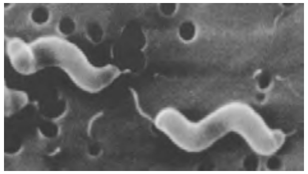


Molecular tools for surveillance

Nigel French

Surveillance workshop
March 2010

Hopkirk Research Institute

EpiCentre

http://epicentre.massey.ac.nz/

Te Kunenga ki Pūrehuroa

MASSEY UNIVERSITY

Overview

- Application of molecular tools for surveillance
  - Event based and policy based
- Brief overview of tools available
- Application of sequence typing and modelling for surveillance of a zoonoses
  - Manawatu sentinel site
  - Space-time indicators (event)
  - Source attribution (policy)

Infectious disease:
need for molecular tools

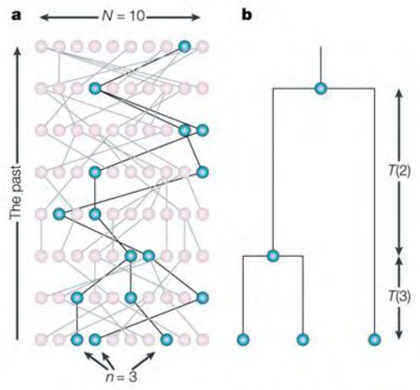
- Pathogen characterisation, detection and diagnosis *(present)*
  - Tracking outbreaks, surveillance
- Origins and ecology *(past)*
  - Evolutionary biology / population genetics
- Emergence *(future)*
  - Identify obstacles to prevention and treatment
    - AR, vaccination.
  - Risks and evolutionary trends – ‘phylodynamics’
    - Where (geographically, hosts species), when, and in what form will pathogens emerge

Use of molecular tools for surveillance

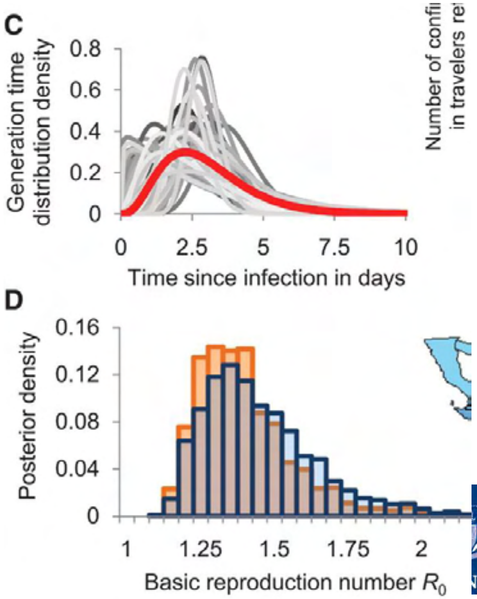
- Event /Control-based surveillance
  - Identify common source
    - Indistinguishable genotypes
      - Clonal, amplifying organisms
      - PulseNet, food chain, VTEC O157, Salmonella.
    - Genotypes indicating common source or pathway
      - Source/pathway signature
      - Host associated genotypes
      - Non-clonal, non amplifying, mixed genotypes
      - Campylobacter, MLST

Use of molecular tools for surveillance

- Policy /Strategy-based surveillance
  - Source attribution
    - Zoonoses
    - Wildlife reservoirs (FMDv buffalo/impala)
  - Vaccination
    - Monitoring of epidemic strains (FMDv)
    - Response to vaccination (MeNZB^{TM})
  - Emergence
    - Time of emergence
  - Invasion
    - Vector surveillance, multiple vs single incursion
  - Estimating R_o from population expansion
    - H1N1 outbreak
  - Compliance / trade
    - NMD



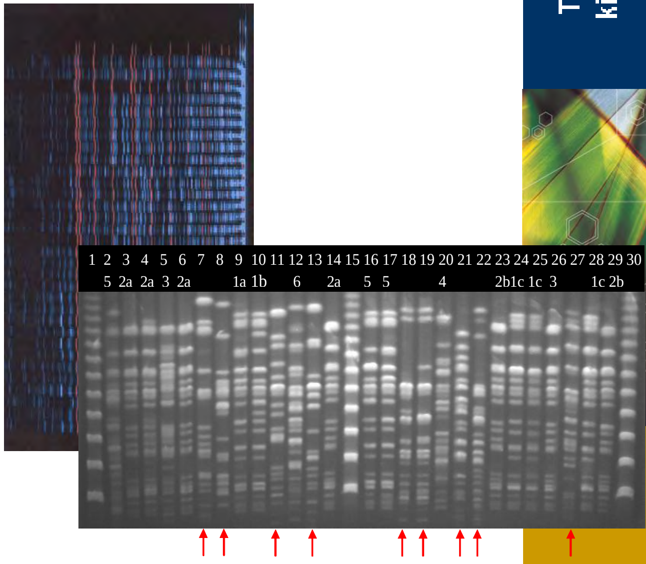
Nature Reviews | Genetics



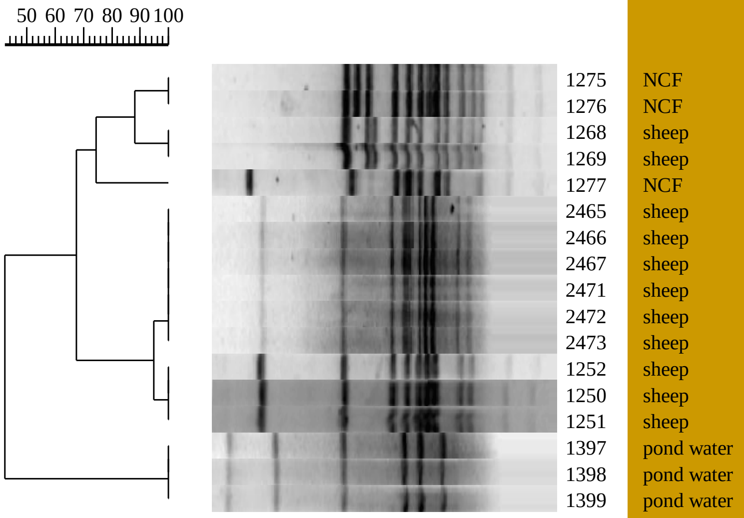
Molecular tools (genotyping)

<u>Tools for measuring ‘relatedness’</u>

- Same/different/relative - polymorphisms
- Based on mutation (mistakes) and recombination (horizontal gene transfer, phage..)

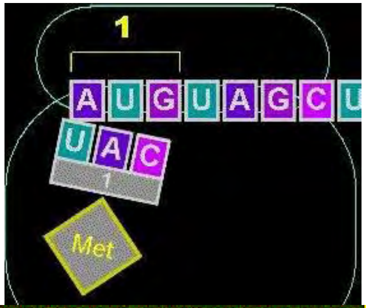


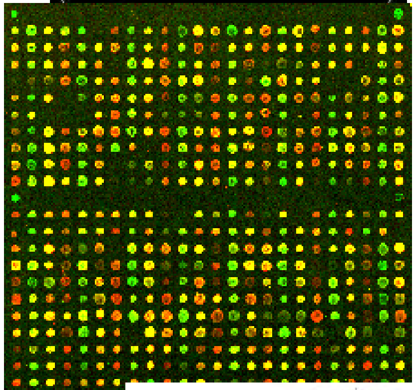
- PCR (Polymerase chain reaction)
  - Random or specific genes/sequences (AFLP, MLG)
  - SNP analysis (RT assays)
  - Binary genomotyping, VNTR
- RFLP (Restriction fragment length polymorphism)
  - Whole genome (PFGE)
  - Individual genes (*fla*A typing)

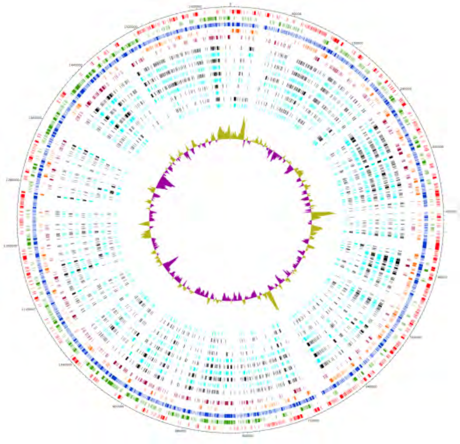


Molecular epidemiology - tools

- Hybridisation techniques
  - Microarray…..
- Sequence analysis
  - Single genes / sequences
  - Multilocus sequence typing
  - Whole genome comparisons

