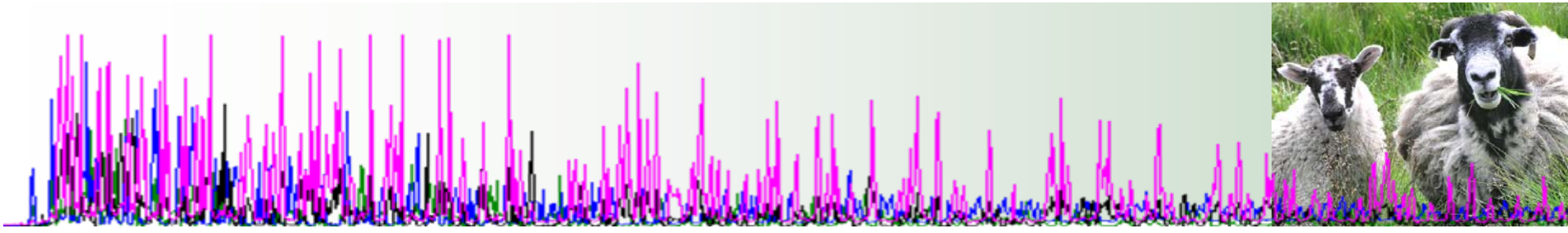
Tracing FMDV in UK outbreaks

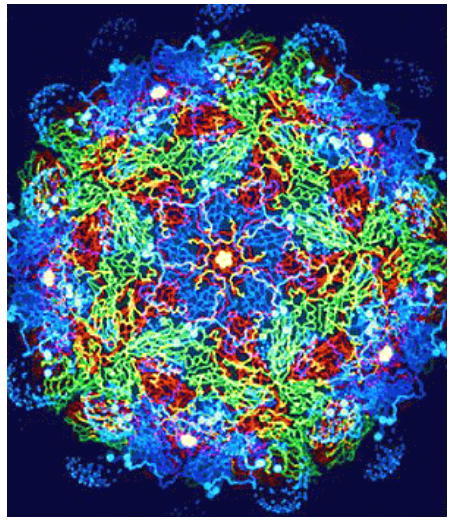
The rationale behind using complete genome sequencing of Foot-and-Mouth Disease Virus as an epidemiological modelling tool

9 May 2008

E.Cottam@uea.ac.uk



FMDV evolution



- Rapid replication rate
- Large population sizes
- High error rate of viral polymerase in the order of 10^{-3} to 10^{-5} misincorporations per nucleotide copied

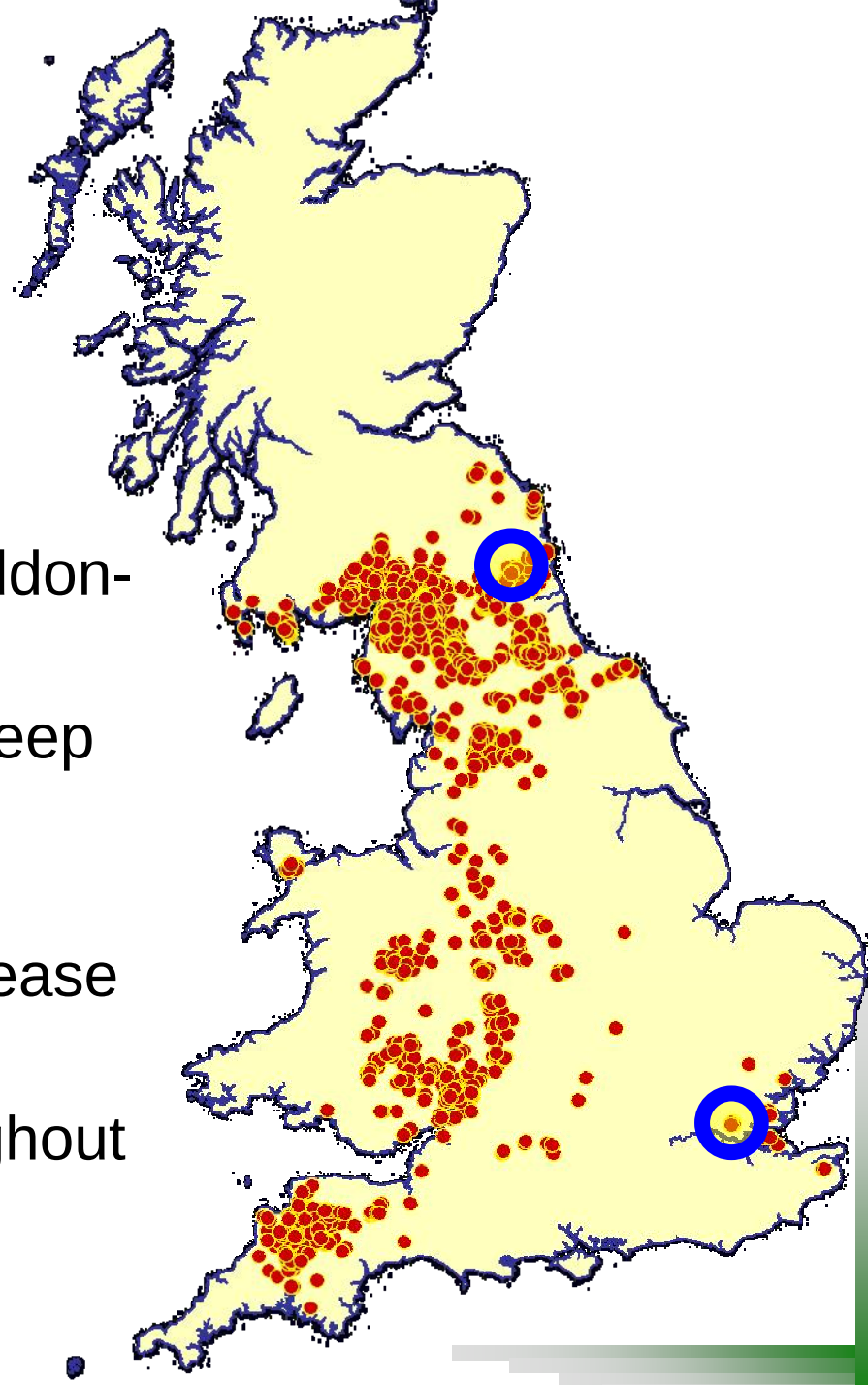
Fast adaptation and evolution

UK 2001 Outbreak

- 2030 infected premises culled
- 8131 culled as dangerous contacts
- > 6.5 million animals culled
- economic cost exceeding £6 billion

UK 2001 Outbreak

- Pan Asia O serotype
- Identified 20th Feb 2001 in Cheale's abattoir Essex
- Source farm identified as Heddon-on-the-Wall.
- Airborne spread to nearby sheep farm.
- Movement of sheep through markets one week before disease identified
- Subsequent outbreaks throughout the UK



Previous work

- The transmission of FMDV is poorly characterised, and information on the direction of farm to farm spread would greatly enhance our understanding and hence control.

- Broad scale genetic tracing of FMDV using VP1 sequences (633nt) between countries

- Attempt in Denmark at intra-epidemic tracing of FMDV using VP1 sequences only*

- Forensic genetic tracing for HIV, Hep C, SARS, Rhinovirus, Norovirus…….

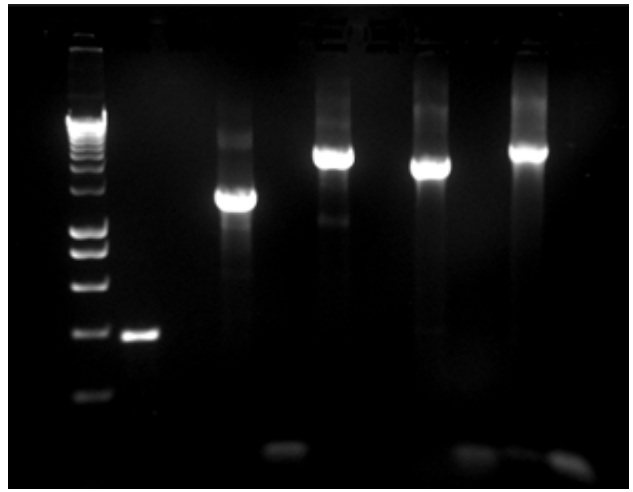
***Christensen, L.S.,** P. Normann, S. Thykier-Nielsen, J. H. Sorensen, K. de Stricker, and S. Rosenorn. **2005**. *Analysis of the epidemiological dynamics during the 1982-1983 epidemic of foot-and-mouth disease in Denmark based on molecular high-resolution strain identification*. **J. Gen. Virol. 86: 2577-2584**

Project Introduction

- Development of method for sequencing complete genome
- Investigation of FMDV evolution during the UK 2001 Outbreak
- Use of complete genome sequence data for forensic genetics of FMDV transmission

