



Public Health
England

Applications of Whole Genome Sequencing in Food Safety Management

Tim Dallman  @timdallman

Gastrointestinal Bacteria Reference Unit
Colindale, London



Who is PHE?

Role of the National Reference Laboratory

- Provide gold standard microbial typing of isolates
- Monitor for emerging threats to public health
- Support national surveillance activities
- Facilitate national outbreak detection and investigation
- Contribute to international infectious disease surveillance and response to outbreak investigation across borders





Public Health
England

Gastrointestinal Bacteria Reference Unit Provide ID & Typing for:

Salmonella (10,000 isolates / yr)

E. coli & *Shigella* (3000 isolates / yr)

Listeria monocytogenes (1000 isolates / yr)

Campylobacter (2000 isolates / yr)



Public Health Questions

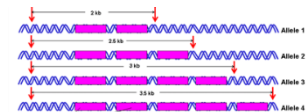
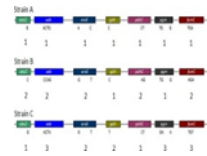
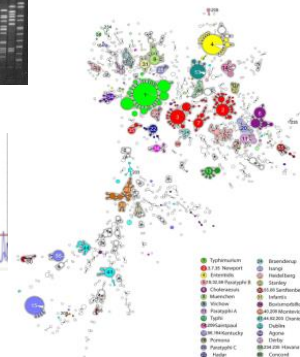
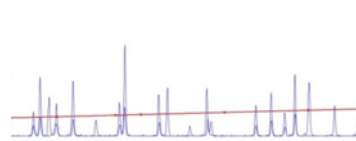
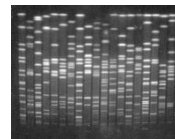
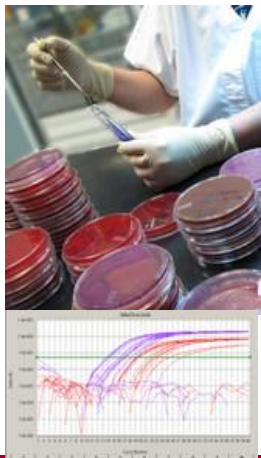
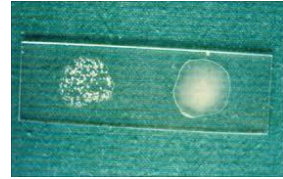
- What species/clone is it? e.g. O157, O26, S. Typhi, S. Enteritidis
- Have we seen this strain before?
- How pathogenic / resistant is this strain to humans?



Conventional ID and typing methods

- Culture-based methods
- PCR specific marker genes
- Biochemistry
- Phenotypic arrays

- Serotyping
- Phage typing
- AMR
- Molecular methods:
PFGE, MLVA, MLST





Move to WGS

- WGS provides a single one step method for identification and typing and provides a wealth of additional information
- Faster TATs, lower costs
- Quality, standardisation, potential for rapid global comparability
- Safety



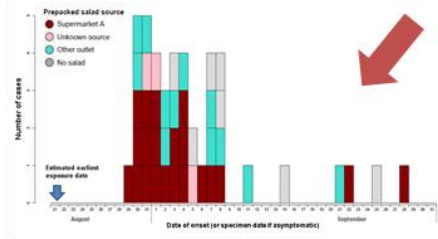
- Improved resolution for strain discrimination
- Able to provide phylogenetic information
- Improved cluster detection
- Provides information on potential sources
- Ability to rapidly screen large number of isolates for virulence genes, AMR genes
- Leads to improved understanding of GI pathogens and OBs



Overview

Cluster 15:2378
1 isolates within search criteria, 3 isolates in the database

SNP Address	Molts ID	Ext ID	Sending Lab	PHE Region	PHE Centre	HPT	Country	Sample Type	Isolate Info	NHS Number	First Name	Surname	Date of Birth	Post Code	Receive Date	Sample Date	Report Date	Foreign Travel	Organism Identified
2.154.315.2041.2320.2378.3088	H160280283	16.0900451	DX5730100-HPA WEST MIDLANDS	Midlands and East of England	West Midlands	West Midlands East	England	Human Faeces							2016-01-15	2016-01-07	2016-01-19	Morocco	Escherichia coli
2.154.315.2041.2320.2378.2814	H153980184	15-2077330-1	DX6120200-HPA S.WEST BRISTOL	South of England	South West	Avon, Gloucestershire and Wiltshire	England	Human Faeces							2015-09-25	2015-09-18	2015-10-06	Morocco	Escherichia coli
2.154.315.2041.2320.2378.2546	H152920519	M.15.1798871	WIMSTON HOSPITAL L35	North of England	North West	Cheshire and Merseyside	England	Human Faeces							2015-07-14	2015-07-09	2015-07-16	Morocco	Escherichia coli



Number of SNP differences

16.09.00451	0
15.2077330-1	1
15.2077330-1	2
15.2077330-1	3
15.2077330-1	4
15.2077330-1	5
15.2077330-1	6
15.2077330-1	7
15.2077330-1	8
15.2077330-1	9
15.2077330-1	10
15.2077330-1	11
15.2077330-1	12
15.2077330-1	13
15.2077330-1	14
15.2077330-1	15
15.2077330-1	16
15.2077330-1	17
15.2077330-1	18
15.2077330-1	19
15.2077330-1	20
15.2077330-1	21
15.2077330-1	22
15.2077330-1	23
15.2077330-1	24
15.2077330-1	25
15.2077330-1	26
15.2077330-1	27
15.2077330-1	28
15.2077330-1	29
15.2077330-1	30
15.2077330-1	31
15.2077330-1	32
15.2077330-1	33
15.2077330-1	34
15.2077330-1	35
15.2077330-1	36
15.2077330-1	37
15.2077330-1	38
15.2077330-1	39
15.2077330-1	40
15.2077330-1	41
15.2077330-1	42
15.2077330-1	43
15.2077330-1	44
15.2077330-1	45
15.2077330-1	46
15.2077330-1	47
15.2077330-1	48
15.2077330-1	49
15.2077330-1	50
15.2077330-1	51
15.2077330-1	52
15.2077330-1	53
15.2077330-1	54
15.2077330-1	55
15.2077330-1	56
15.2077330-1	57
15.2077330-1	58
15.2077330-1	59
15.2077330-1	60
15.2077330-1	61
15.2077330-1	62
15.2077330-1	63
15.2077330-1	64
15.2077330-1	65
15.2077330-1	66
15.2077330-1	67
15.2077330-1	68
15.2077330-1	69
15.2077330-1	70
15.2077330-1	71
15.2077330-1	72
15.2077330-1	73
15.2077330-1	74
15.2077330-1	75
15.2077330-1	76
15.2077330-1	77
15.2077330-1	78
15.2077330-1	79
15.2077330-1	80
15.2077330-1	81
15.2077330-1	82
15.2077330-1	83
15.2077330-1	84
15.2077330-1	85
15.2077330-1	86
15.2077330-1	87
15.2077330-1	88
15.2077330-1	89
15.2077330-1	90
15.2077330-1	91
15.2077330-1	92
15.2077330-1	93
15.2077330-1	94
15.2077330-1	95
15.2077330-1	96
15.2077330-1	97
15.2077330-1	98
15.2077330-1	99
15.2077330-1	100