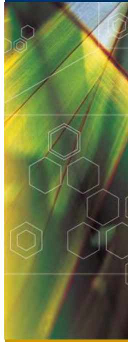




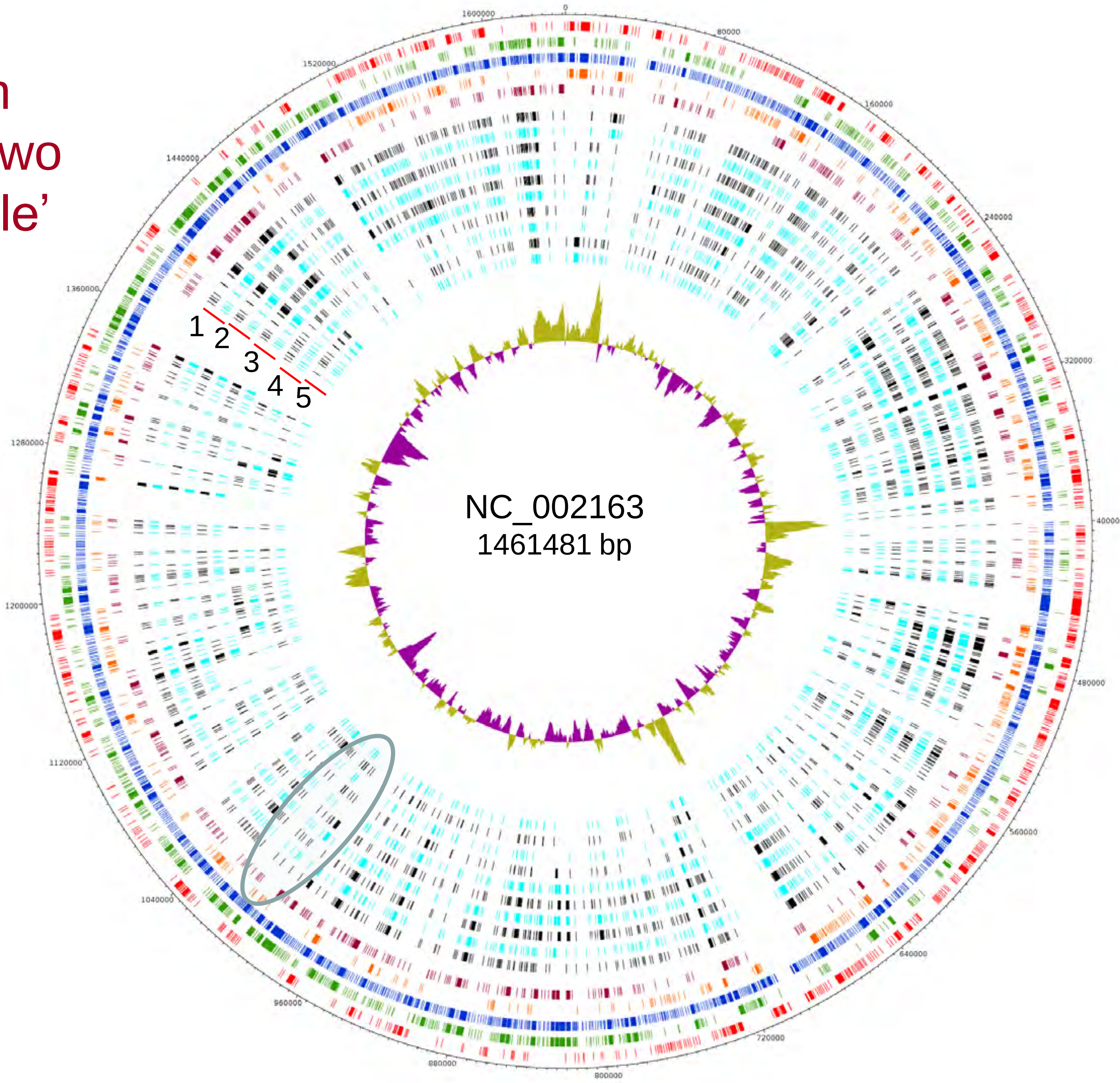


What makes a good molecular tool for surveillance?

- Good typeability (proportion that can be typed)
- Discriminatory
  - Optimum level of discrimination depends on setting (lumping and splitting, event and policy based surveillance)
- Epidemiologically useful
  - Identifies host association
  - Pathway association (food vs direct contact)
  - Informs choice of vaccine
- Rapid, inexpensive
- Comparable and portable
  - YATM?
  - International databases



Discrimination: Next generation sequencing of two ‘indistinguishable’ strains



Over 100 core genes with SNPs + insertions and deletions

Pooled DNA for SNP identification?

What makes a good molecular tool for surveillance?

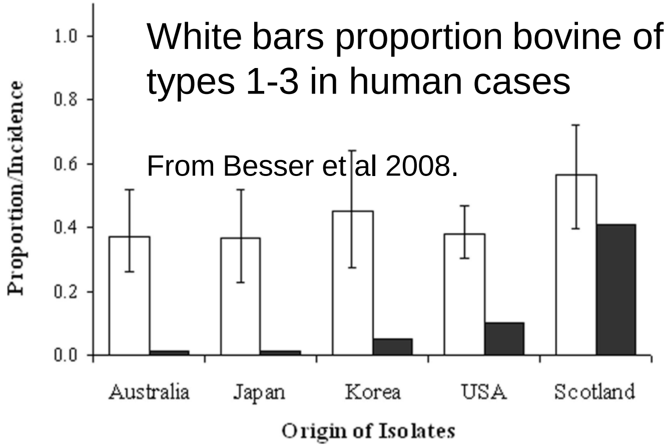
- Good typeability (proportion that can be typed)
- Discriminatory
  - Optimum level of discrimination depends on setting (lumping and splitting)
- Epidemiologically useful
  - Identifies host association
  - Pathway association (food vs direct contact)
  - Informs choice of vaccine
- Rapid, inexpensive
- Comparable and portable
  - YATM?
  - International databases

STEC O157

stx insertion typing identifies “bovine-biased” and “clinical biased” lineages

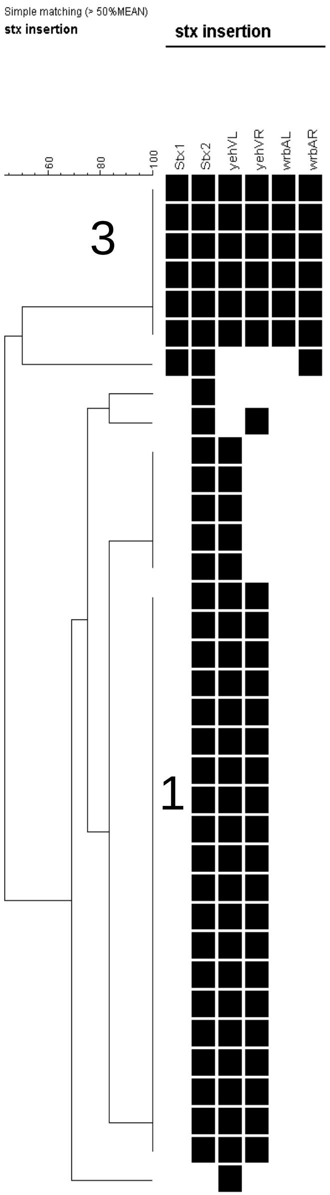
Useful for identifying both source and pathway?

Food vs environmental



White bars proportion bovine of types 1-3 in human cases

From Besser et al 2008.