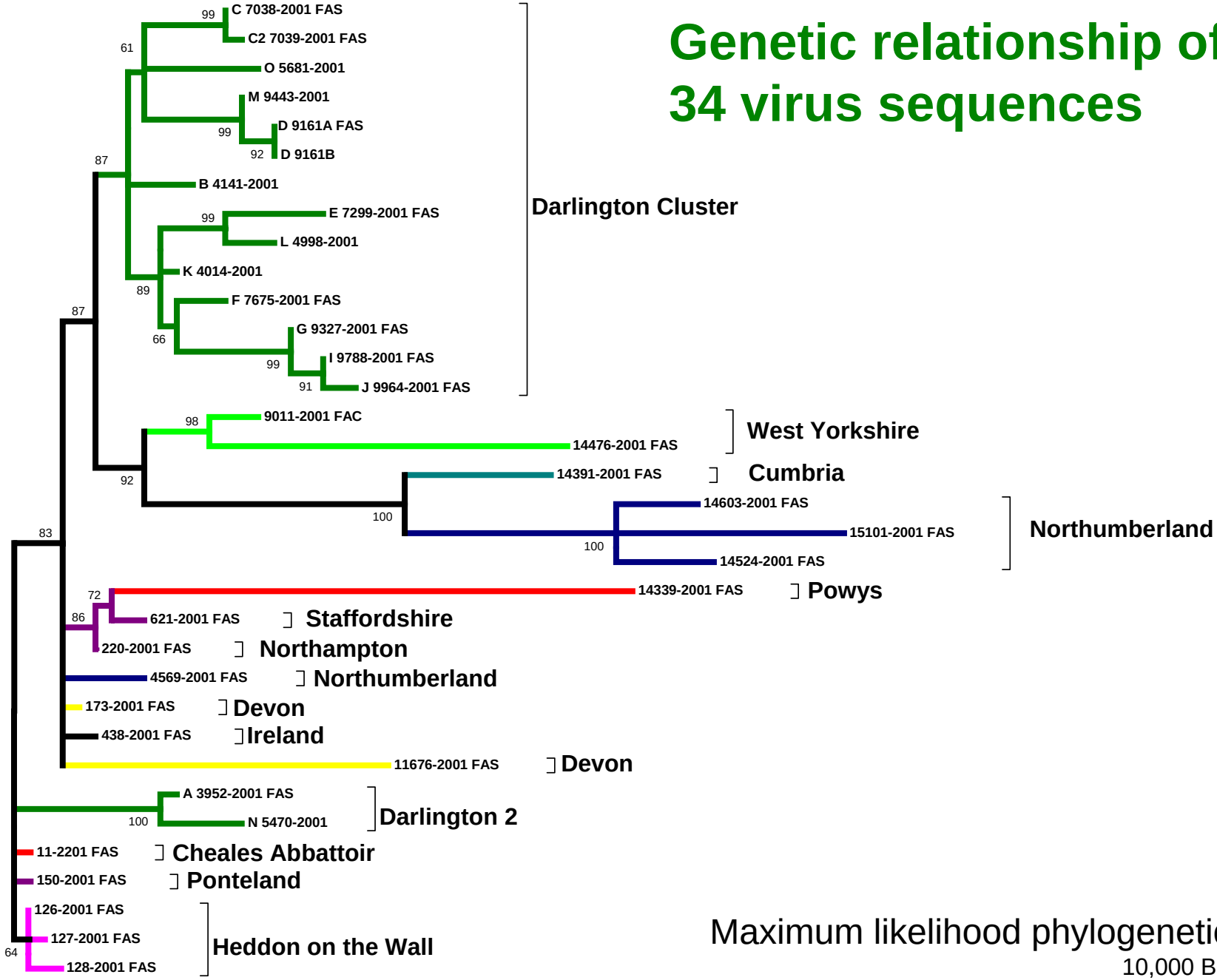
Complete genome sequencing

- Direct from epithelium
- RT-PCR
  - 5 overlapping fragments
  - 84 sequencing reactions
- Use of QIAGEN robot
- Complete sequence within 3 days



Cottam *et al.* 2006 J Virol 80(22), 11274-82

Genetic relationship of 34 virus sequences



Maximum likelihood phylogenetic tree

10,000 Bootstrap

Observed Variation

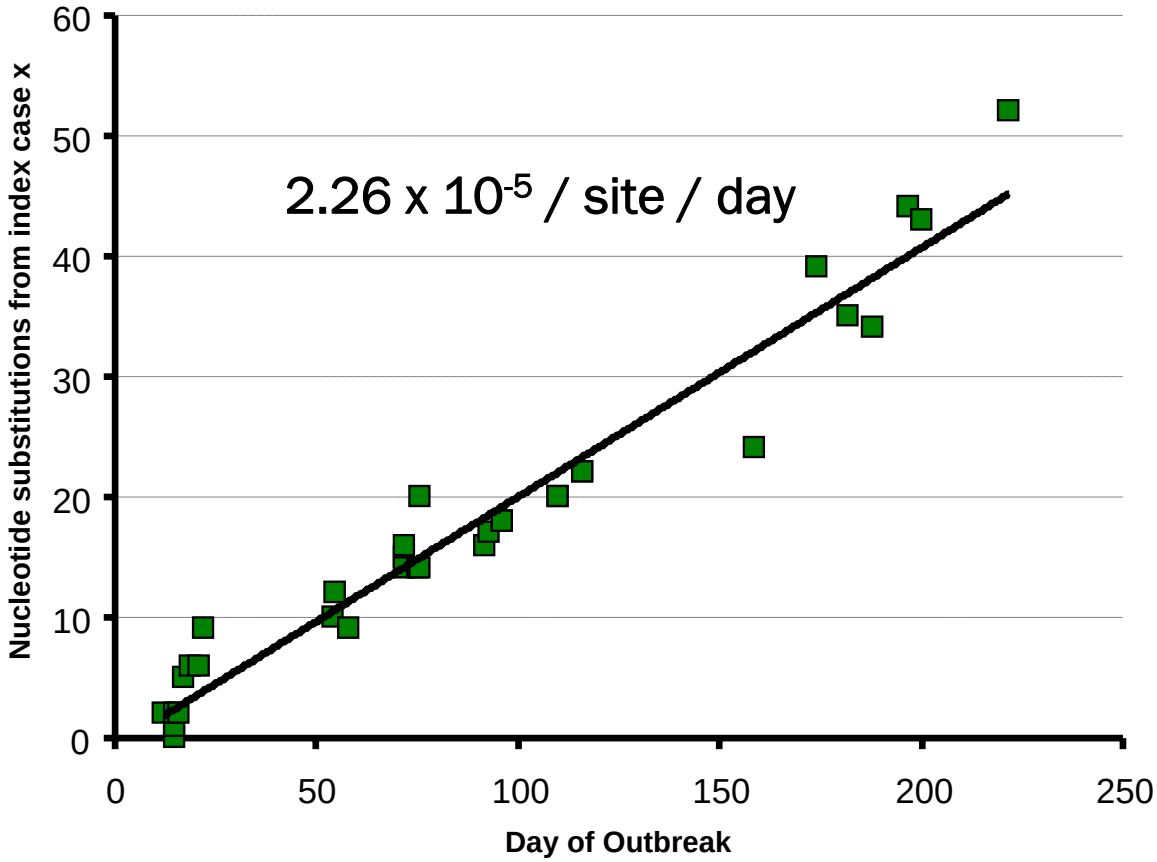
- 52 nucleotide substitutions between source and final case
  - 8 were amino acid changes