Sequencing
4 - 7 days



Sub-culture and extraction
1 - 2 days



Bioinformatics pipeline
1 - 5 days

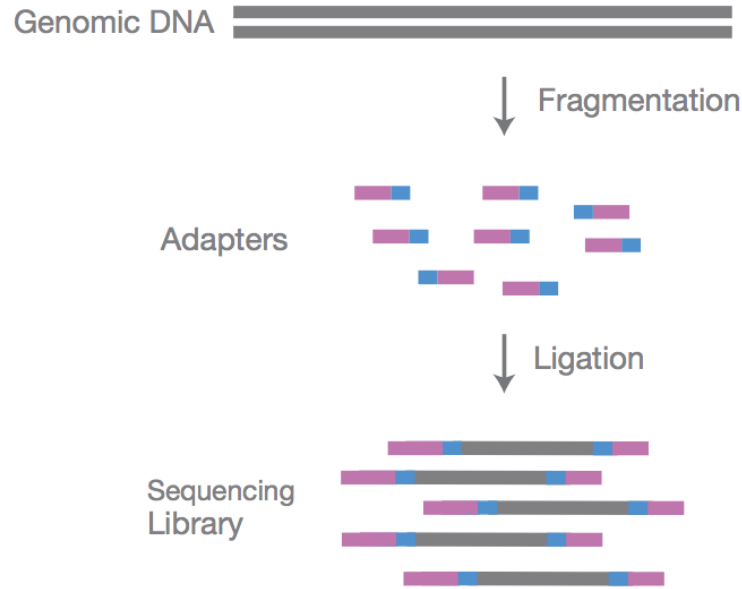


SNP address on GDW



The (very) basics of sequencing

A. Library Preparation

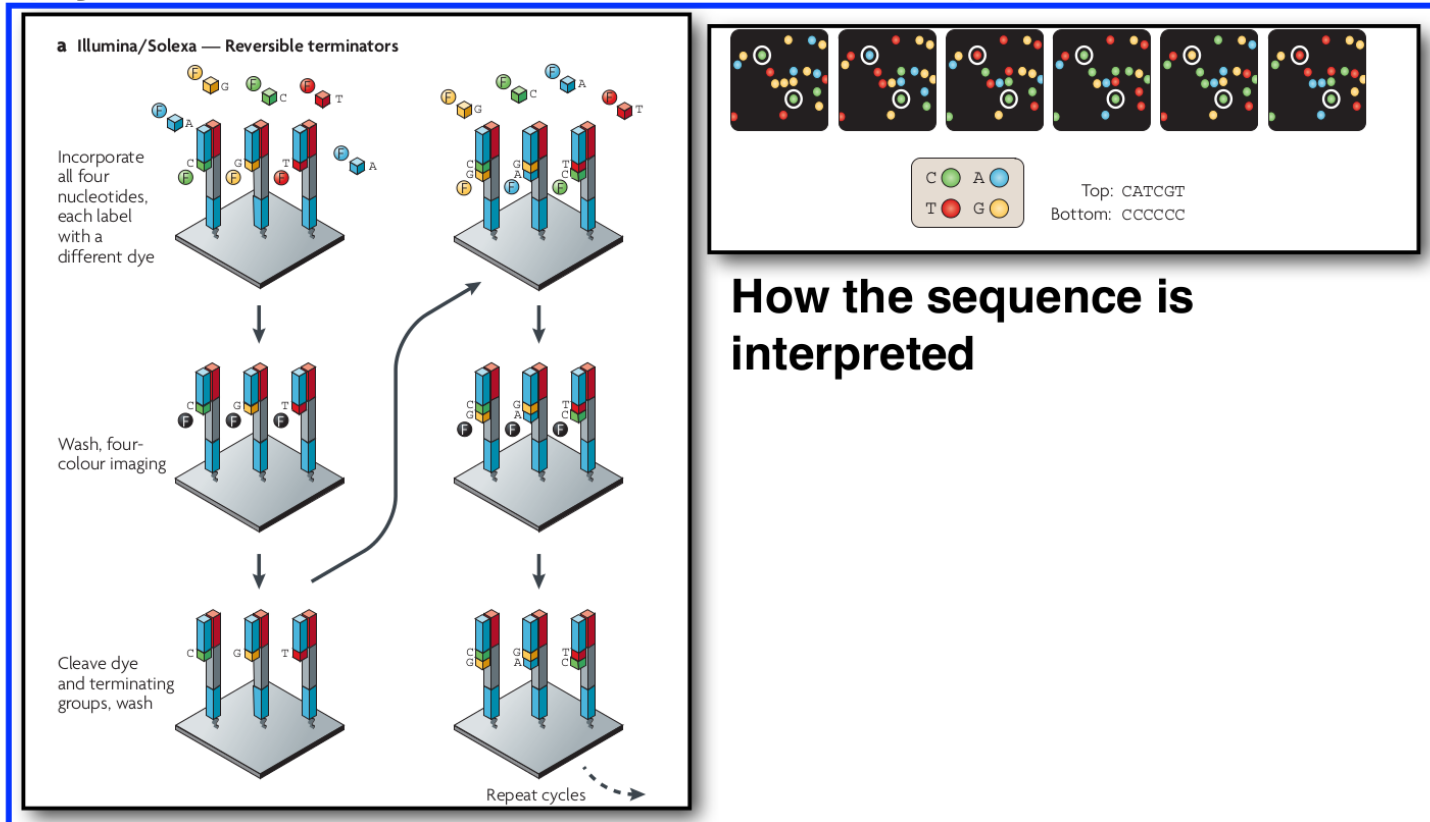


NGS library is prepared by fragmenting a gDNA sample and ligating specialized adapters to both fragment ends.