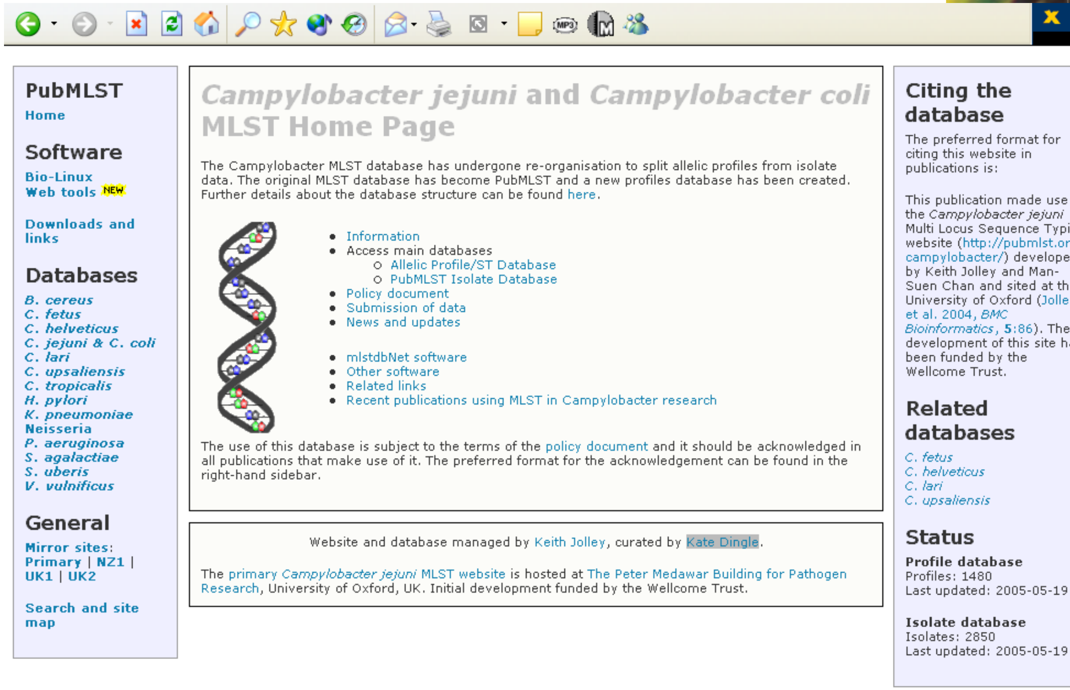
Key		Type	Source	concatenate	genotype
🟩	12	O157:H7	Human	111111	3
🟩	13	O157:H7	Human	111111	3
🟩	15	O157:H7	Human	111111	3
🟥	26	O157	Bovine	111111	3
🟥	27	O157	Bovine	111111	3
🟥	29	O157	Bovine	111111	3
🟩	3	O157	Human	110001	U
🟩	2	O157:H7	Human	010000	11
🟥	31	O157	Bovine	010100	19
🟩	17	O157:H7	Human	011000	5
🟩	18	O157:HNM	Human	011000	5
🟩	19	O157:HNM	Human	011000	5
🟦	20	O157:H7	Spring Water	011000	5
🟩	21	O157:H16	Human	011000	5
🟩	10	O157:H7	Human	011100	1
🟩	11	O157:H7	Human	011100	1
🟦	14	O157:H7	Spring Water	011100	1
🟩	16	O157:H7	Human	011100	1
🟥	22	O157:H7	Bovine	011100	1
🟥	23	O157	Bovine	011100	1
🟥	24	O157	Bovine	011100	1
🟥	25	O157	Bovine	011100	1
🟥	28	O157	Bovine	011100	1
🟥	30	O157	Bovine	011100	1
🟥	32	O157	Bovine	011100	1
🟥	33	O157	Bovine	011100	1
🟥	34	O157	Bovine	011100	1
🟥	35	O157	Bovine	011100	1
🟩	4	O157:H7	Human	011100	1
🟩	5	O157:H7	Human	011100	1
🟩	6	O157:H7	Human	011100	1
🟩	7	O157:H7	Human	011100	1
🟩	8	O157:HNM	Human	011100	1
🟩	9	O157:H7	Human	011100	1
🟩	1	O157:H7	Human	001000	17

What makes a good molecular tool for surveillance?

- Good typeability (proportion that can be typed)
- Discriminatory
  - Optimum level of discrimination depends on setting (lumping and splitting)
- Epidemiologically useful
  - Identifies host association
  - Pathway association (food vs direct contact)
  - Informs choice of vaccine
- Rapid, inexpensive
- Comparable and portable
  - YATM?
  - International databases

International databases

- International initiatives (harmonisation, comparison)
  - PulseNet
    - *E. coli*
    - *Campylobacter*
  - MLST
    - Oxford website
    - *Neisseria*
    - *Streptococcus*
    - *Bacillus*



PubMLST
Home

Software
Bio-Linux
Web tools NEW

Downloads and links

Databases
B. cereus
C. fetus
C. helveticus
C. jejuni & C. coli
C. lari
C. upsaliensis
C. tropicalis
H. pylori
K. pneumoniae
Neisseria
P. aeruginosa
S. agalactiae
S. uberis
V. vulnificus

General
Mirror sites:
Primary | NZ1 |
UK1 | UK2

Search and site map

Campylobacter jejuni and *Campylobacter coli* MLST Home Page

The Campylobacter MLST database has undergone re-organisation to split allelic profiles from isolate data. The original MLST database has become PubMLST and a new profiles database has been created. Further details about the database structure can be found here.

- Information
- Access main databases
  - Allelic Profile/ST Database
  - PubMLST Isolate Database
- Policy document
- Submission of data
- News and updates

- mlstdbNet software
- Other software
- Related links
- Recent publications using MLST in Campylobacter research

The use of this database is subject to the terms of the policy document and it should be acknowledged in all publications that make use of it. The preferred format for the acknowledgement can be found in the right-hand sidebar.

Website and database managed by Keith Jolley, curated by Kate Dingle.

The primary *Campylobacter jejuni* MLST website is hosted at The Peter Medawar Building for Pathogen Research, University of Oxford, UK. Initial development funded by the Wellcome Trust.

Citing the database

The preferred format for citing this website in publications is:

This publication made use
the *Campylobacter jejuni*
Multi Locus Sequence Typi
website (http://pubmlst.or
campylobacter/) develope
by Keith Jolley and Man-Suen Chan and sited at th
University of Oxford (Jolle
et al. 2004, *BMC Bioinformatics*, **5**:86). The
development of this site h
been funded by the Wellcome Trust.

Related databases
C. fetus
C. helveticus
C. lari
C. upsaliensis

Status

Profile database
Profiles: 1480
Last updated: 2005-05-19

Isolate database
Isolates: 2850
Last updated: 2005-05-19

Application of molecular tools for the surveillance of campylobacteriosis

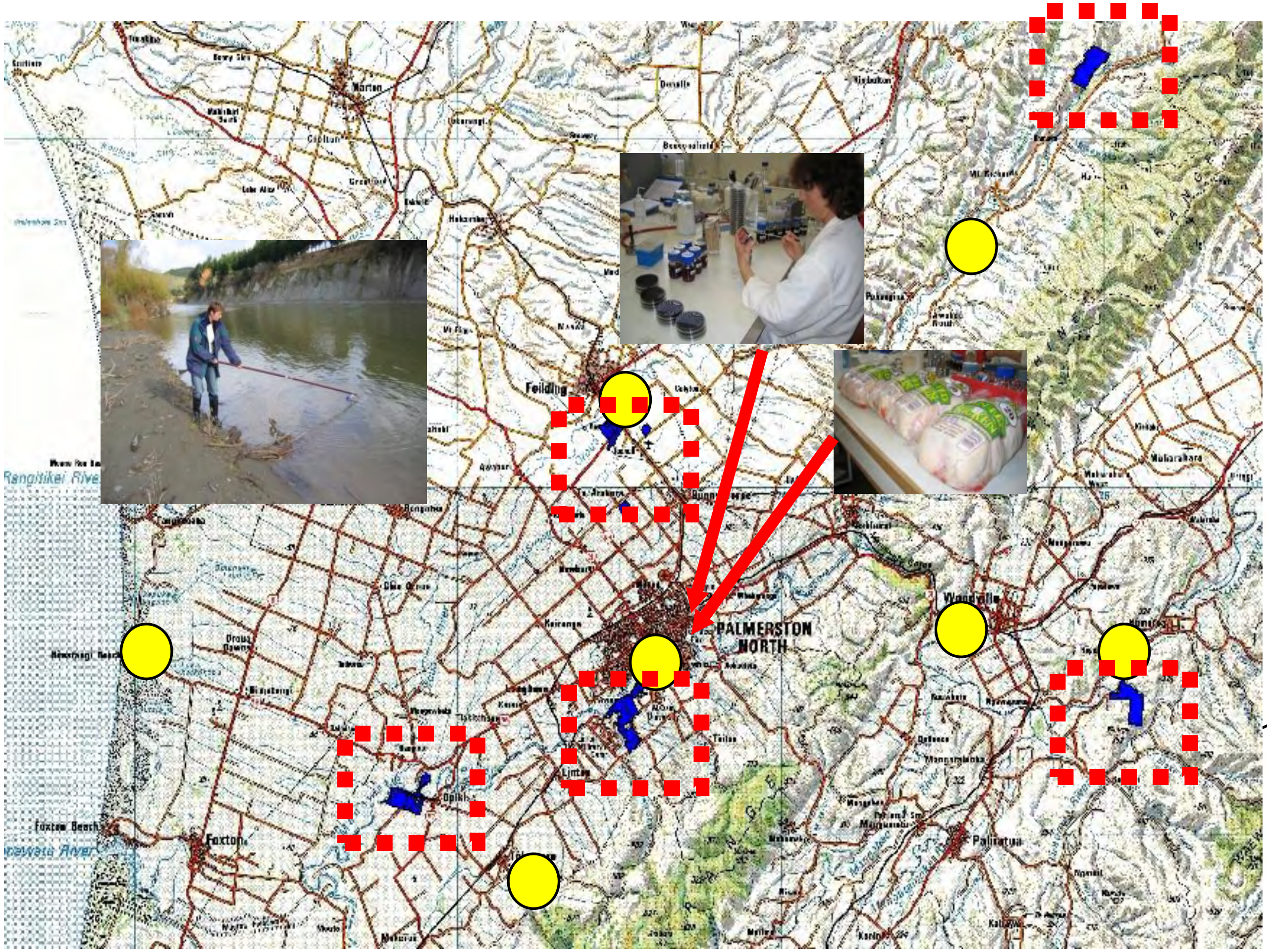
- Identifying and characterising outbreak events
  - Common source?
- Source attribution
  - Microbial subtyping / source tracking
  - Applying molecular tools, population genetics and epidemiological modelling to inform public health policy
- NZFSA funded