- 83 variant nucleotides between Wales and Northumberland

Cottam *et al* 2006 J Virol 80(22), 11274-82

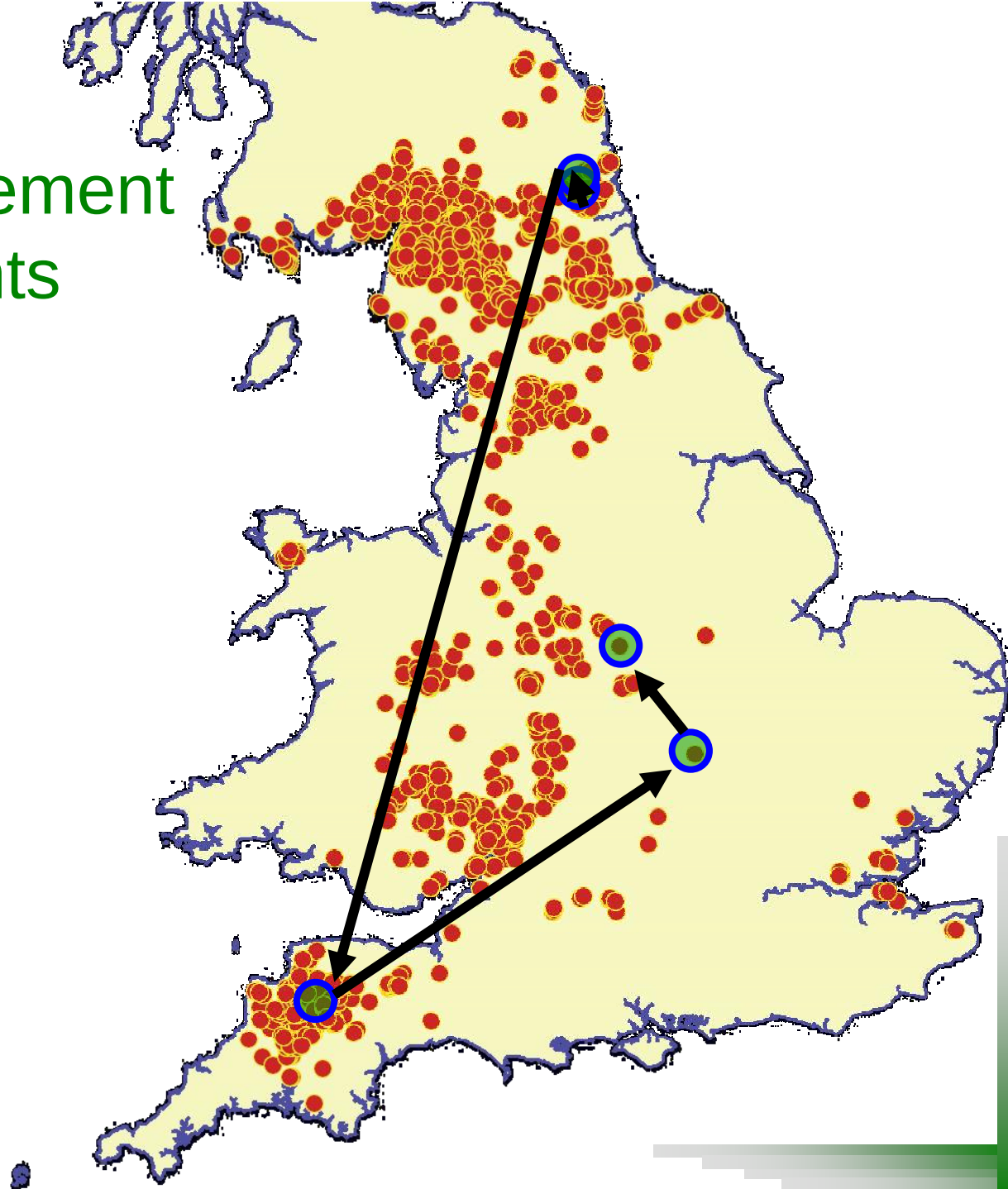
Molecular clock

Nucleotide changes accrue linearly with time and are inherited

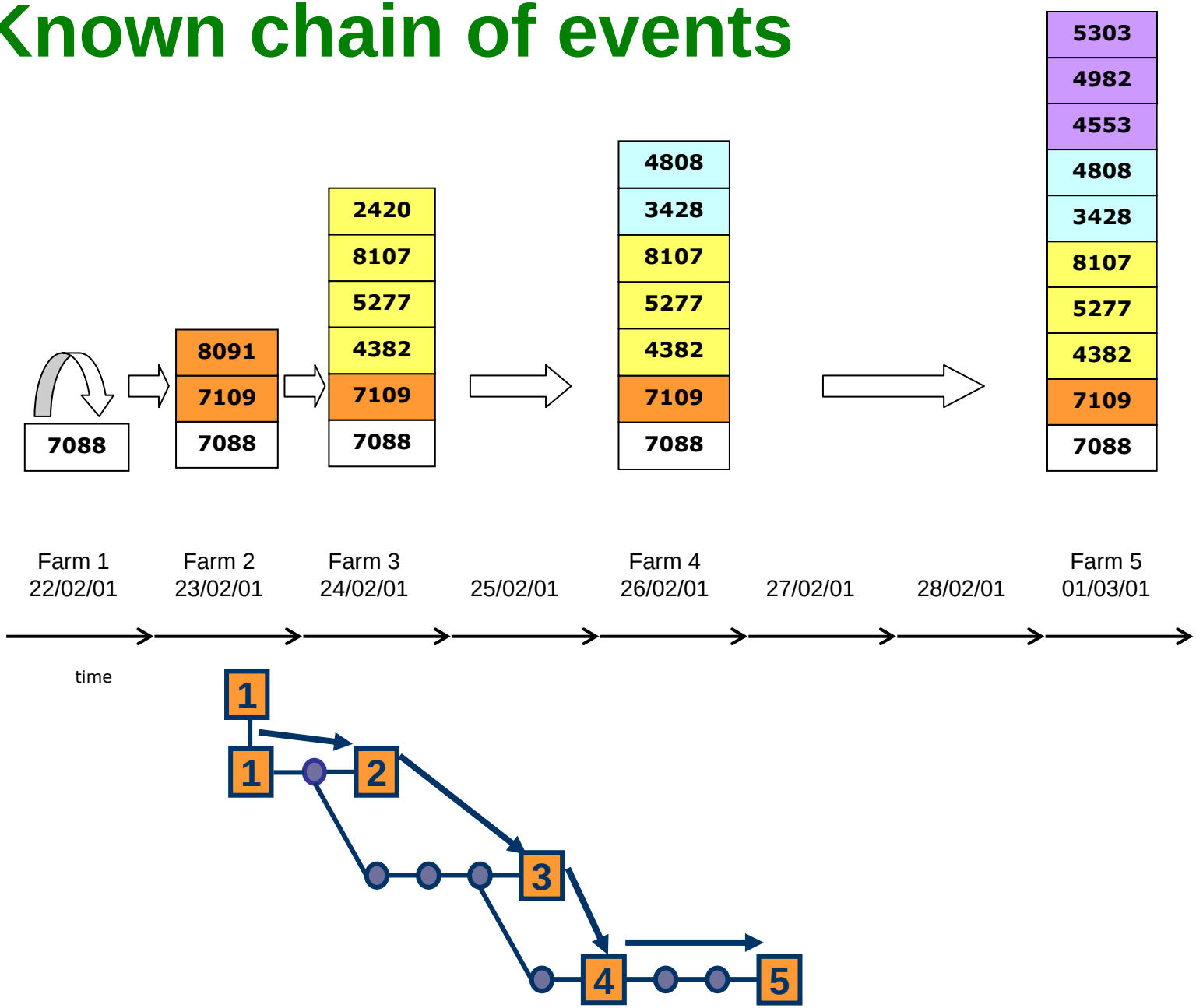


Cottam *et al* 2006 J Virol 80(22), 11274-82

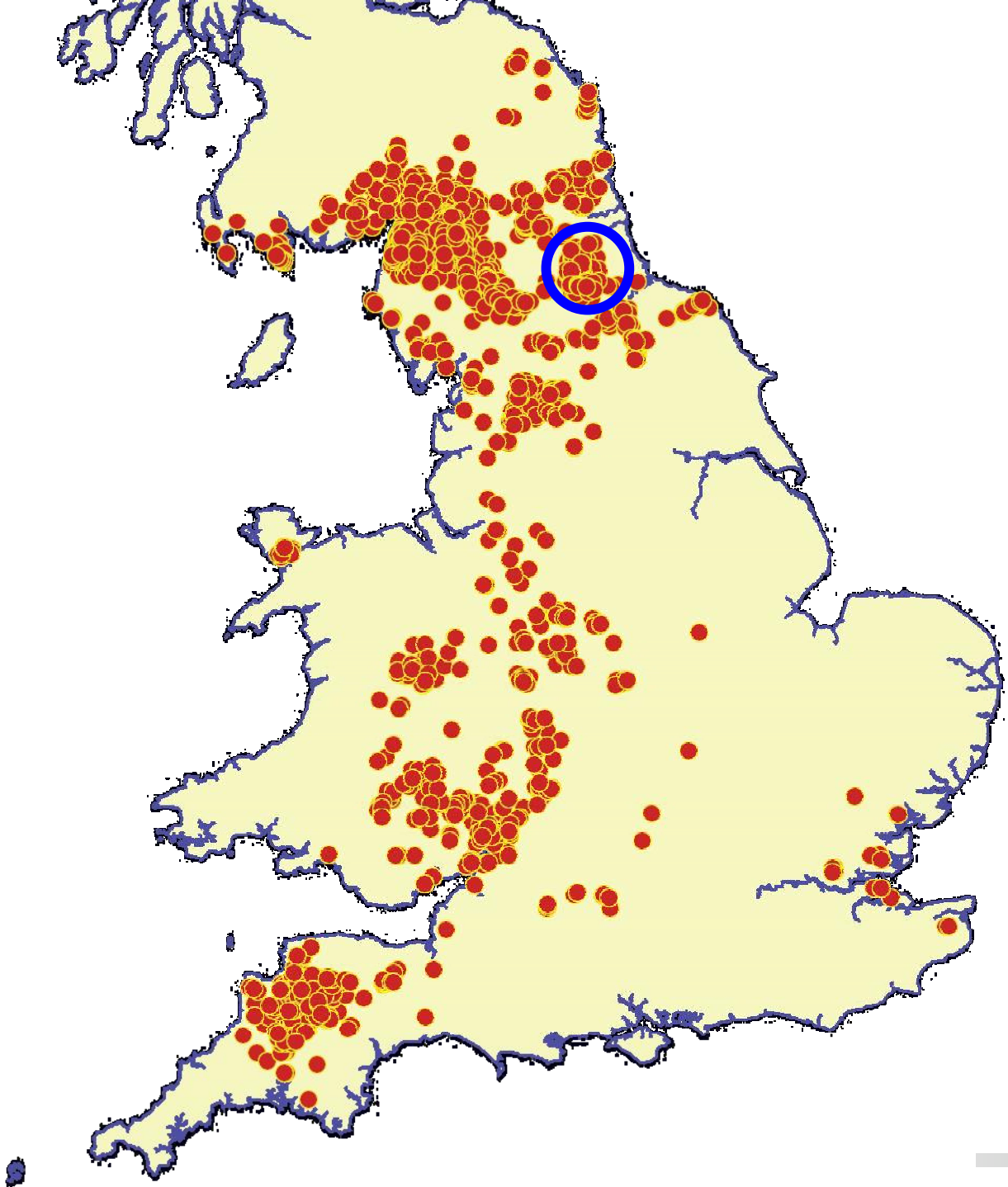
Early animal movement transmission events between 5 farms.



