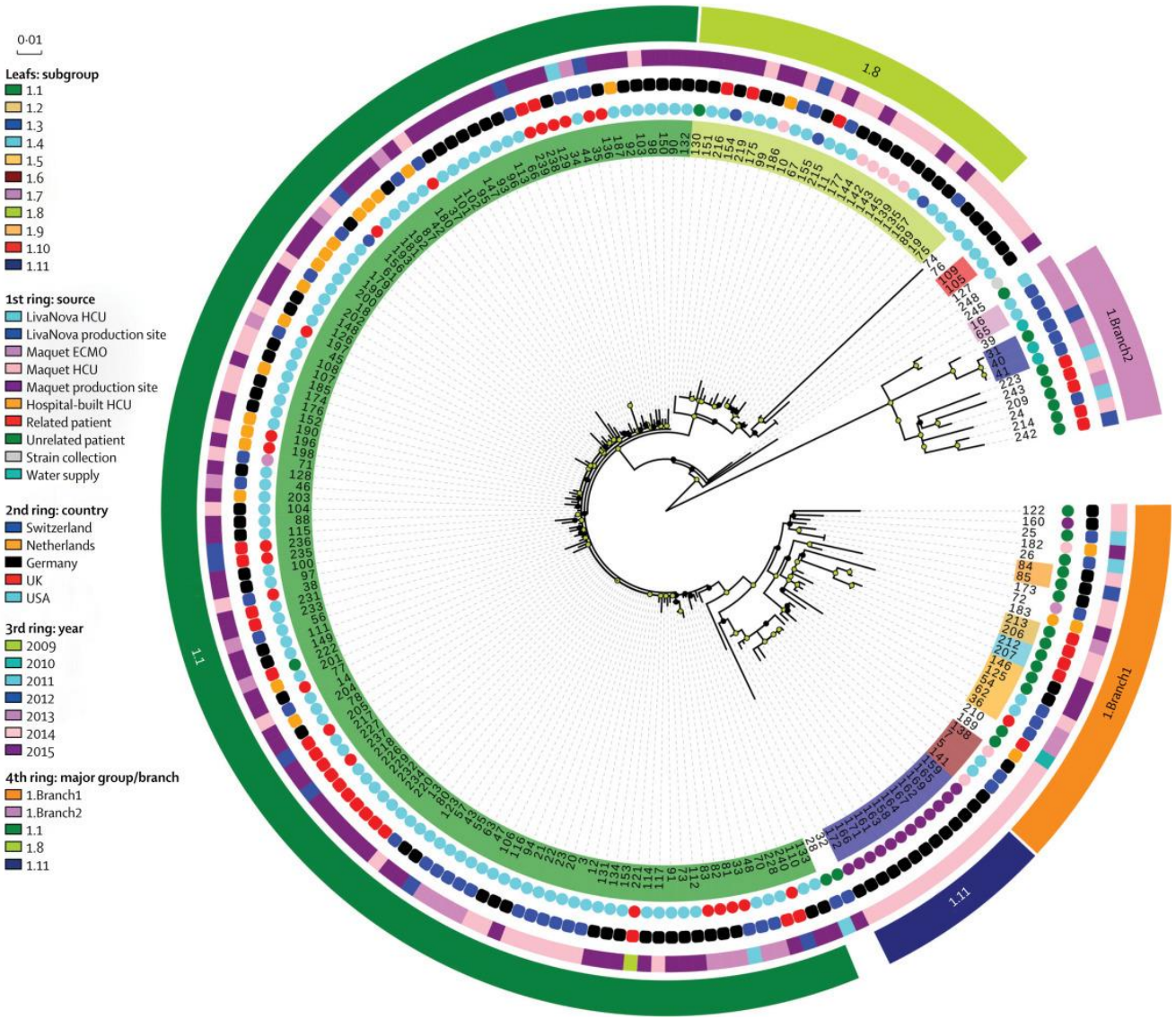
N Engl J Med 2013; 369:1195-1205

UNIVERSITY OF OXFORD

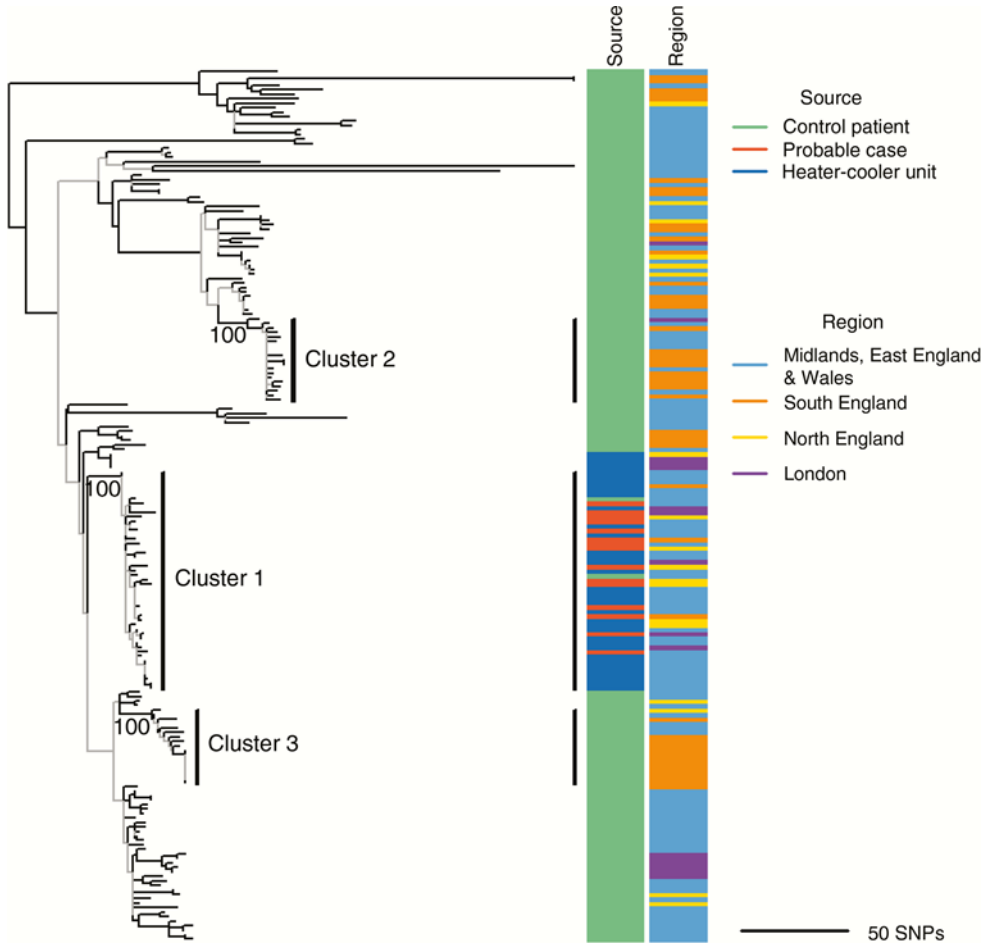
Public Health England

Mycobacterium chimera

Mycobacterium chimera



The Lancet Infectious Diseases, Available online 12 July 2017



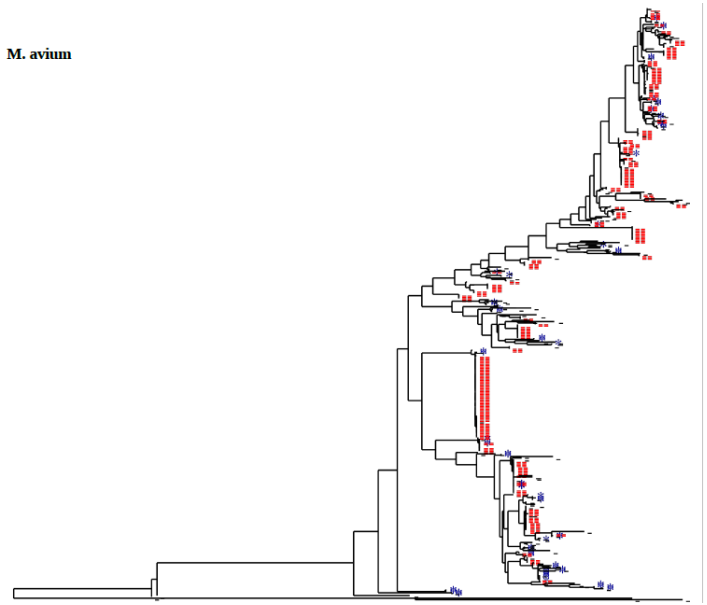
Clin Infect Dis. 2016;64(3):335-342

Clusters being no greater than ~20 SNPs apart

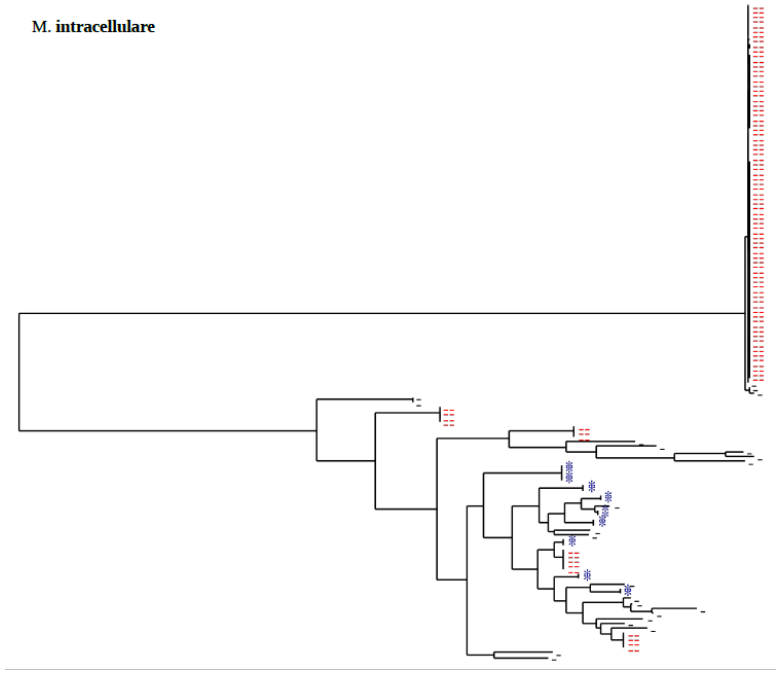
UNIVERSITY OF OXFORD

Public Health England

Three other NTMs



M. avium



M. intracellulare

M. abscessus