Manawatu sentinel surveillance site

- Collaboration between:
  - MedLab Central
  - Public Health Unit
  - Human health surveillance unit: ESR
  - Regulator: NZFSA
  - Industry body: PIANZ / suppliers
  - Dairy and sheep farmers
  - Regional council
  - Institutes at Massey University

Manawatu study 2005-2009

- Sentinel site (4 yrs)



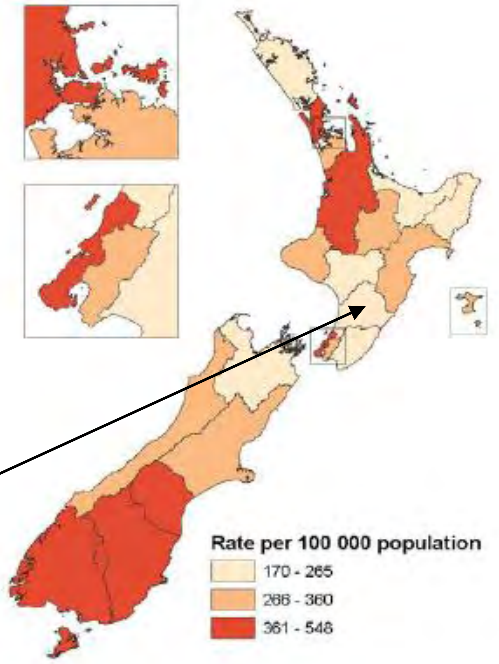


Figure 7. Campylobacteriosis notifications by DHB, 2004

Numbers of samples/isolates: *C. jejuni*

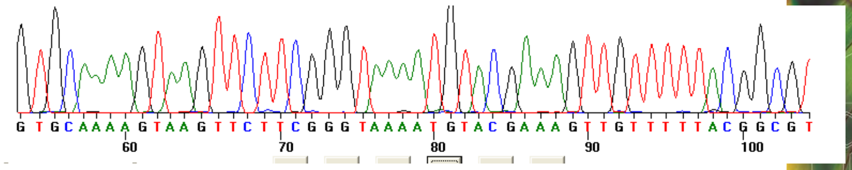
Human	520 (770 samples)	
Poultry	562 samples	75% +ve
Red meat	1312 samples	12% +ve
Ruminant faeces	278 samples	58% +ve
Env. Water	335 samples	30% +ve
Wild bird	192 samples	13% +ve

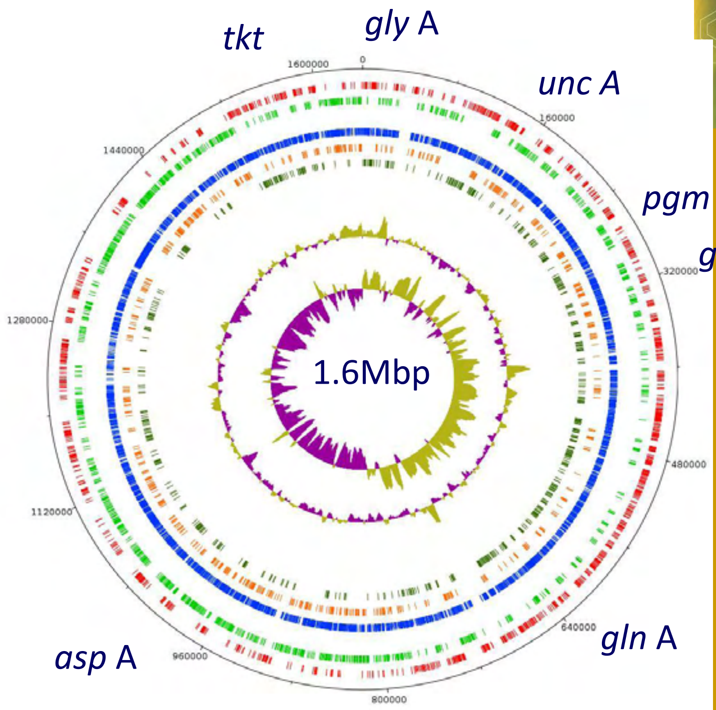
March 1st 2005 to Feb 29th 2008

Multi Locus Sequence Typing

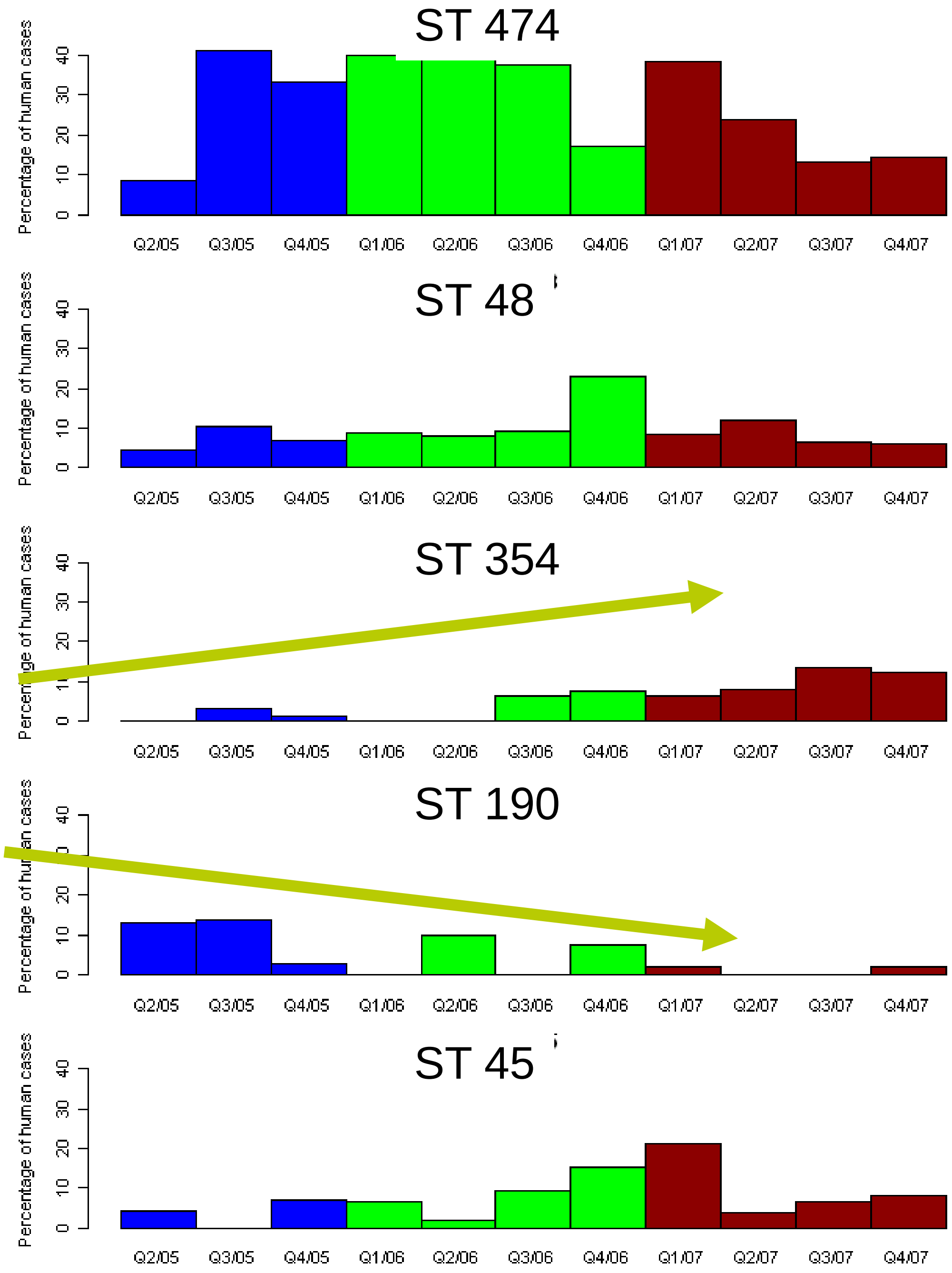
- PCR highly conserved genes
- 7 housekeeping genes
- Use allelic variation to describe subtypes:
  - ST = sequence type – unique pattern of 7 alleles
  - Clonal complex = group of related STs identified by progenitor ST
  - Website: Oxford University