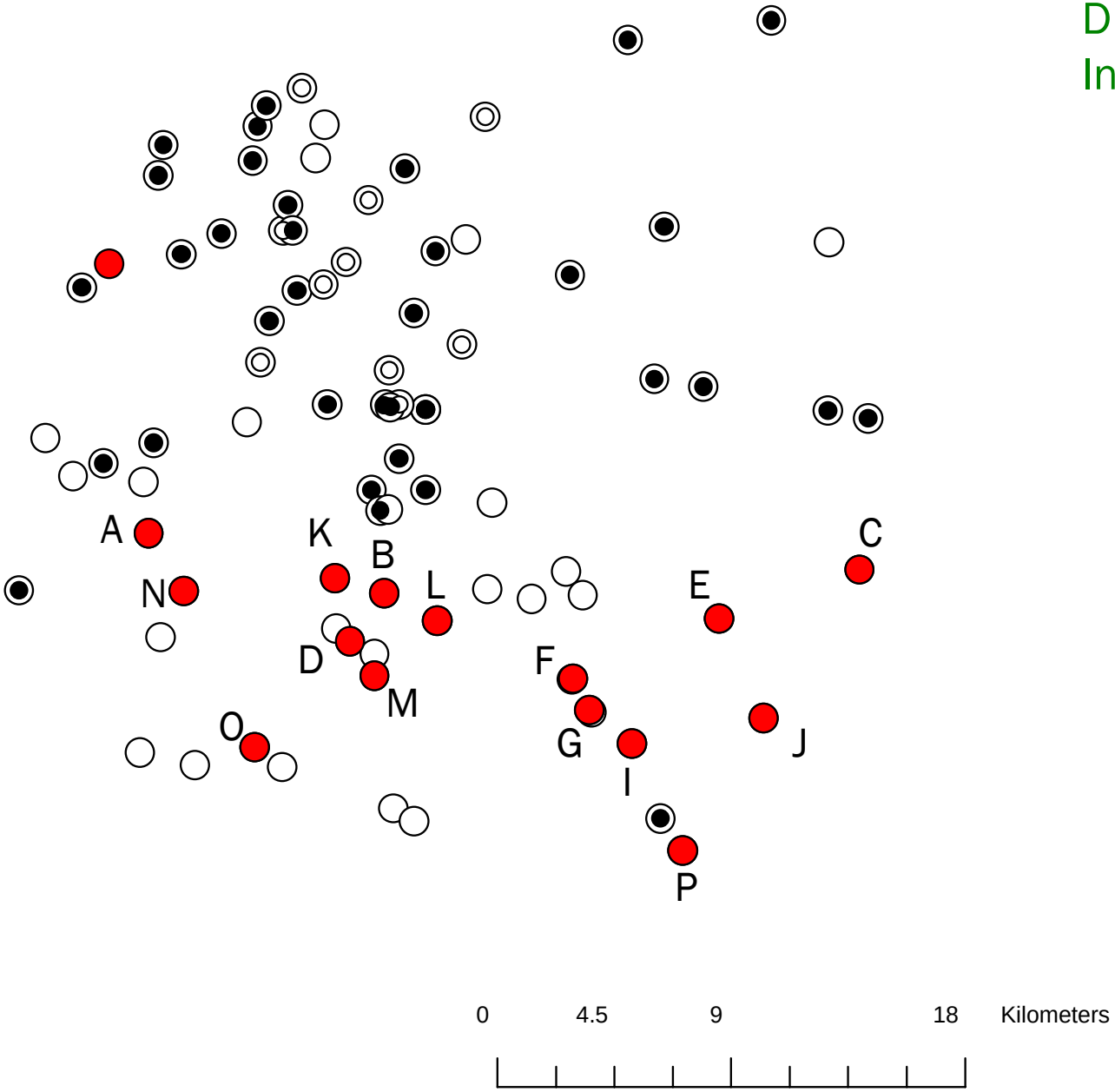
Known chain of events



Cottam *et al* 2006 J Virol 80(22), 11274-82

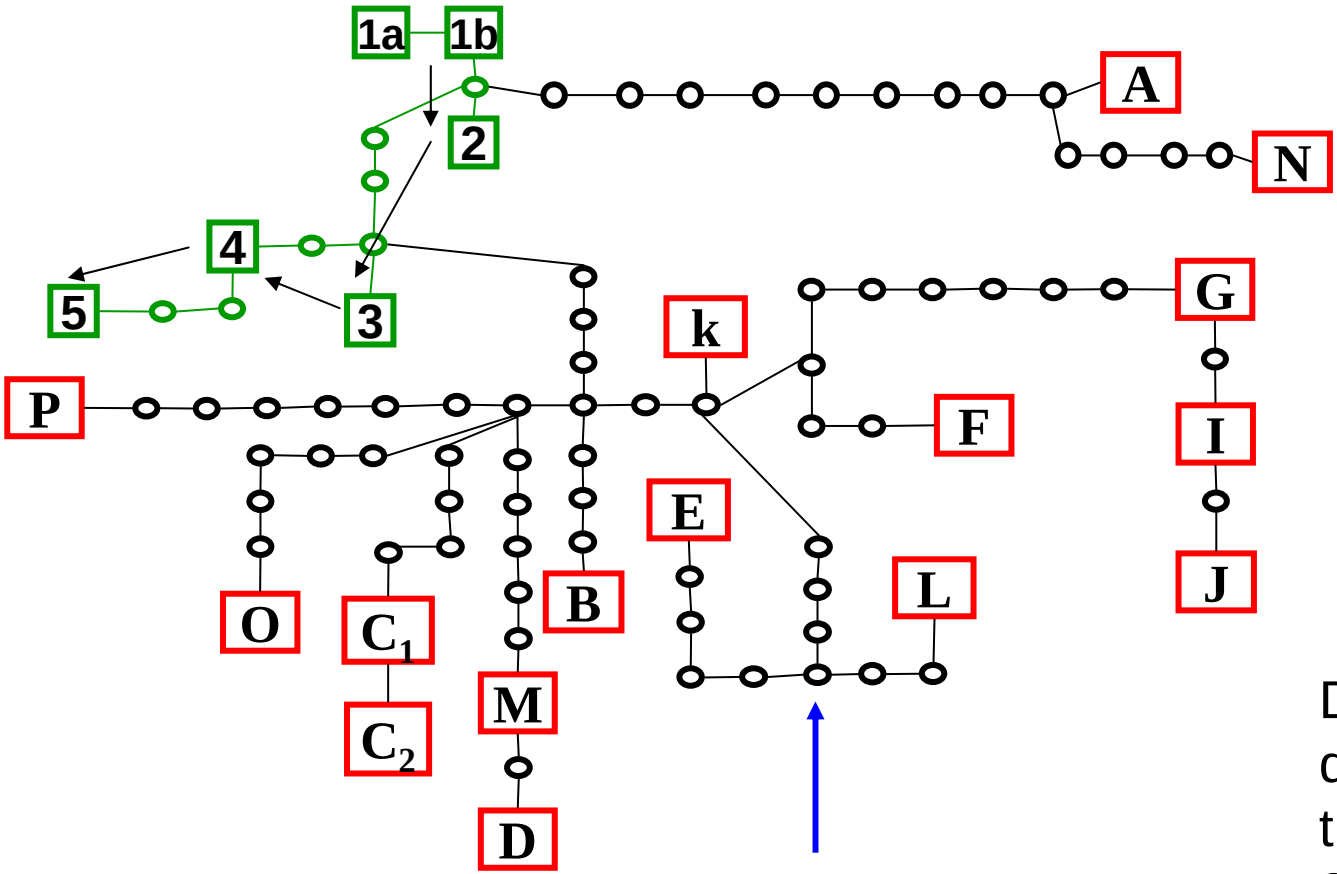


Durham cluster of Infected Premises



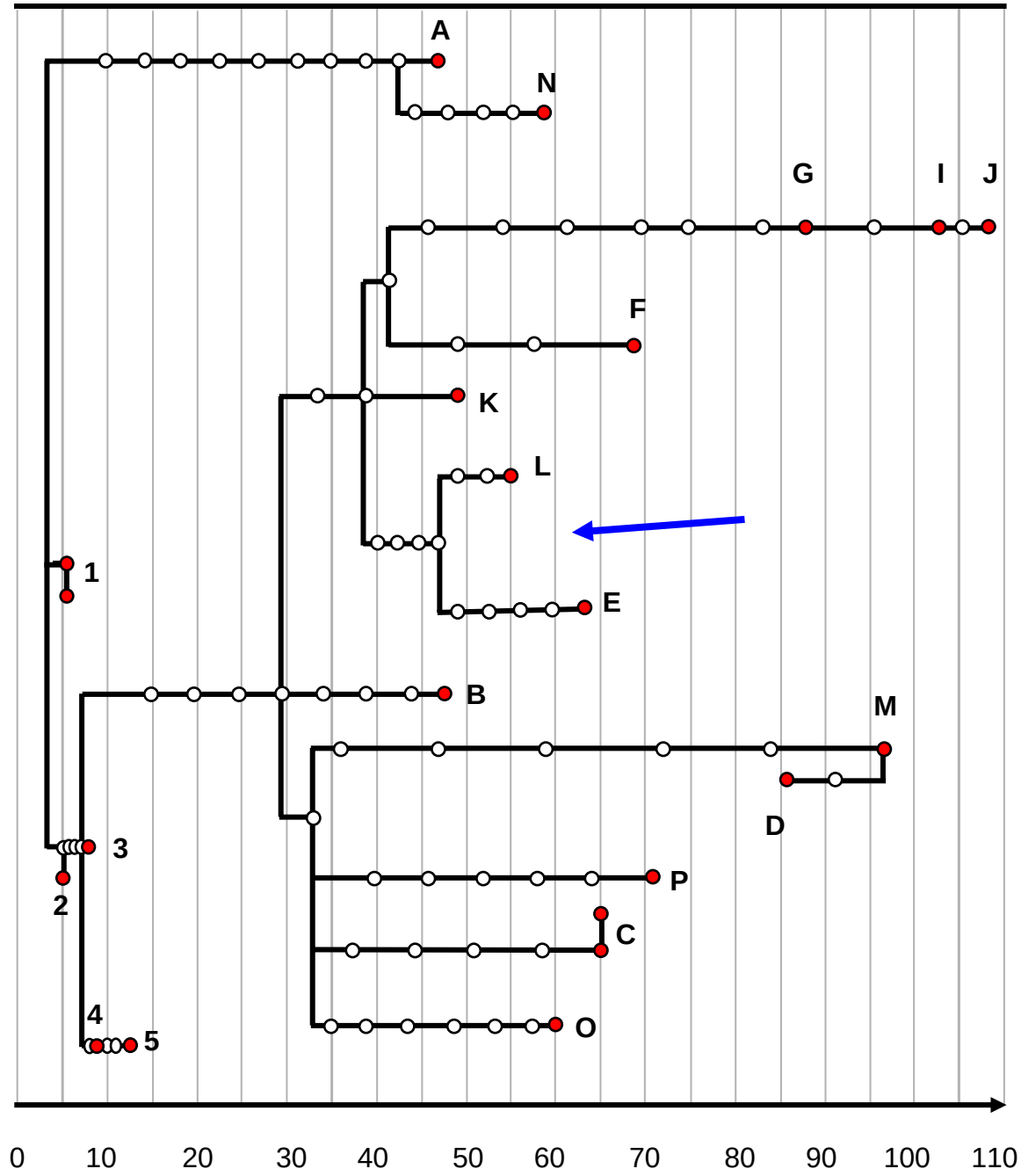
Cottam *et al*. 2008 Proc Biol Sci. 22;275(1637):887-95.

Statistical parsimony analysis of sequence data from IPs by TCS



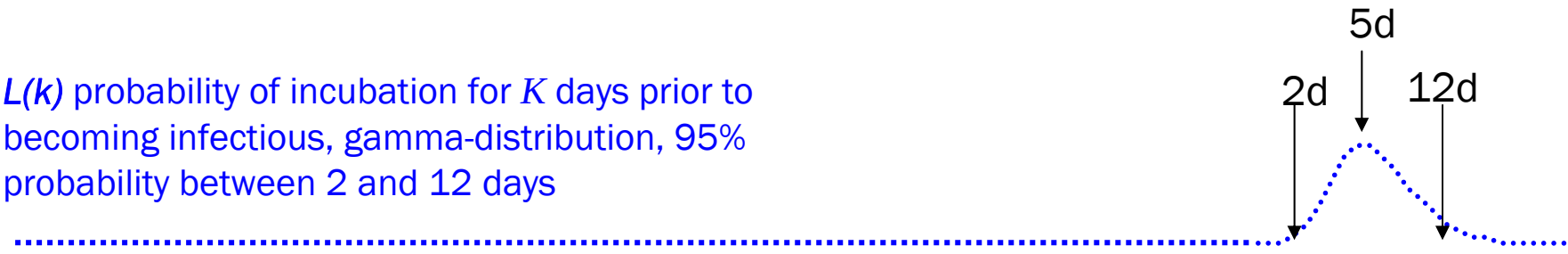
Difficult to determine transmission tree – 839808 different trees consistent with genetic data

Cottam *et al*. 2008 Proc Biol Sci. 22;275(1637):887-95.



- Restricted by date of cull
- Individual farm infection profiles?
- Statistical weighting for transmission tree – how confident are we?

Cottam *et al*. 2008 Proc Biol Sci. 22;275(1637):887-95.