Illumina Sequencing



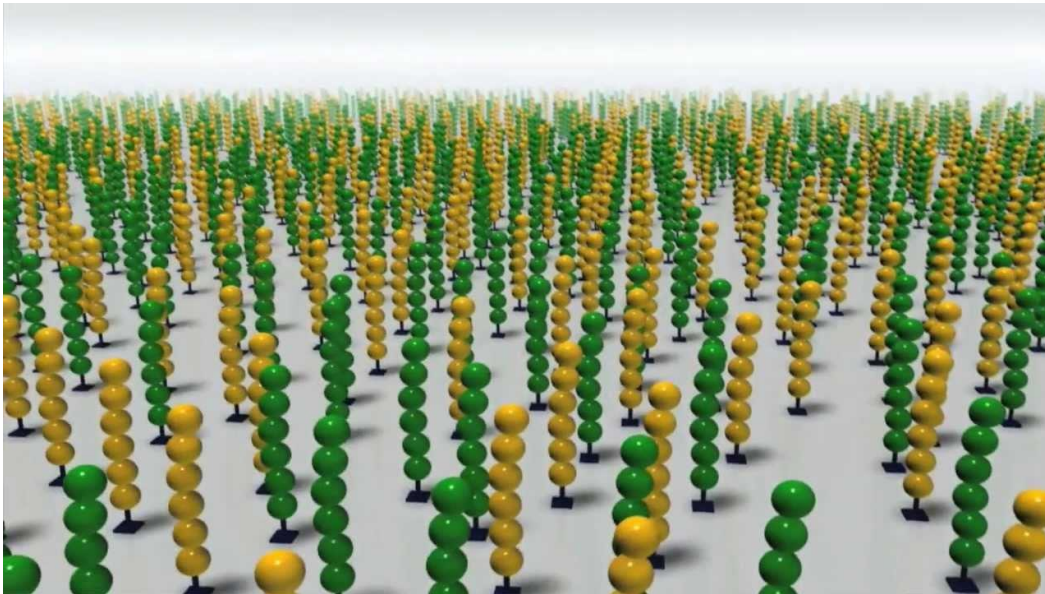
Sequencing and Imaging (Cyclic reversible termination - Illumina)