http://campylobacter.mlst.net



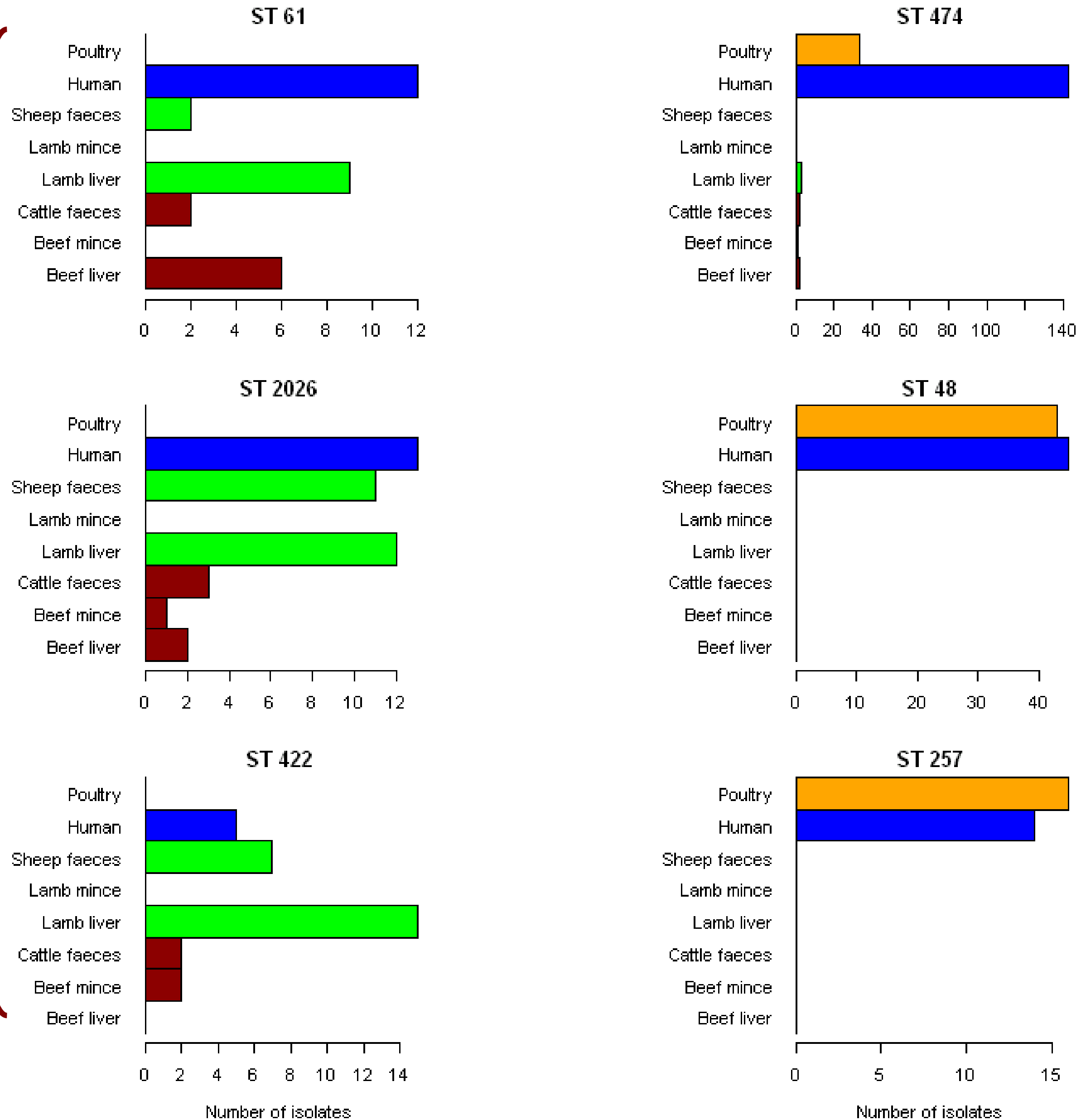


Human cases over 3-year period



Host associated sequence types in NZ

Ruminant associated



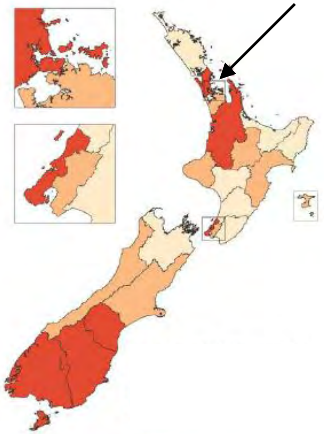
Poultry associated

Application of MLST and modelling

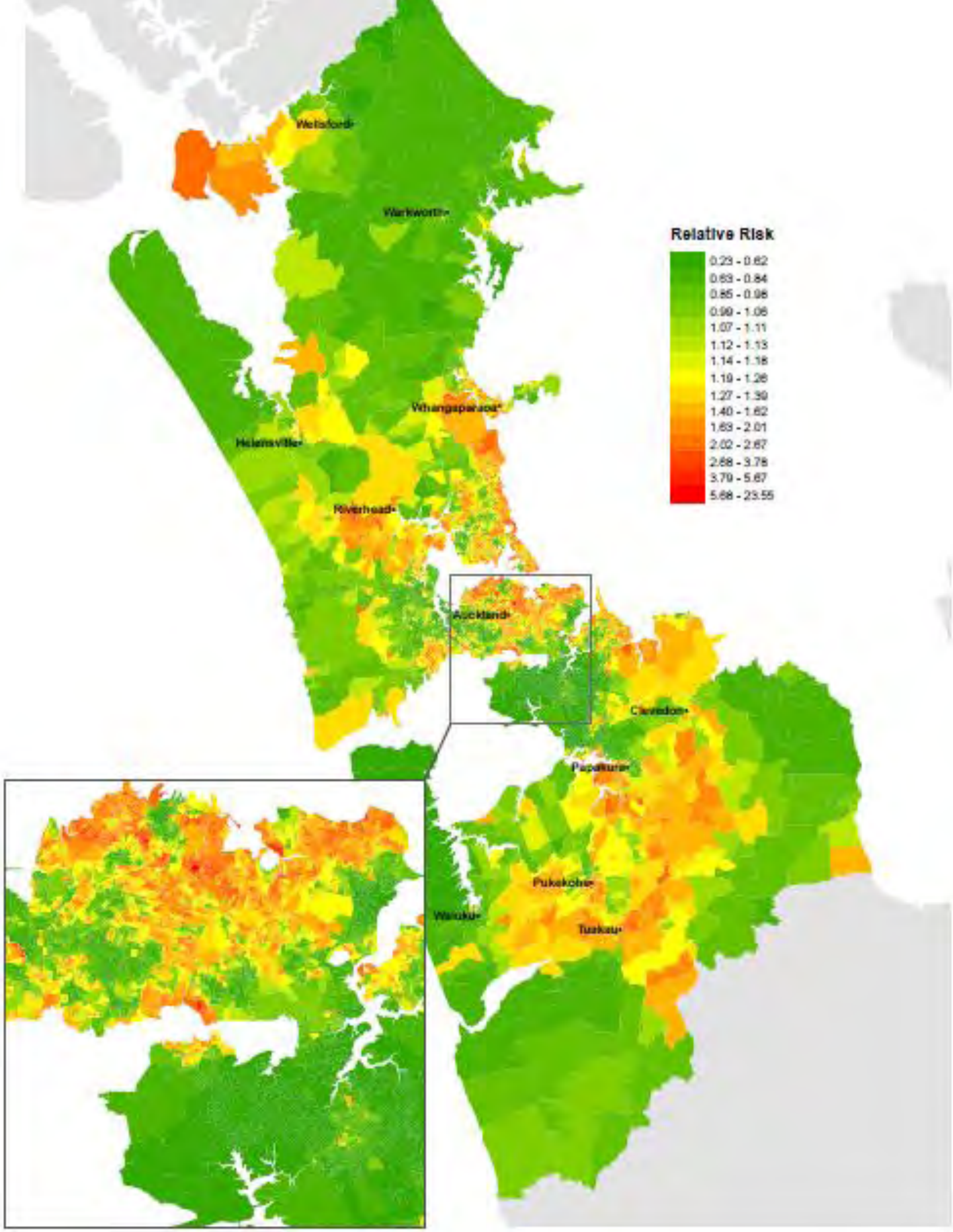
Event-based surveillance

- Identify space-time cluster in ‘real time’
  - SatScan
  - Knorr-Held and Richardson model
- Type samples from outbreak (+controls)
  - Indistinguishable?
  - Common source (ruminant, poultry)?
  - Common pathway (food, water..)?