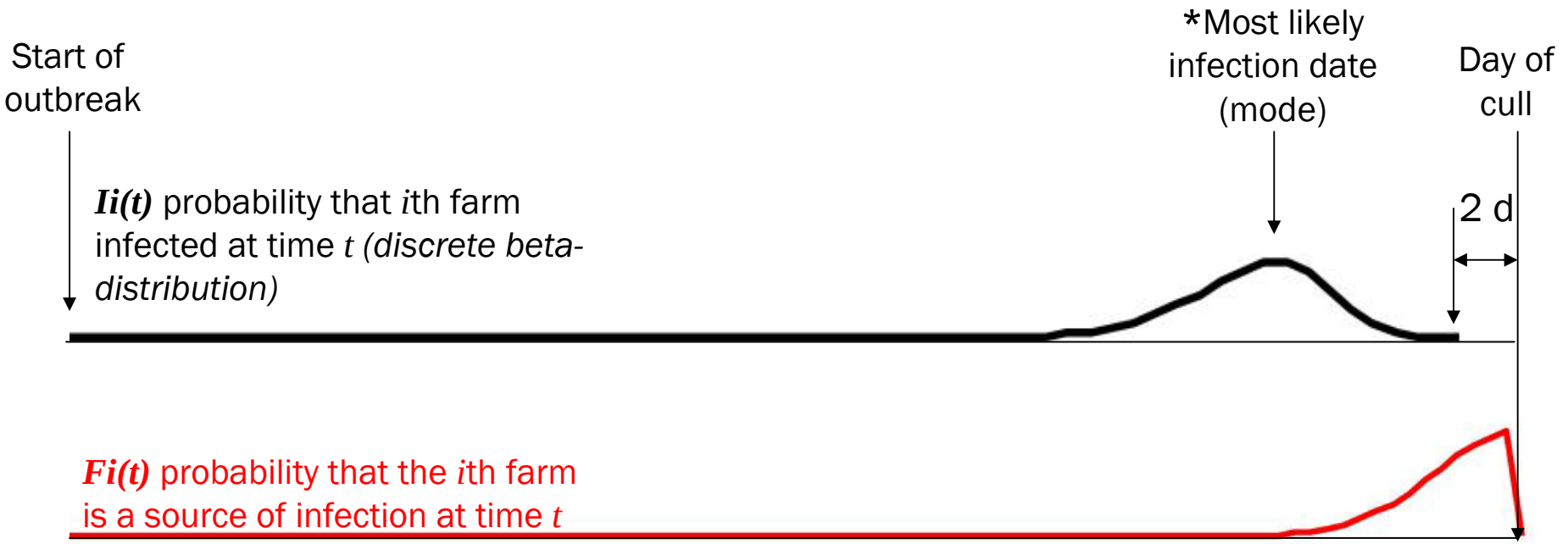


*date of confirmation minus oldest lesion minus 5d incubation

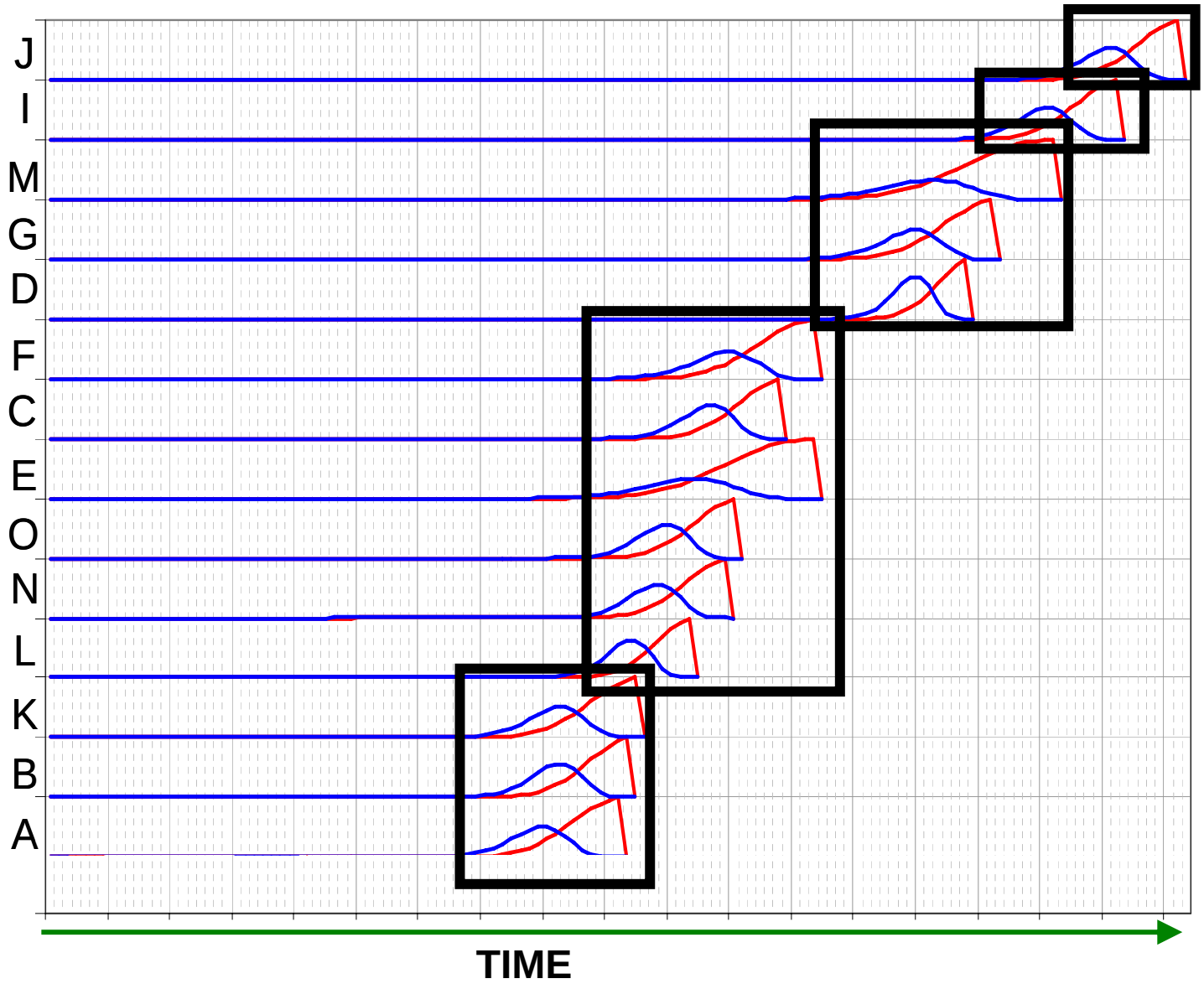
Probability that the *i*th farm is a source of infection at time t :

$$F_i(t) = \sum_{j=0}^{C_i} \left(I_i(j) \cdot \left(\sum_{k=1}^{t-j} L(k) \right) \right)$$

Cottam *et al*. 2008 Proc Biol Sci. 22;275(1637):887-95.


Start of outbreak
*Most likely infection date (mode)
Day of cull
2 d
Ii(t) probability that ith farm infected at time t (discrete beta-distribution)
Fi(t) probability that the ith farm is a source of infection at time t

Infection profiles of farms



Cottam *et al*. 2008 Proc Biol Sci. 22;275(1637):887-95.

From these infection profiles we can calculate a likelihood for each hypothesised transmission event

Likelihood that farm *i* infected farm *j* :

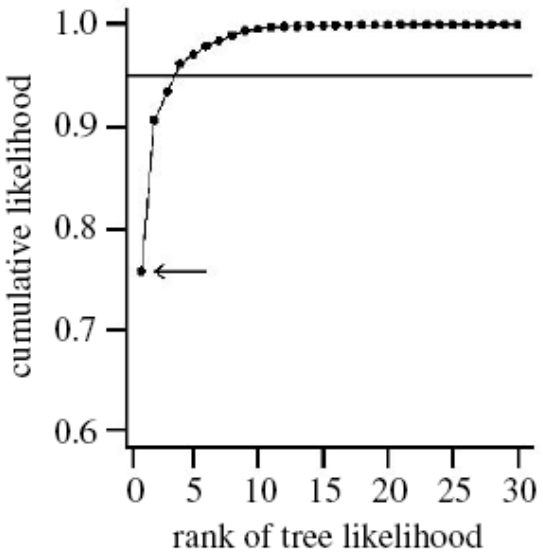
$$\ell_{ij} = \sum_{t=0}^{C_j} I_i(t) \cdot F_j(t) \Big/ \sum_{\substack{k=1 \\ k \neq i}}^{n} \sum_{t=0}^{C_k} I_i(t) \cdot F_k(t)$$

It is now possible to determine the transmission tree with the highest likelihood from all the trees consistent with the genetic data

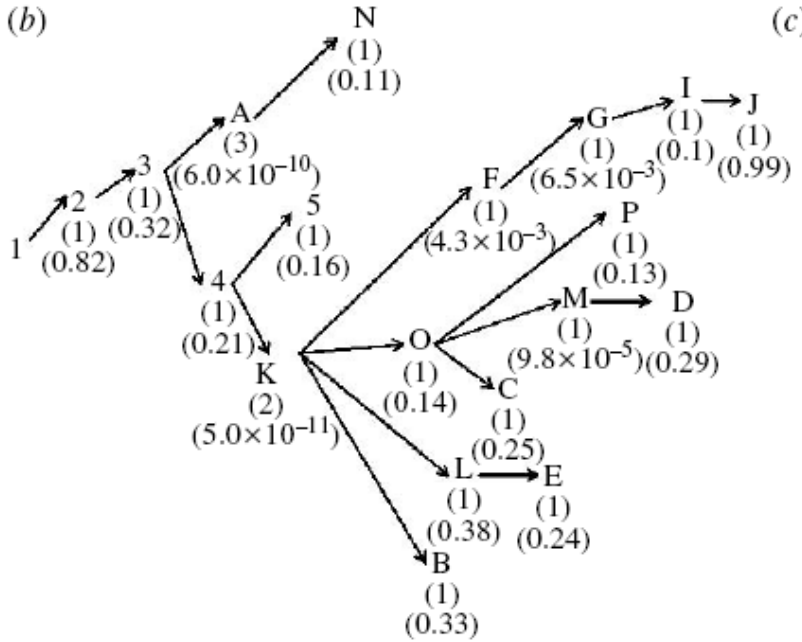
Cottam *et al*. 2008 Proc Biol Sci. 22;275(1637):887-95.