Clusters are all no greater than 38 SNPs apart

-- = 3 or more clustered isolates

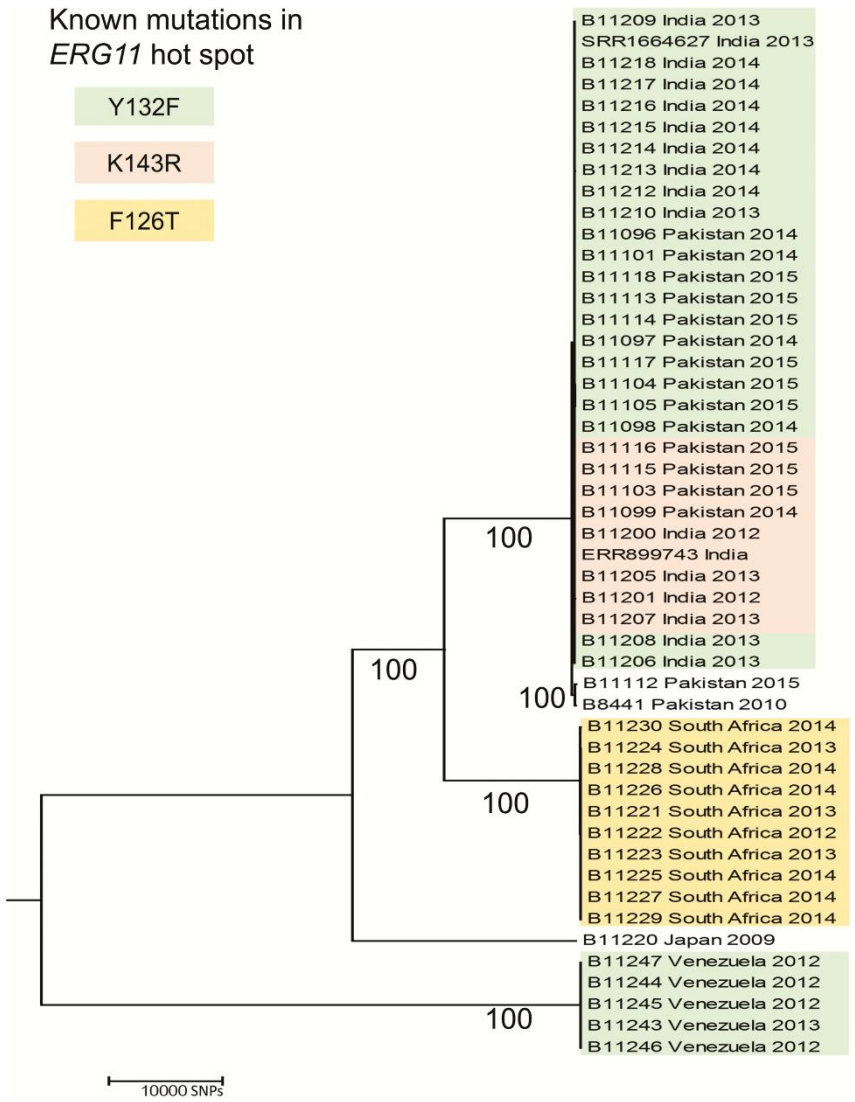
• = 2 clustered isolates

- = a single isolate

UNIVERSITY OF OXFORD

Public Health England

Candida auris



Haploid Eukaryotic organism of ~12 Mb
First reported in in Japan 2009
Resistant to azoles

Independent emergences structured by geography

What is selecting for these emergences?

UNIVERSITY OF OXFORD

Public Health England

Future

- We don’t have an operational solution for routine practice

- Key challenges:

  - Sample preparation
  - Speed, cost and accuracy