Spatio-temporal modelling



meshblocks



Spatio-temporal modelling

$$Y_{i,t} \sim \text{Pois}(n_i \lambda_{i,t})$$

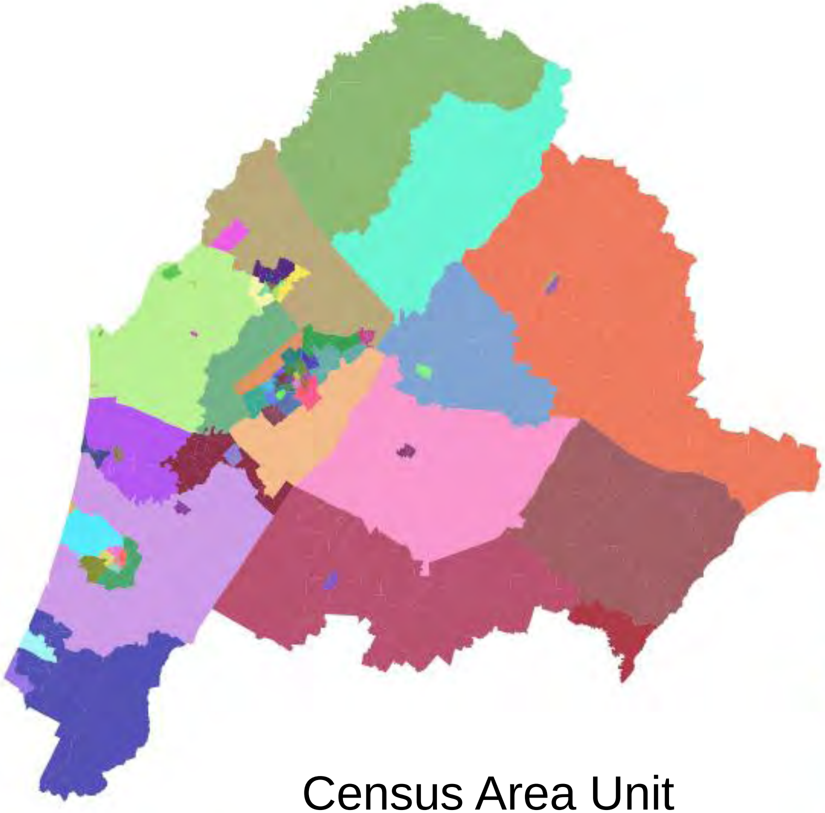
$$\log(\lambda_{i,t}) = R_t + U_i$$

Add space-time interaction term for ‘epidemic indicator’ model

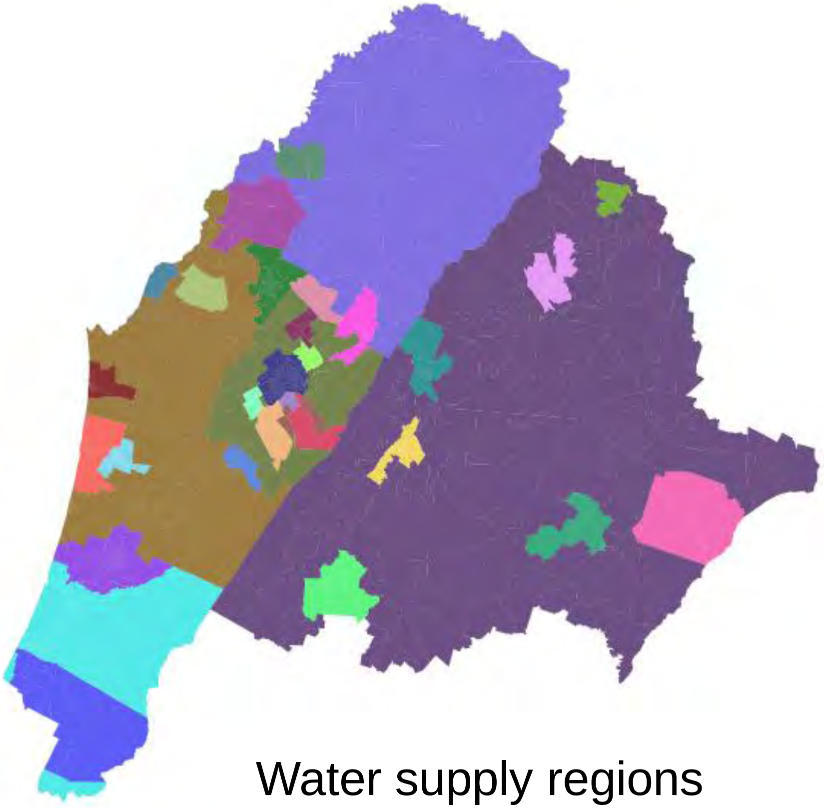
Temporal prior	Change in risk this week is similar to change in risk last week. (Second Order Gaussian Random Walk)
Spatial prior	Risk in a meshblock is similar to risk in neighbouring meshblocks. (Gaussian Markov Random Field)

Appropriate spatial resolution?

>9000 meshblocks in Auckland)