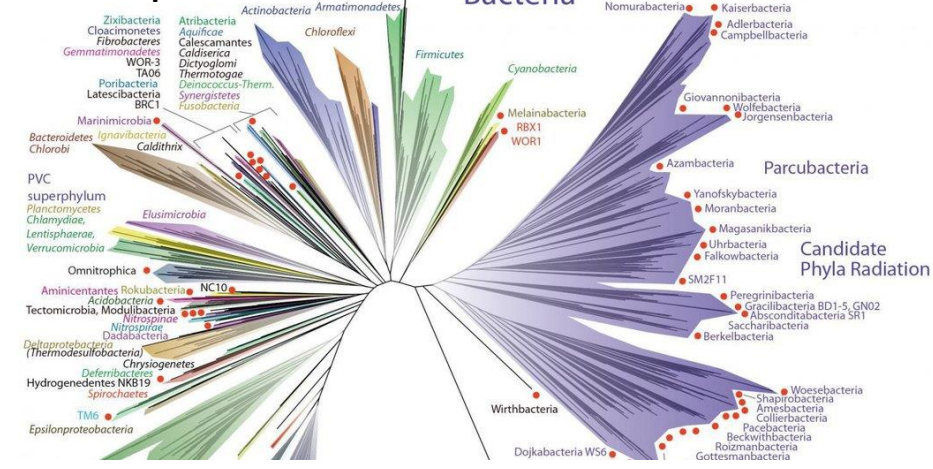
Public Health
England

Illumina – sequencing by synthesis

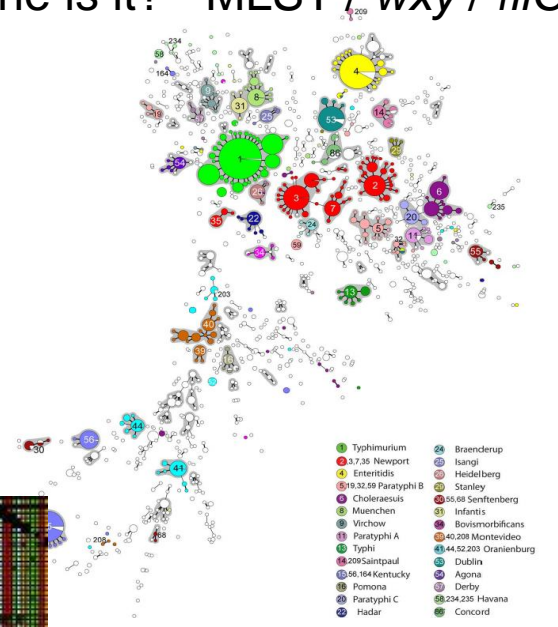


98% Accuracy
100-300 bp read length
1.5 billion reads per flow cell
27 hrs

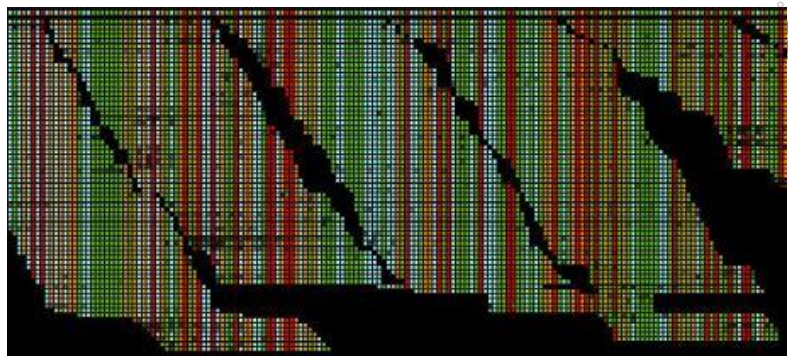
What species is it? - Kmer



What clone is it? - MLST / wxy / fliC

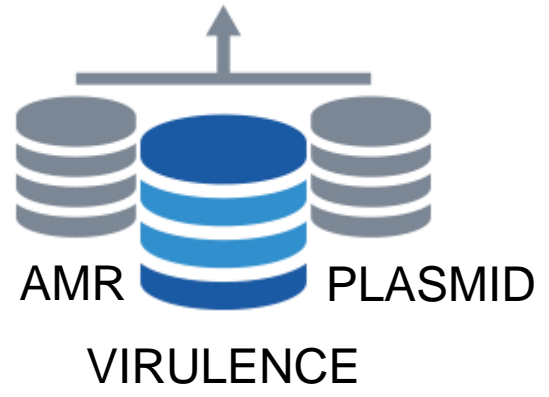


Illumina Short Reads



A	G	T	C	G	C	G	T	A	T	G	T	C	T	G	A	C	C	C	
A	G	T	C	G	C			A	T	G	T	A	T						
			C	G	C	G	T	A				T	A	T	G	A	C	C	C
A	G	T	C	G	C	G						A	T	G	A	C	C	C	
						G	T	A	T	G	T	A							
		T	C	G	C	G	T					G	T	A	T	G	A		
A	G	T	C	G	C	G						A	T	G	A	C			
				G	C	G	T	A	T										
									T	G	T	A	T						
				C	G	T	A	T	G										

How similar / different is it? - SNP

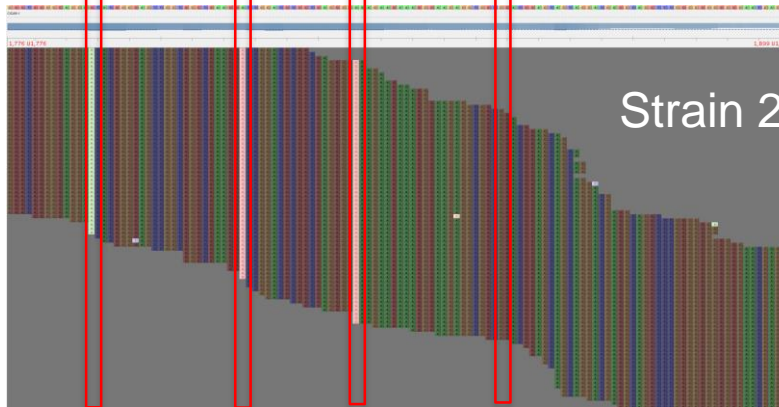
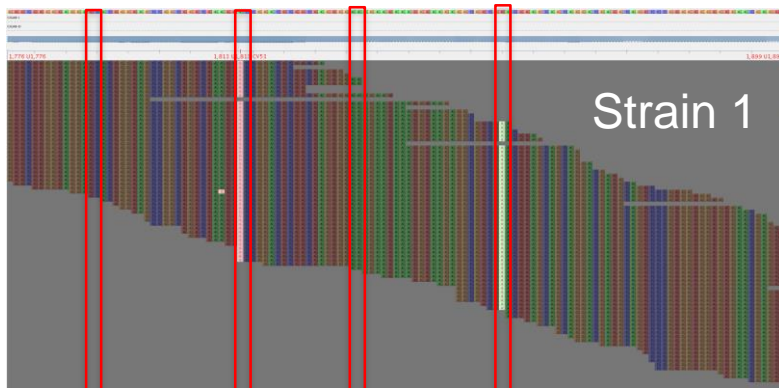


How bad is it?

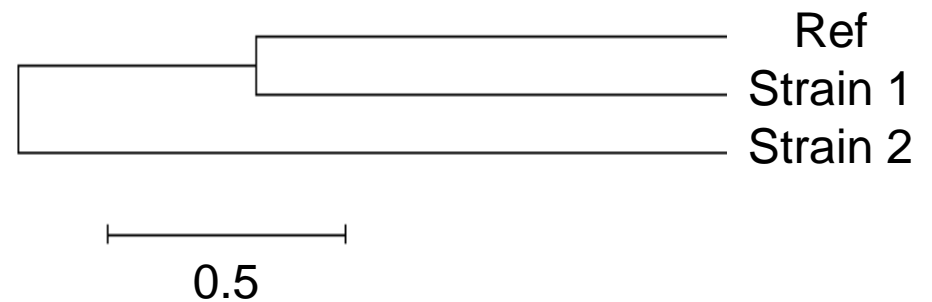


How do we assess similarity between genomes?

Reference



Ref	G	T	A	G
Strain 1	G	G	A	A
Strain 2	A	G	G	G



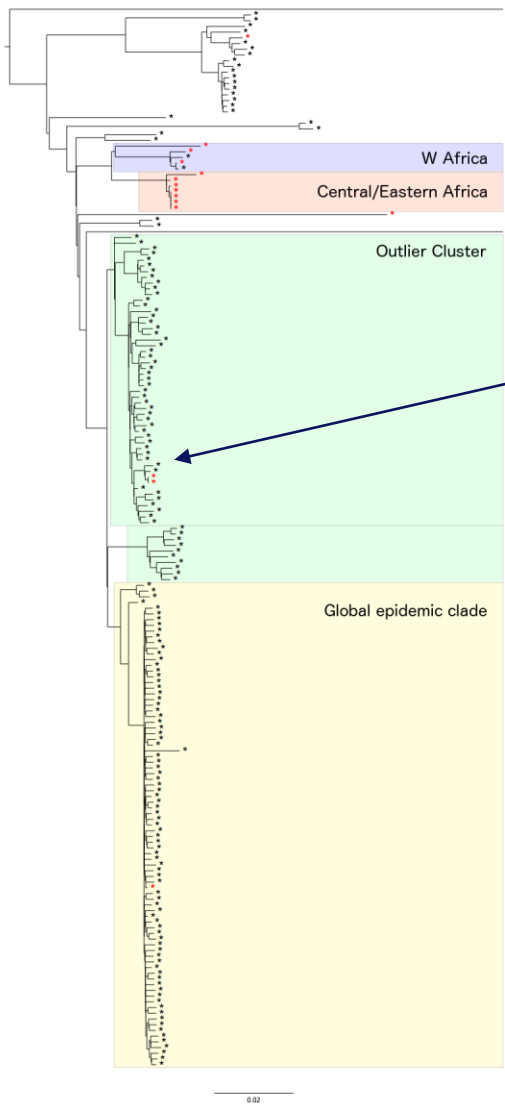


Public Health
England



Salmonella Enteritidis PT14b

‘How Deep Down the Rabbit Hole Do We Really Need to Go?’