Most likely tree

(a)



(b)

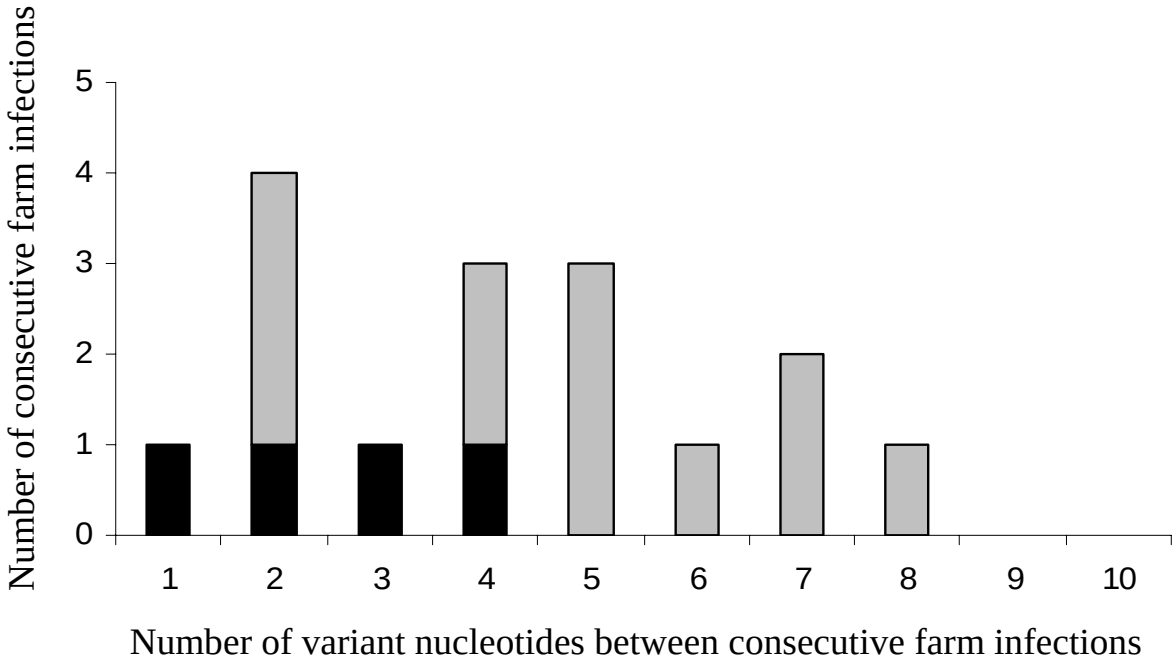


(c)

source farm	recipient farm	likelihood ratio
3	A	27.3
4	K	5.1
A	N	2.1×10^{16}
F	G	8.3×10^{11}
K	B	168
K	F	4.5×10^{3}
K	L	94.7
K	O	84.6
L	E	94.7
O	C	84.6
O	M	84.6
O	P	84.6

Cottam *et al*. 2008 Proc Biol Sci. 22;275(1637):887-95.

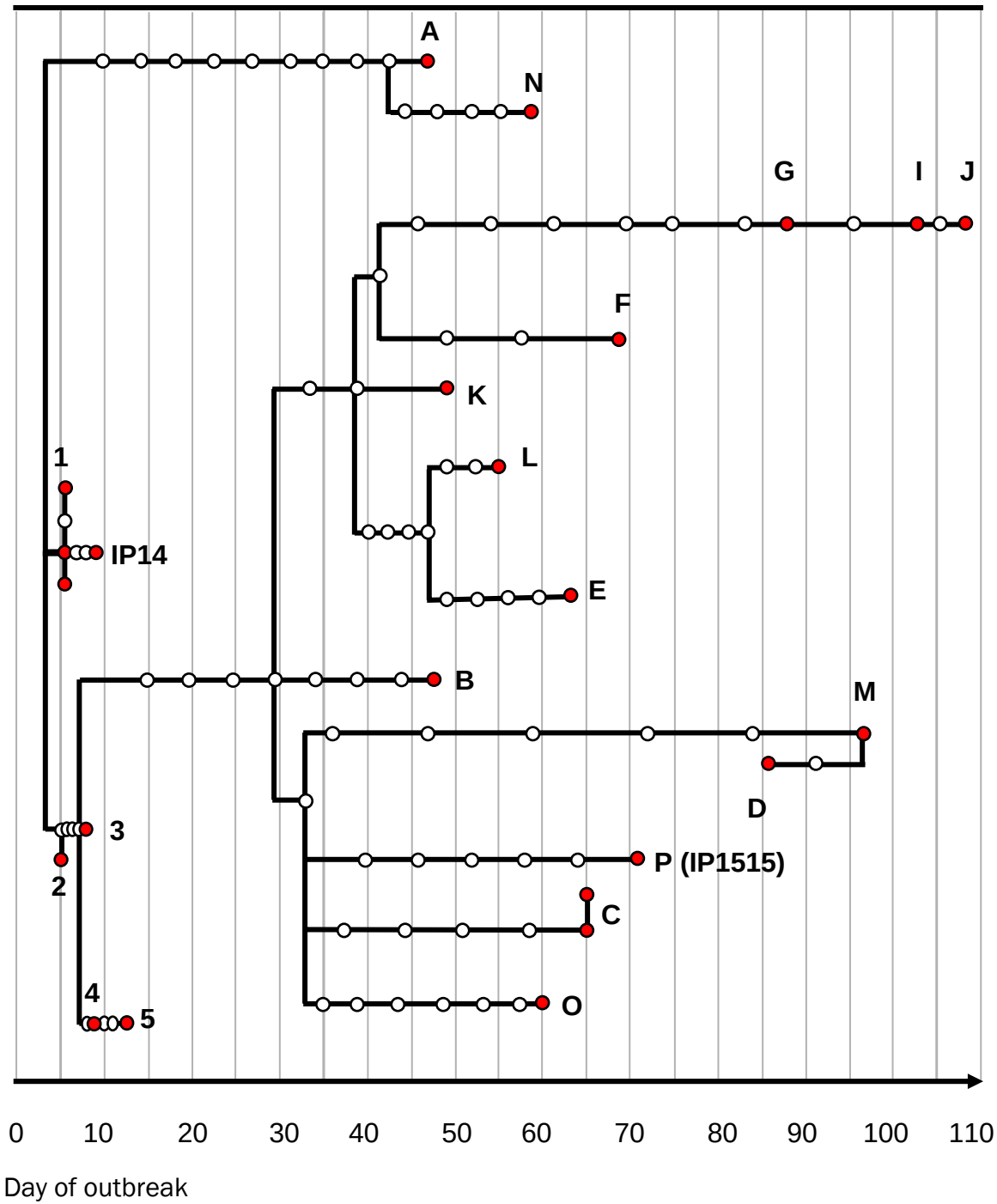
Distribution of the number of changes that occur upon transmission of virus between farms (n=16)



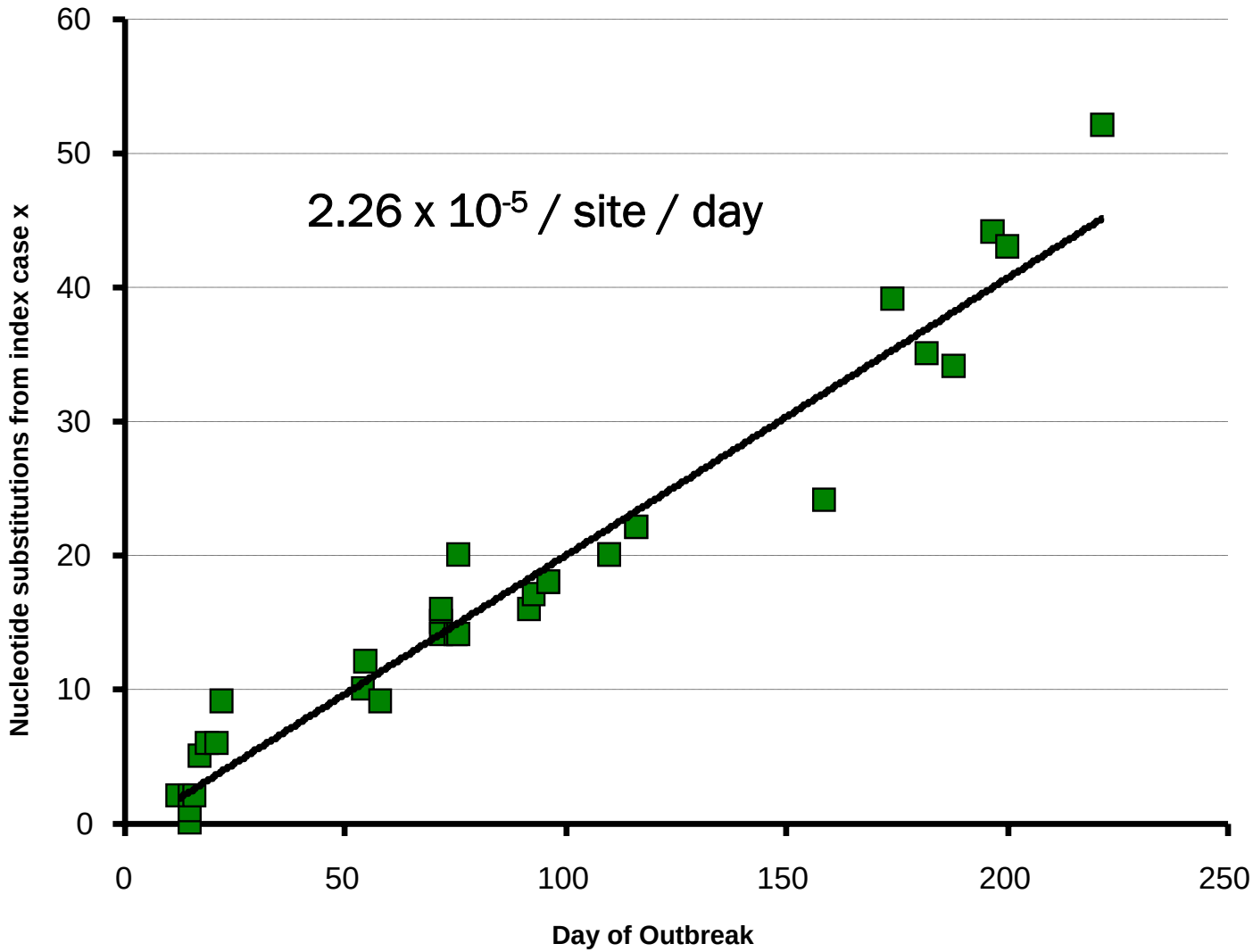
Mean 4.2 (SD 2.1) nucleotide changes per transmssion event

Thus if the number of changes seen is outside this distribution, a missing intermediate farm would be suspected

Cottam *et al*. 2008 Proc Biol Sci. 22;275(1637):887-95.