  - Software for processing, analysis and reporting sequence data
  - Building, updating, accredited knowledge bases (e.g. resistance determinants)
  - How to scale pairwise difference analyses within current computational resources
  - Achieving and storing “open data” for analysis

Acknowledgements

- Sarah Walker
- Rosalind Harding
- Tim Peto
- Neil Woodford
- Mark Wilcox – Leeds
- Grace Smith – Birmingham
- Philip Monk - Leicester
- John Paul – Brighton
- Martin Llewellyn – Brighton
- Research Fellows (6)
- Tim Davies
- Amy Mathers - Uva, USA

Microbiology, DNA preparation

- Dai Griffiths
- Kate Dingle
- Nicole Stoesser
- Alison Vaughan
- Bernadette Young
- Claire Gordon

International

Oxford High Throughput Sequencing Hub team

People participating in the studies

Informatics

- David Wyllie
- John Finney
- Milind Achyria
- Laura Madrid
- Infections in Oxfordshire Research Database Team

Bioinformatics and Population Biology

- Danny Wilson
- Carlos del Ojo Elias
- Saheer Gharbia
- Tanya Golubchik
- Anna Sheppard
- Dilrini de Silva
- Xavier Didelot
- Jess Hedge

UNIVERSITY OF OXFORD

Public Health England