N= 9017

Routine Analysis WGS of all *Salmonella* isolates since 2014

Outbreak detected within inpatients at hospital



Public Health
England

Salmonella Enteritidis 14B

Heartlands Hospital
Birmingham

Figure 1

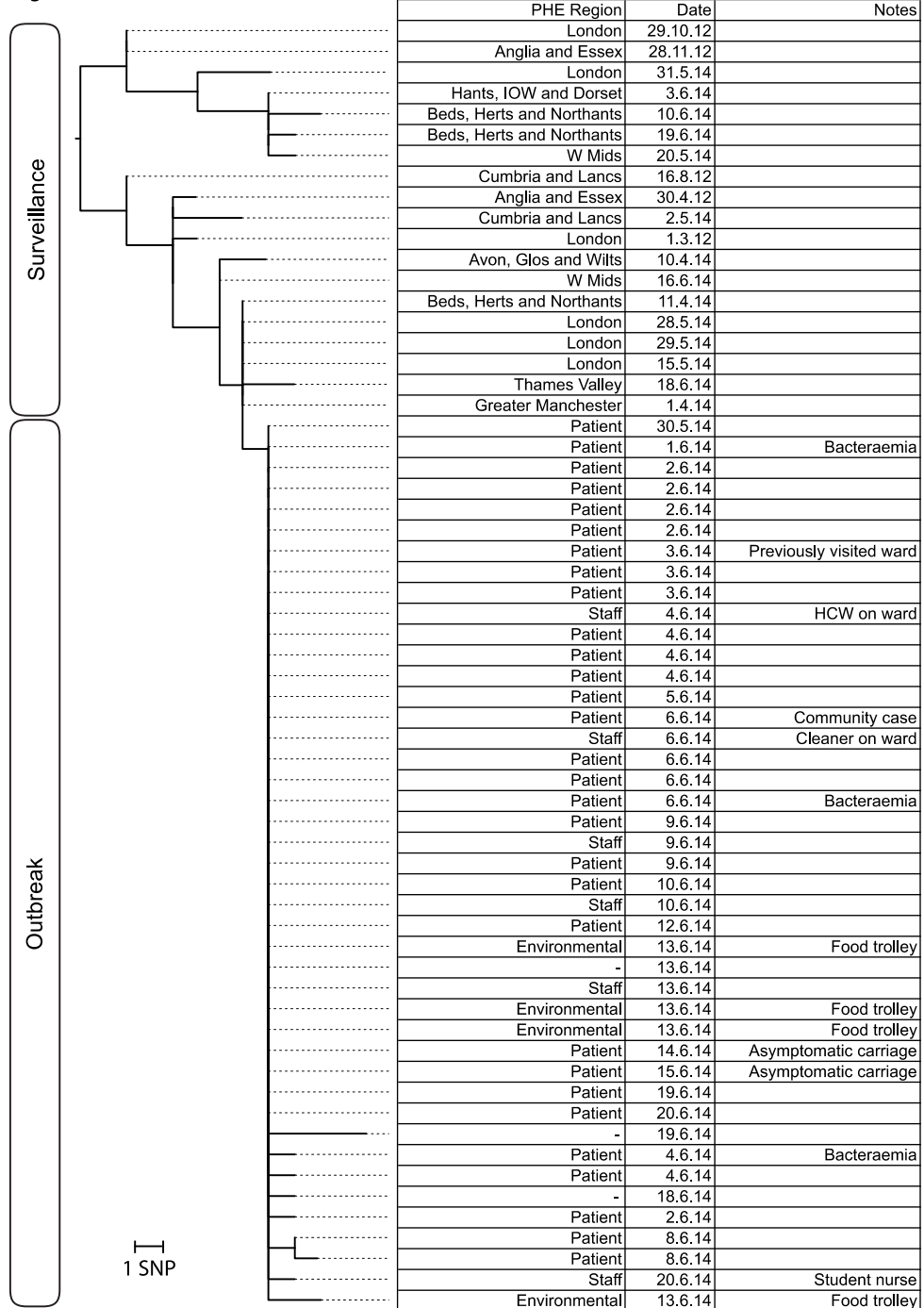
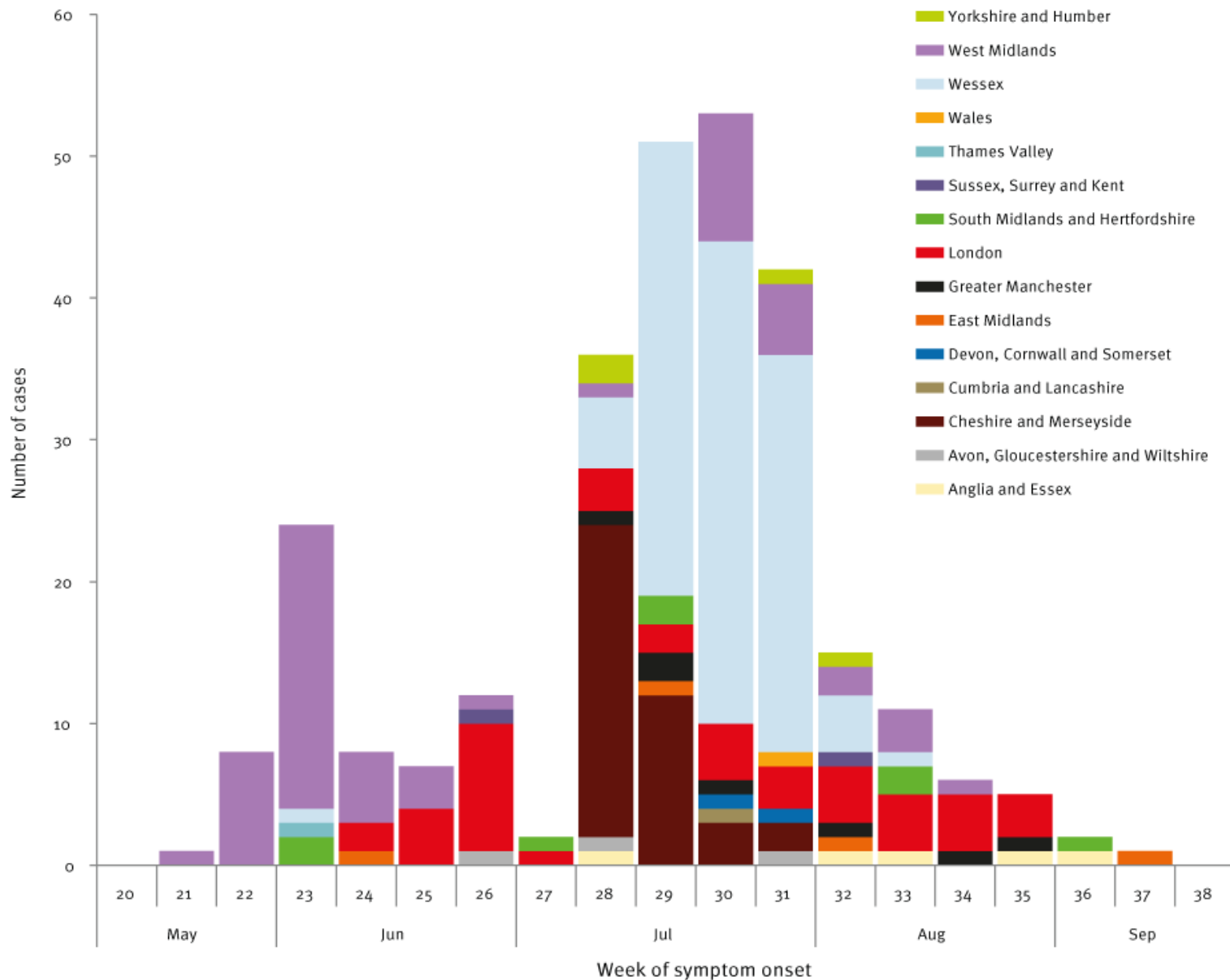