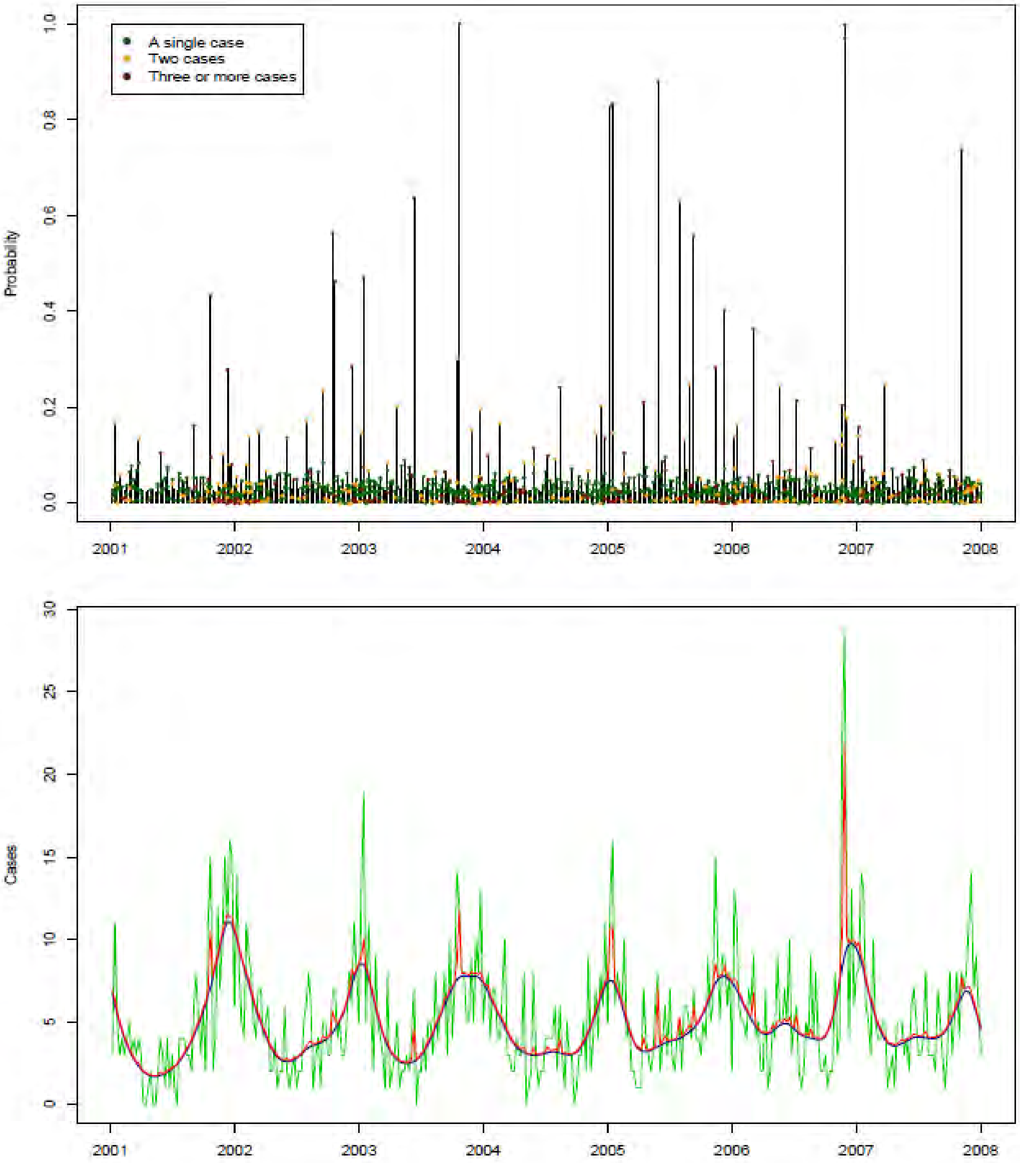


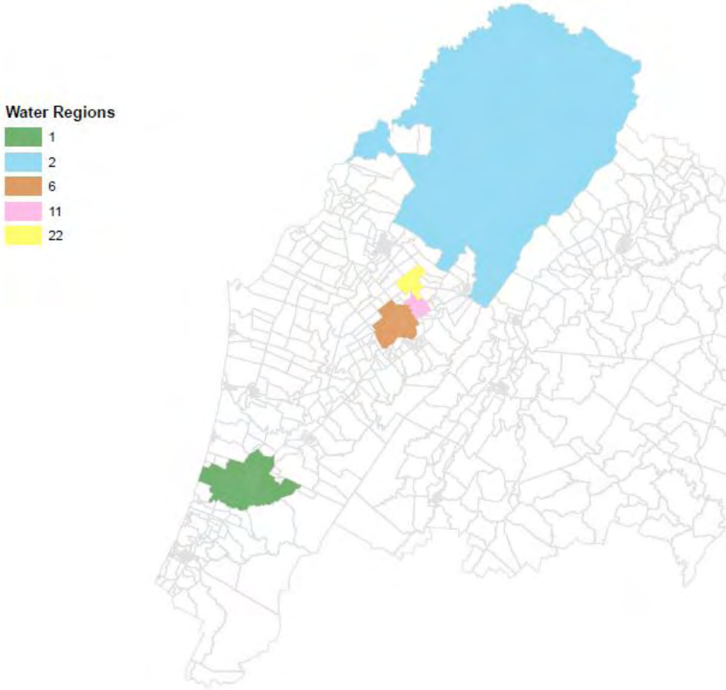
Census Area Unit

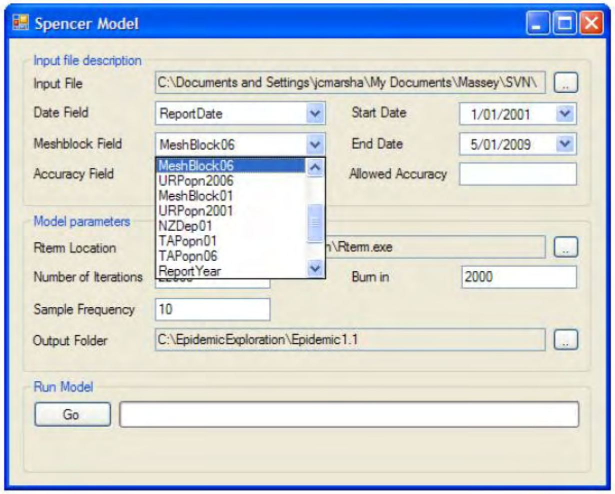


Water supply regions

Identifying potential outbreaks in water supply regions







Input File: C:\Documents and Settings\jcmarsha\My Documents\Massey\SVN\
Date Field: ReportDate
Start Date: 1/01/2001
Meshblock Field: MeshBlock06
End Date: 5/01/2009
Accuracy Field:
Allowed Accuracy:
Rterm Location: \Rterm.exe
Number of Iterations:
Burn in: 2000
Sample Frequency: 10
Output Folder: C:\EpidemicExploration\Epidemic1.1

Known outbreaks

Outbreak	Region(s) Affected	Number of Cases
11 January 2005	1	8
25 May 2005	6	7
15 November 2006	2, 6	8
21 November 2006	4, 6, 11, 22	23

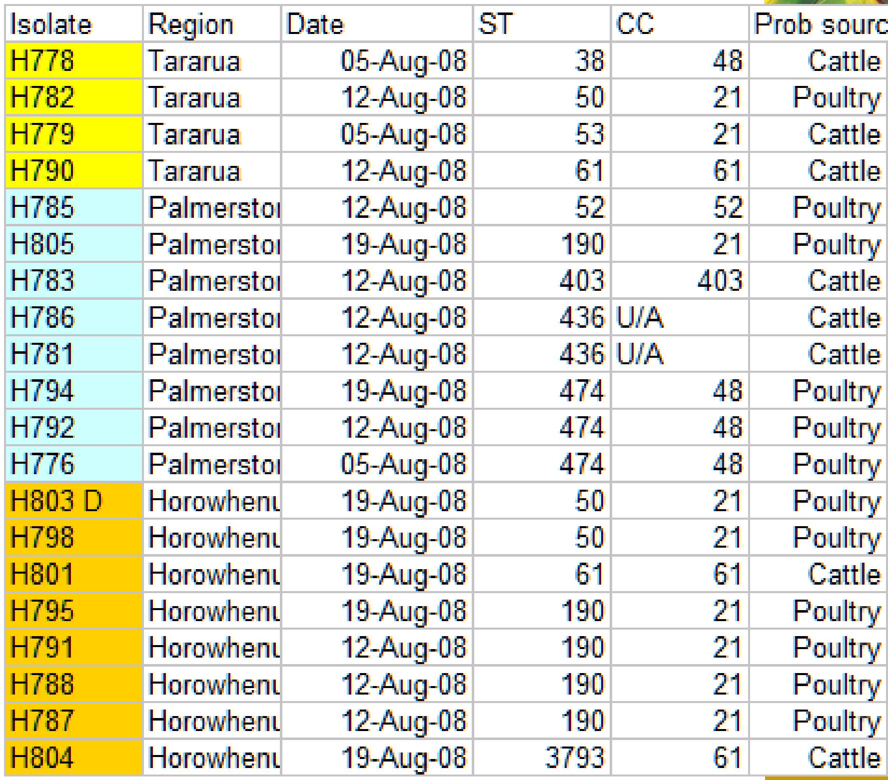
Outbreaks identified by model

Outbreak	Region	Probability	Number of Cases
11 January 2005	1	1.0000	10
25 May 2005	6	0.4785	8
15 November 2006	6	0.948	9
	2	0.620	6
	1*	0.104	2
21 November 2006	6	1.000	26
	11	0.999	6
	22	0.303	2
	1*	0.121	4