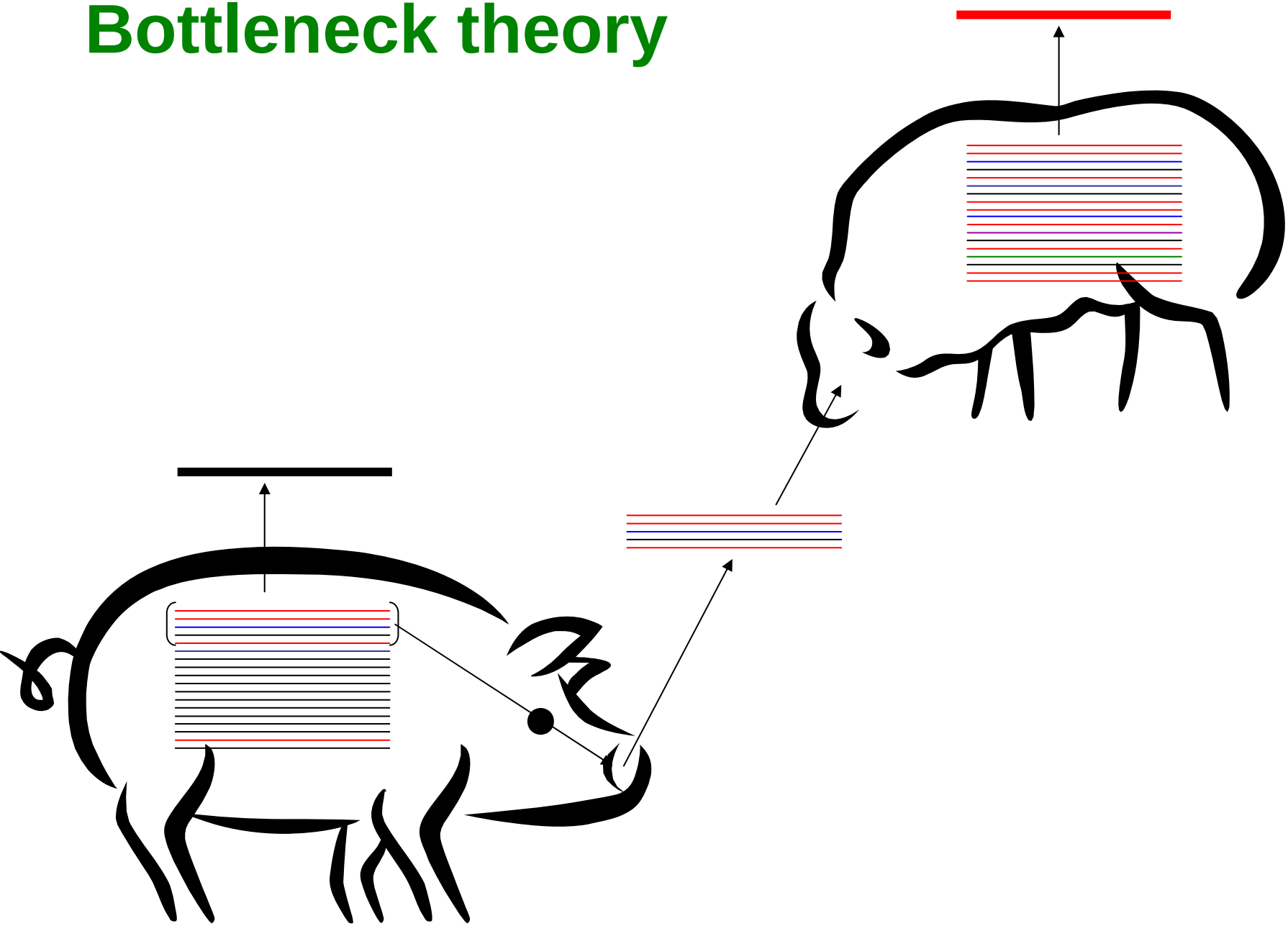
Rate of nucleotide change seems to vary.....


2.26 x 10-5 / site / day
Nucleotide substitutions from index case x
Day of Outbreak
0
10
20
30
40
50
60
0
50
100
150
200
250

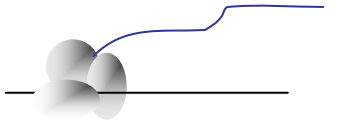
- Previously linked mutation rate with time

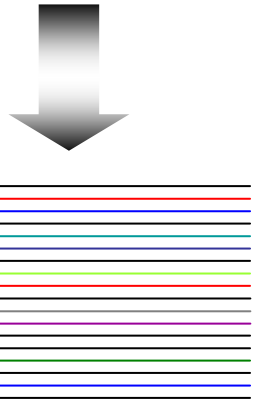
- Was the rate of genetic evolution of FMDV in the epidemic due in part to the rate of transmission and spread of the virus?