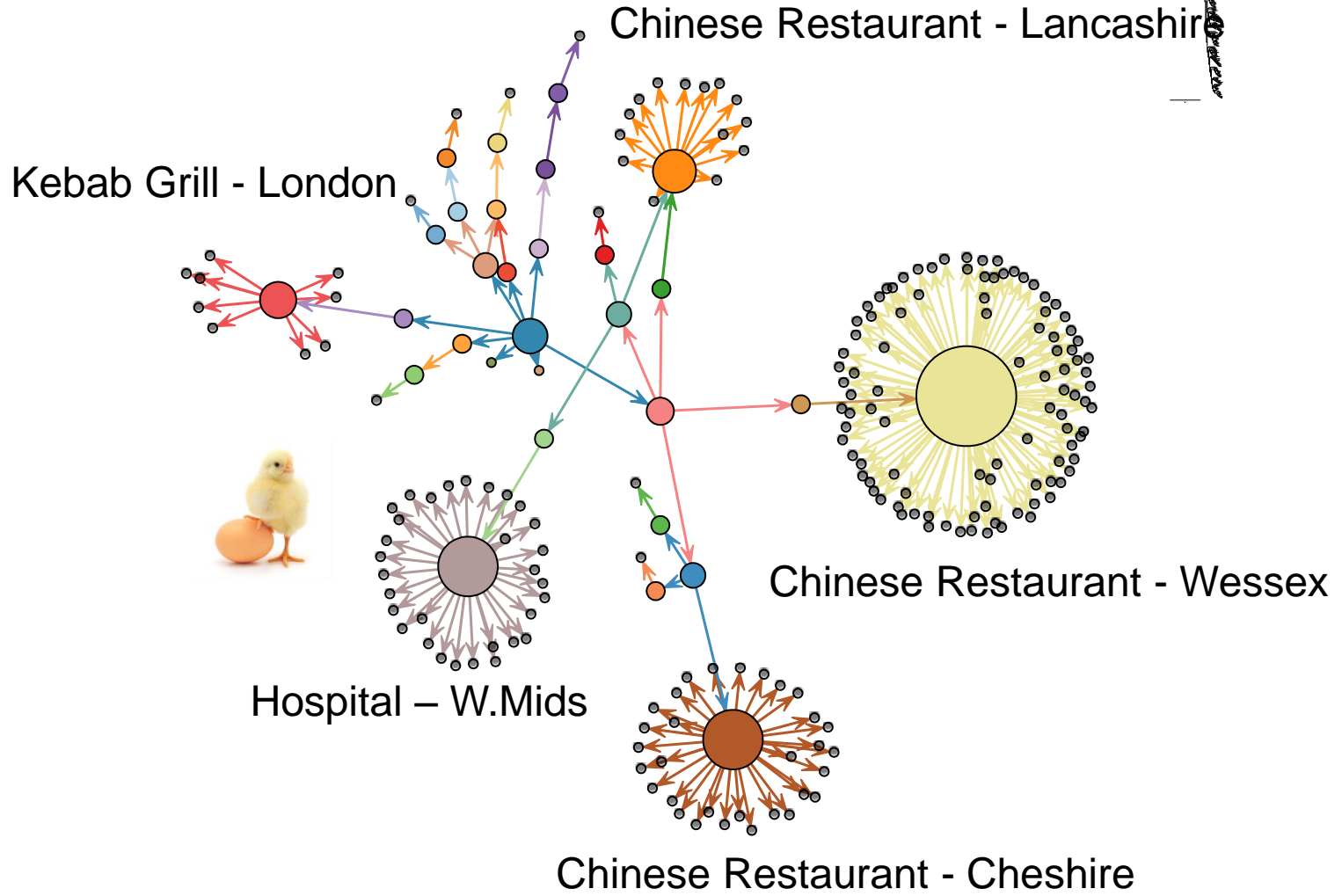
FIGURE 2

Distribution of *Salmonella* Enteritidis phage type 14b, by calendar week of symptom onset and Public Health England centre of residence, England and Wales, week 20–38 2014 (n=284)^a



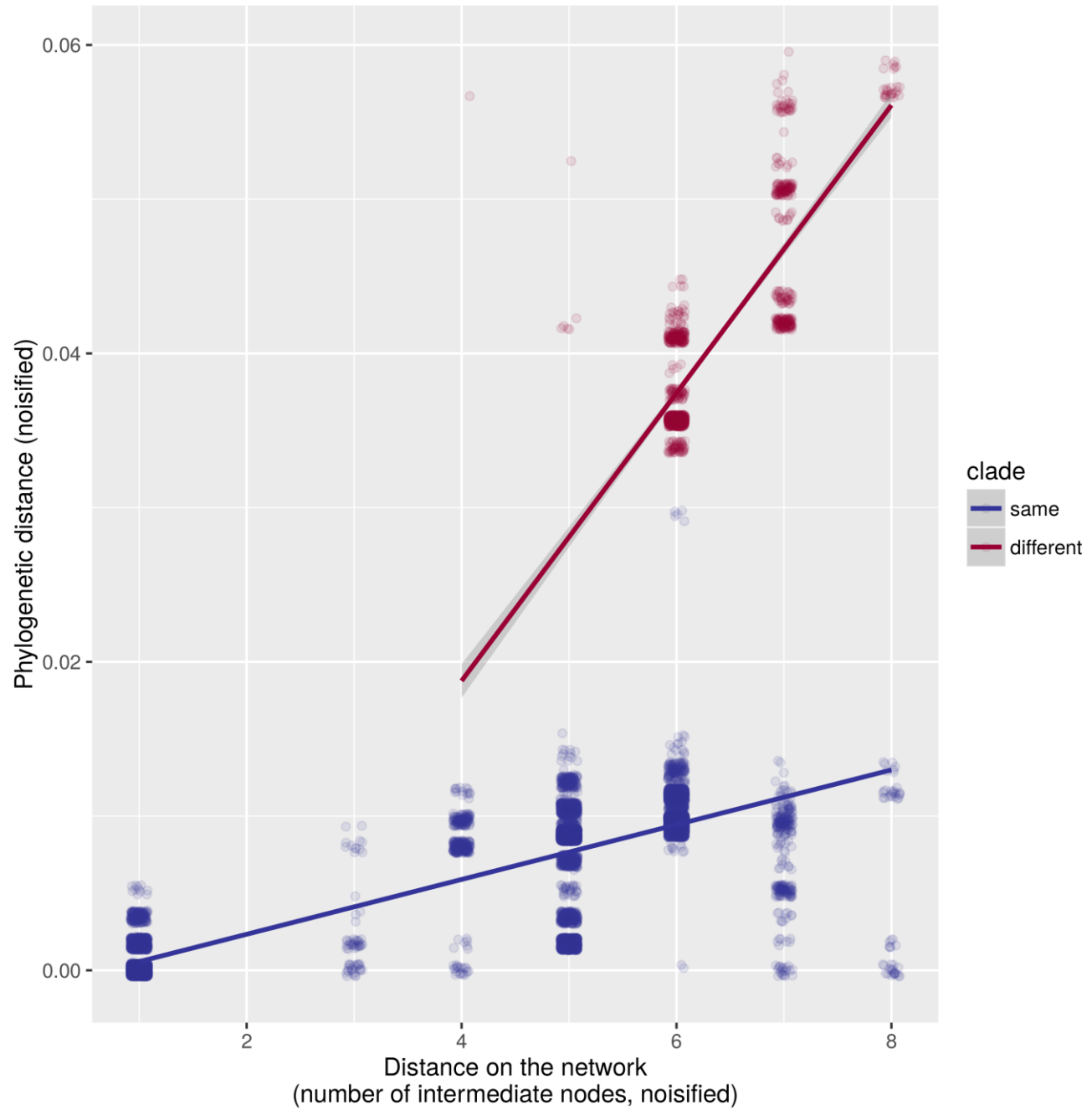
^a Symptom onset dates were not available for three of the 287 outbreak cases.

22 SNPs across outbreak





Public Health
England





Can we explain this association?

- Trace back in UK – All outbreaks involve eggs from same German supplier
- Cases in Austria, Germany, France and Luxembourg
- Trace back investigations lead to sampling of egg production premises in Germany
- Company has 4 separate egg production plants supplied with pullet flocks from single supplier
- European isolates sequenced – cases / eggs