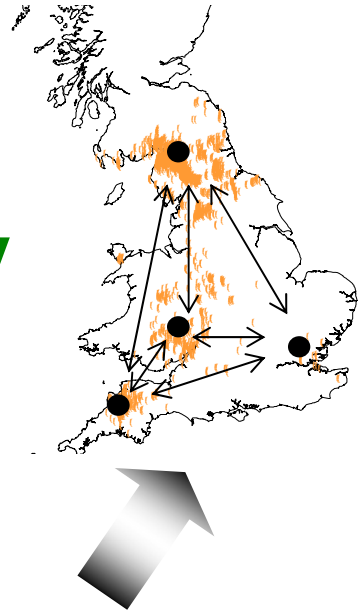
Bottleneck theory

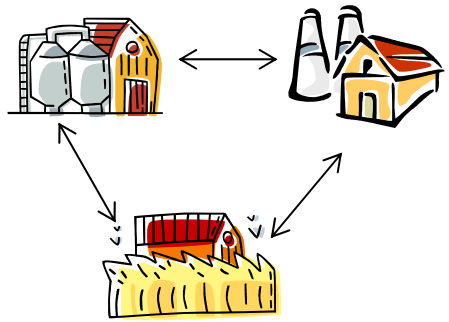


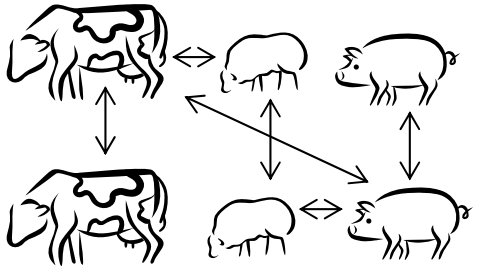
Genetic evolution of FMDV



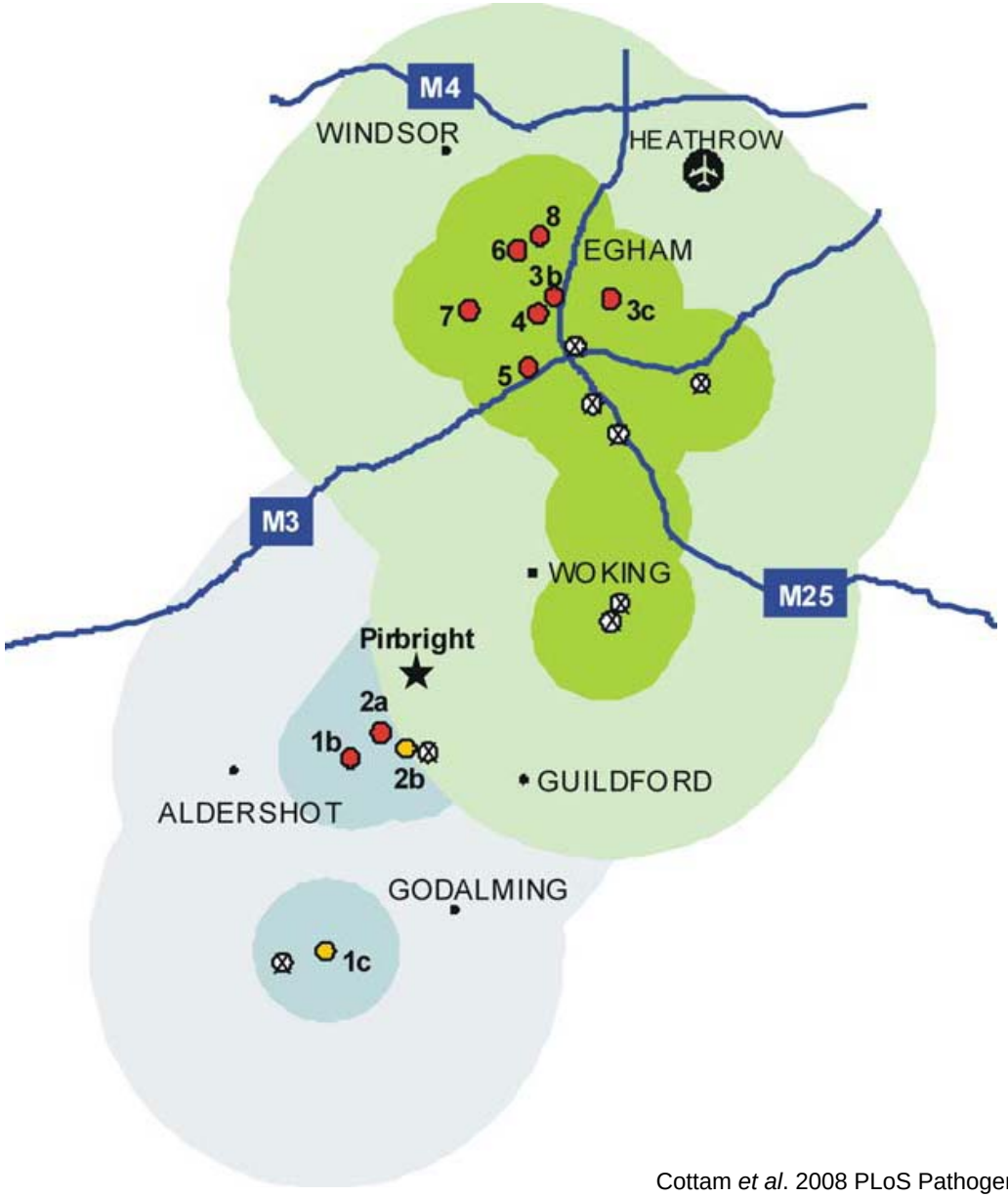






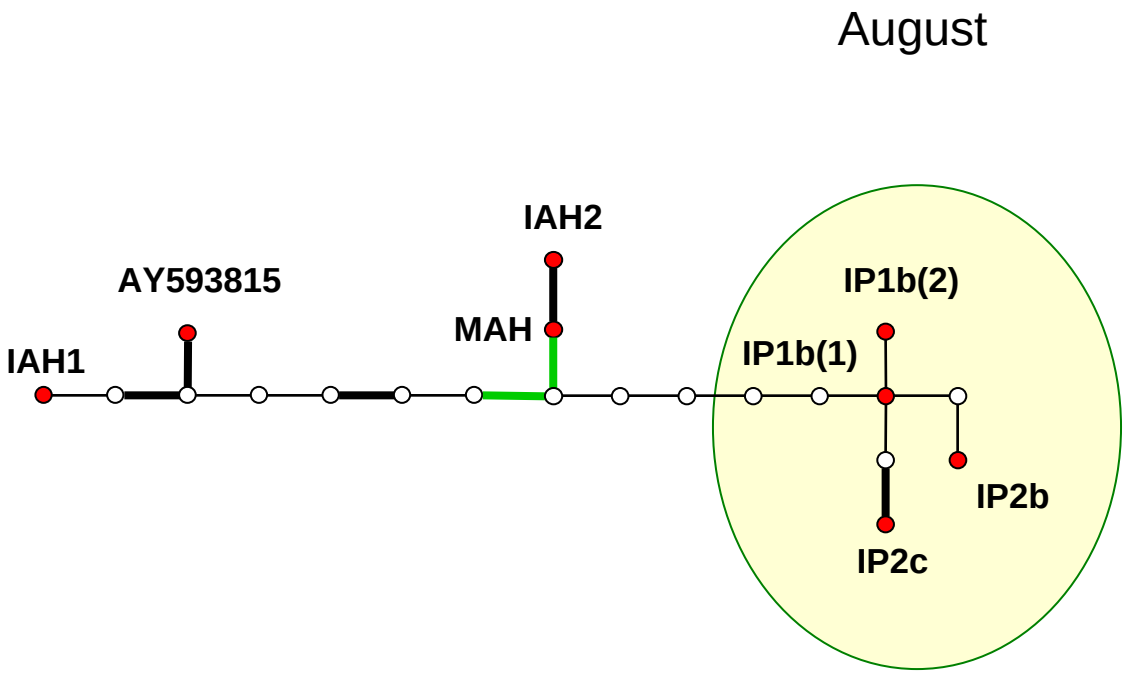


2007 UK FMDV Outbreaks



Cottam *et al*. 2008 PLoS Pathogens 4(4): e1000050

2007 UK FMDV Outbreaks

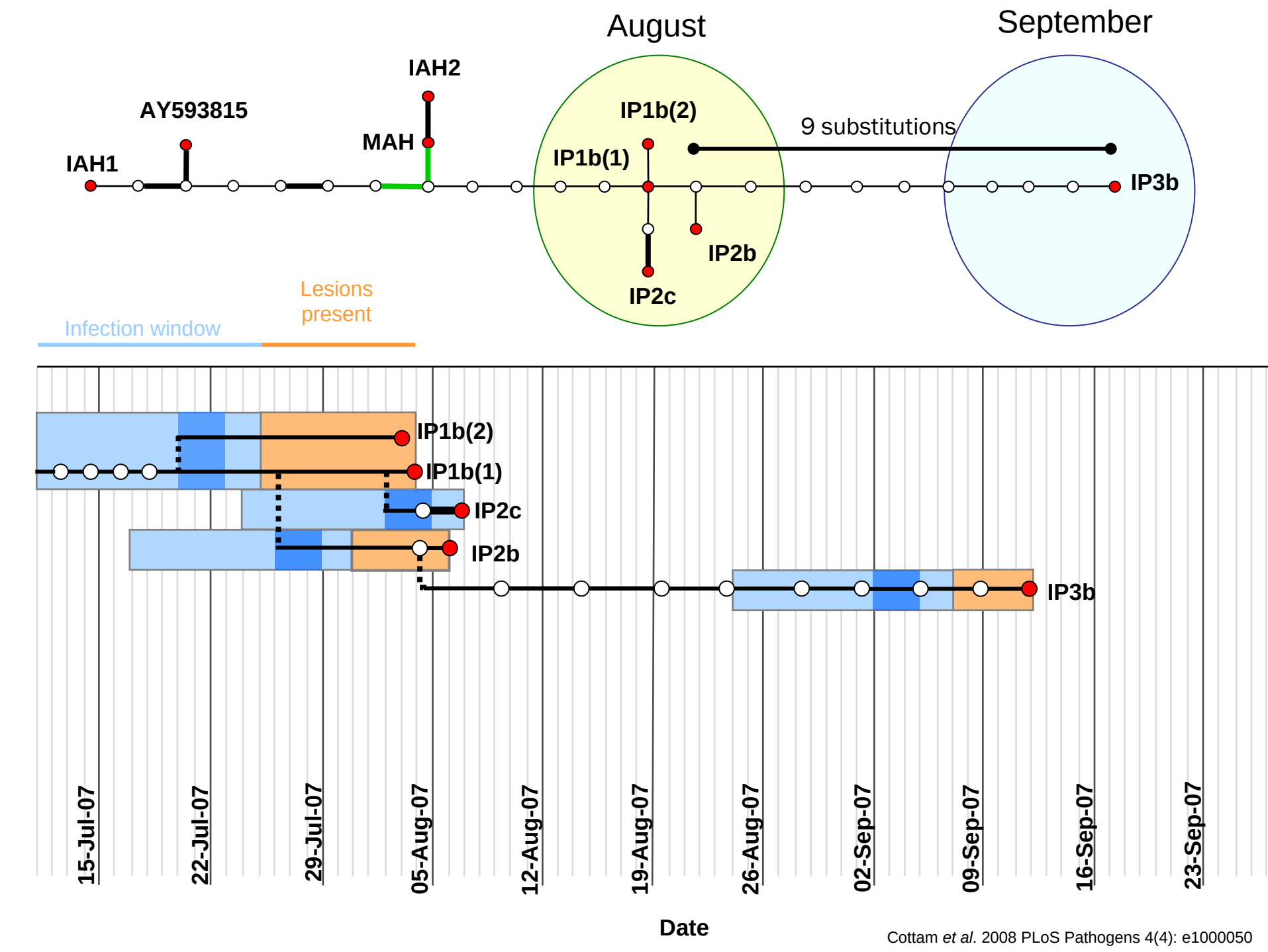


Synonymous substitution

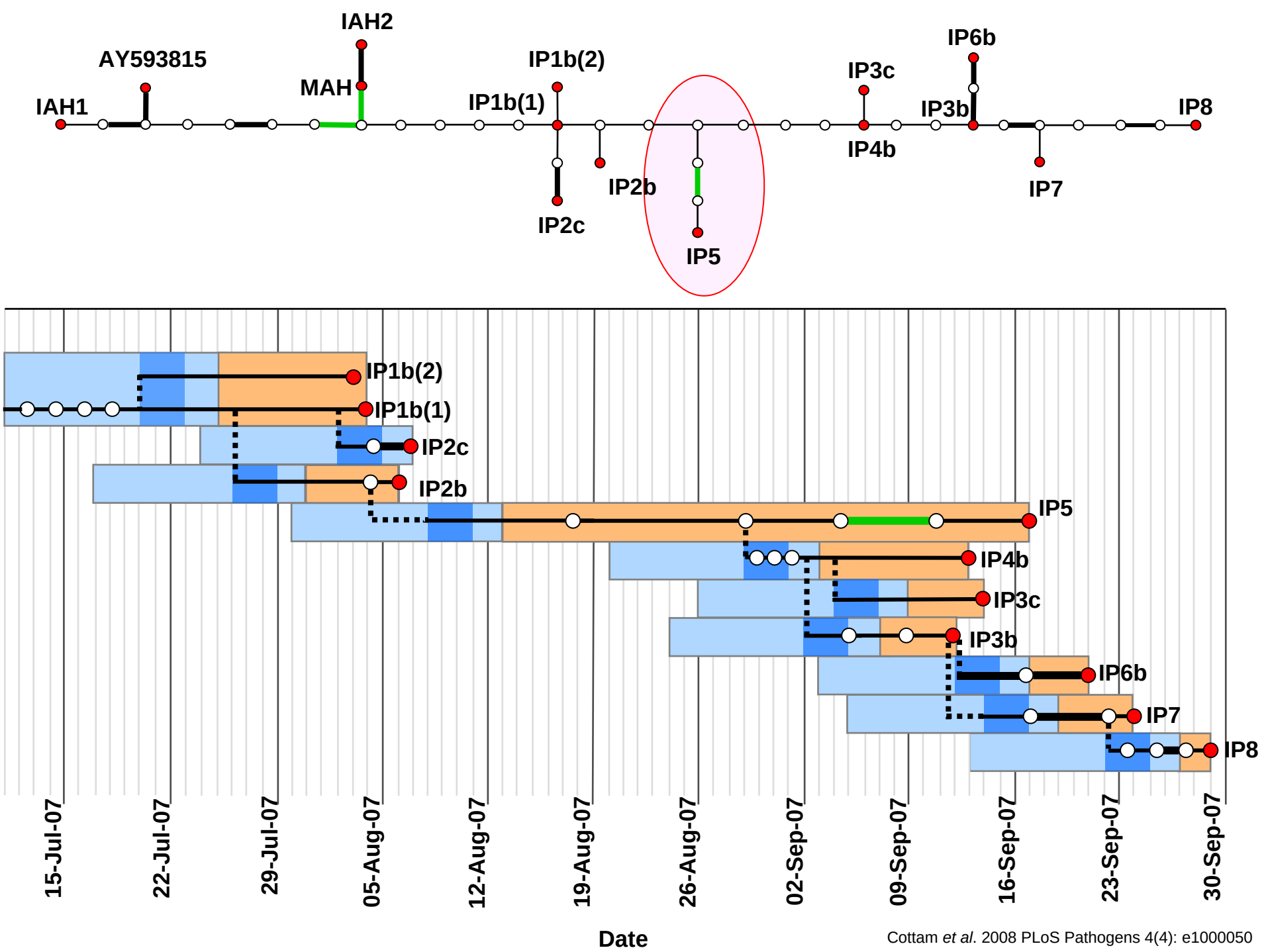
Non-synonymous substitution

Non-synonymous substitution important in cell culture adaptation

Cottam *et al*. 2008 PLoS Pathogens 4(4): e1000050



Cottam *et al*. 2008 PLoS Pathogens 4(4): e1000050



Cottam *et al*. 2008 PLoS Pathogens 4(4): e1000050

Summary

- Complete genome sequencing data can assist in our understanding of the spread of FMDV
- Combining genetic and epidemiological data gives greater resolution.
- More research is needed to help refine the use of these data in real time.

Thank you

Don King
David Paton
Nick Knowles
Nigel Ferris
Geoff Hutchings
Andrew Shaw
Jemma Wadsworth
John Gloster

Dan Haydon
Gael Thébaud
John Wilesmith
Sam Mansley