Public Health
England

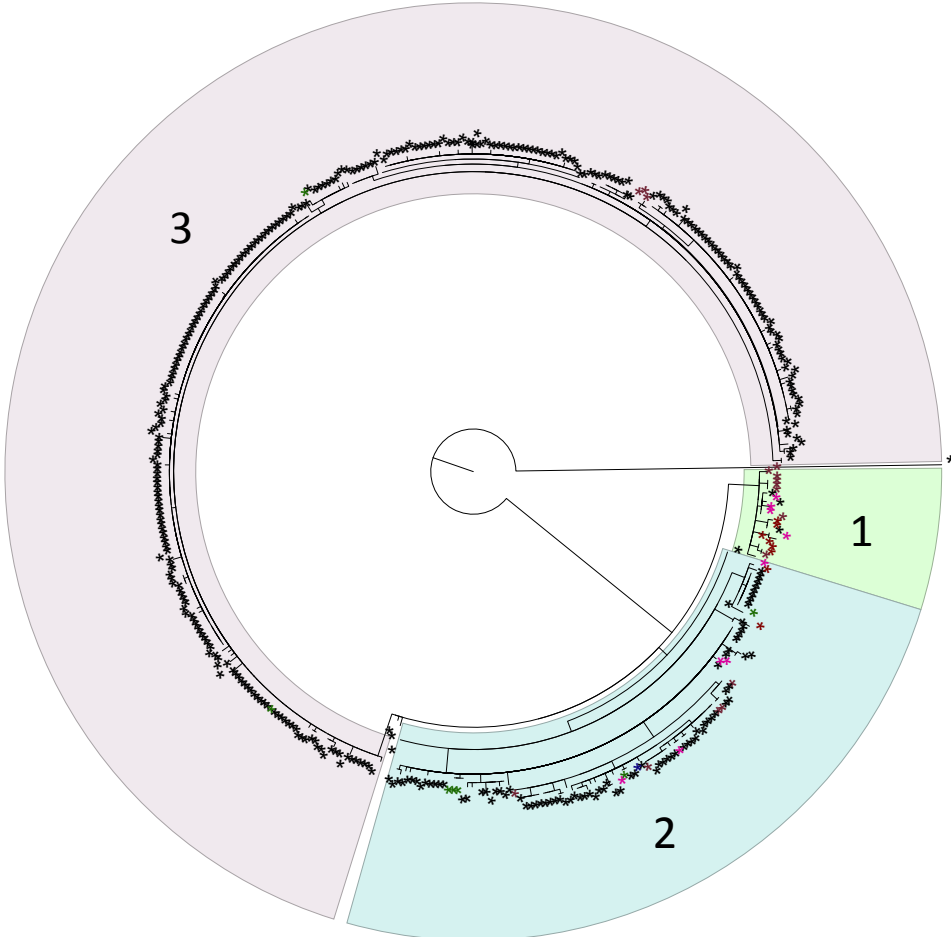
3 years



Timed phylogeny

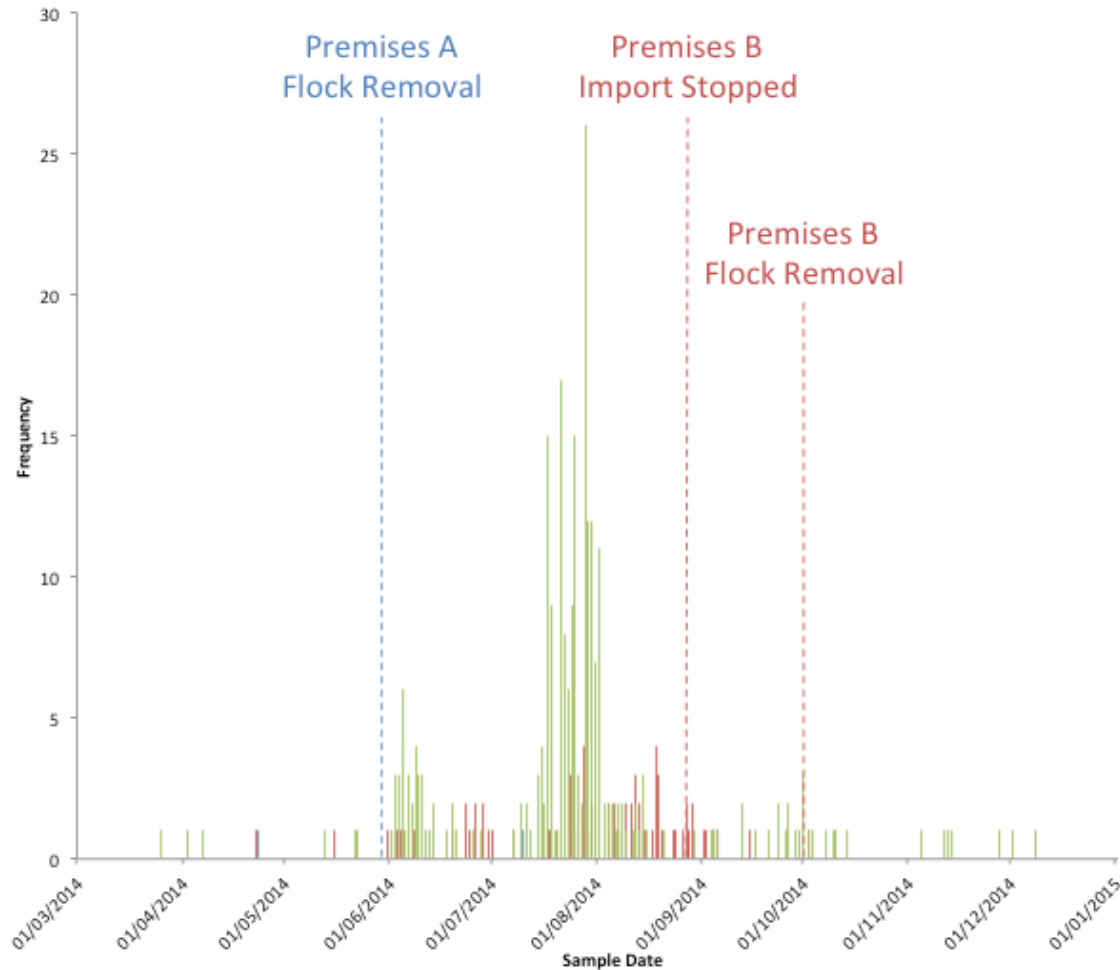
Changes during the outbreak?

Sampling from a diversified source?



Clade 1 – 12 egg isolates + traceback (all site A)
Clade 2 – 2 egg isolates + traceback (all site B)
Clade 3 – 0 egg isolates, 0 traceback to site A or B

Clade 1 – 12 egg isolates + traceback (site A)
Clade 2 – 2 egg isolates + traceback (site B)
Clade 3 – 0 egg isolates, 0 traceback (site A or B)





Summary

- WGS for routine surveillance of infectious disease is a reality.
- WGS allows identification of clusters with unparalleled specificity
- WGS provides framework for risk assessment / attribution.
- WGS allows the underlying epidemiology within an outbreak to be revealed.



Acknowledgments

Enteritidis 14b Outbreak Control Team - Paul Cleary, Tom Inns and International collaborators

Staff in Gastrointestinal Bacteria Reference Unit, Genome Sequencing and Development Unit & Infectious Disease Informatics.

Questions?