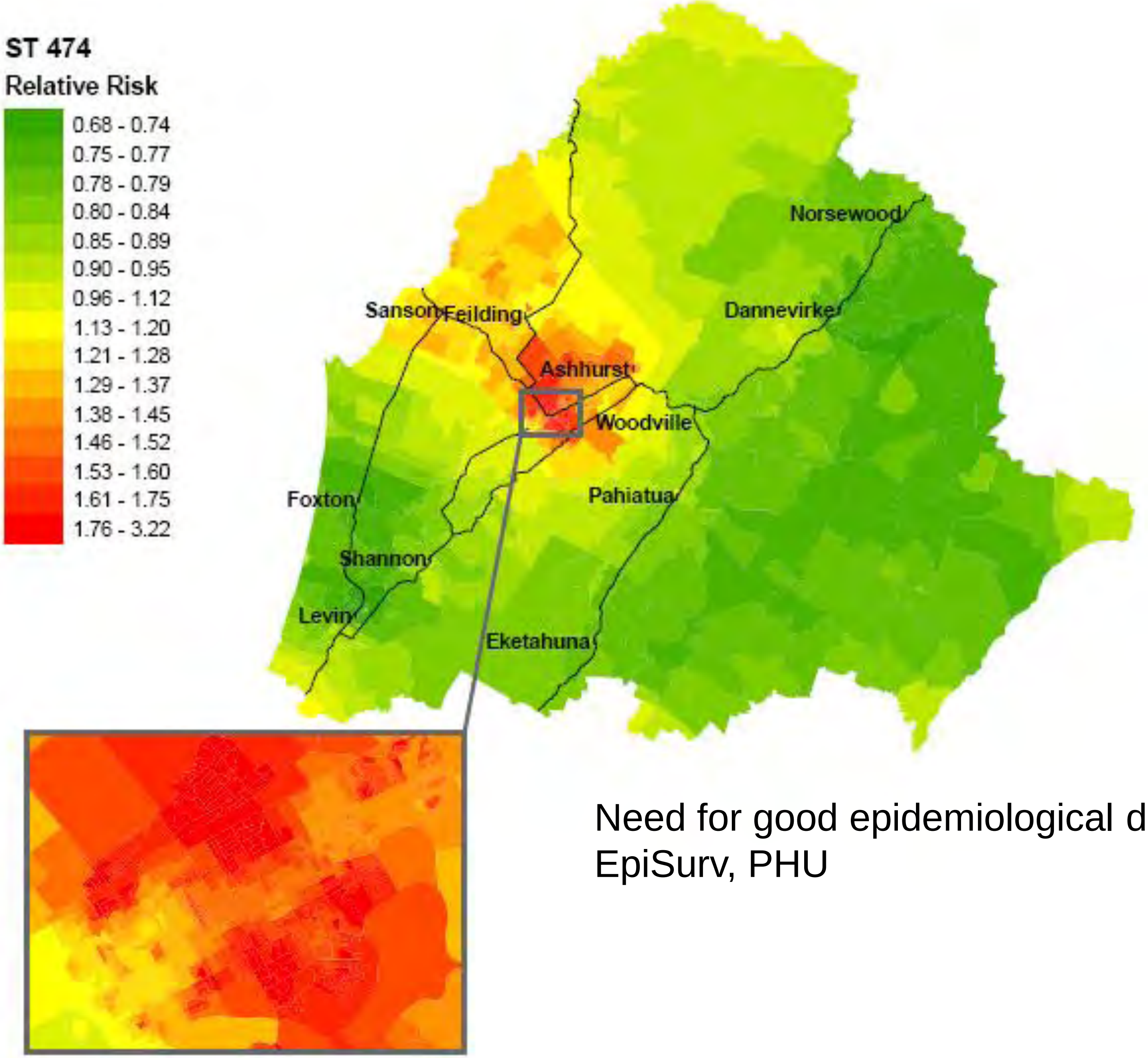
MLST

- Genotype clusters
  - Poultry and cattle associated
- Some use but not really discriminatory enough?
  - Add *fla*A SVR
- Not rapid enough with current resources
  - Not ‘real time’
  - 2-3 weeks

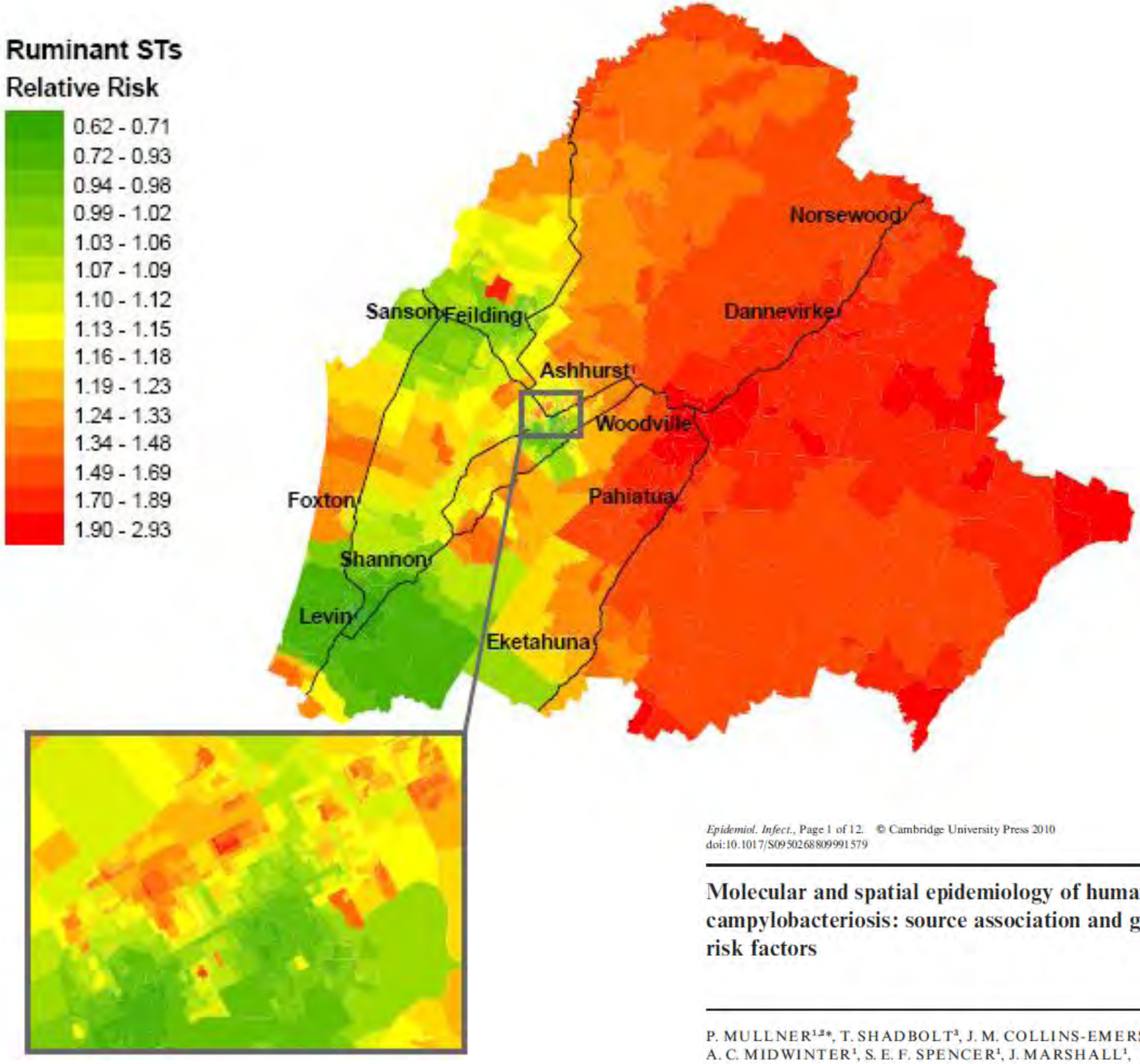
- More useful for understanding epidemiology
  - Spatial models combined with MLST



Isolate	Region	Date	ST	CC	Prob sourc
H778	Tararua	05-Aug-08	38	48	Cattle
H782	Tararua	12-Aug-08	50	21	Poultry
H779	Tararua	05-Aug-08	53	21	Cattle
H790	Tararua	12-Aug-08	61	61	Cattle
H785	Palmersto	12-Aug-08	52	52	Poultry
H805	Palmersto	19-Aug-08	190	21	Poultry
H783	Palmersto	12-Aug-08	403	403	Cattle
H786	Palmersto	12-Aug-08	436	U/A	Cattle
H781	Palmersto	12-Aug-08	436	U/A	Cattle
H794	Palmersto	19-Aug-08	474	48	Poultry
H792	Palmersto	12-Aug-08	474	48	Poultry
H776	Palmersto	05-Aug-08	474	48	Poultry
H803 D	Horowhenu	19-Aug-08	50	21	Poultry
H798	Horowhenu	19-Aug-08	50	21	Poultry
H801	Horowhenu	19-Aug-08	61	61	Cattle
H795	Horowhenu	19-Aug-08	190	21	Poultry
H791	Horowhenu	12-Aug-08	190	21	Poultry
H788	Horowhenu	12-Aug-08	190	21	Poultry
H787	Horowhenu	12-Aug-08	190	21	Poultry
H804	Horowhenu	19-Aug-08	3793	61	Cattle



Need for good epidemiological data
EpiSurv, PHU



Epidemiol. Infect., Page 1 of 12. © Cambridge University Press 2010
doi:10.1017/S0950268809991579

Molecular and spatial epidemiology of human campylobacteriosis: source association and genotype-related risk factors

P. MULLNER[1,2*], T. SHADBOLT[3], J. M. COLLINS-EMERSON[1],
A. C. MIDWINTER[1], S. E. F. SPENCER[1], J. MARSHALL[1],
P. E. CARTER[4], D. M. CAMPBELL[2], D. J. WILSON[5], S. HATHAWAY[2],
R. PIRIE[4] AND N. P. FRENCH[1]

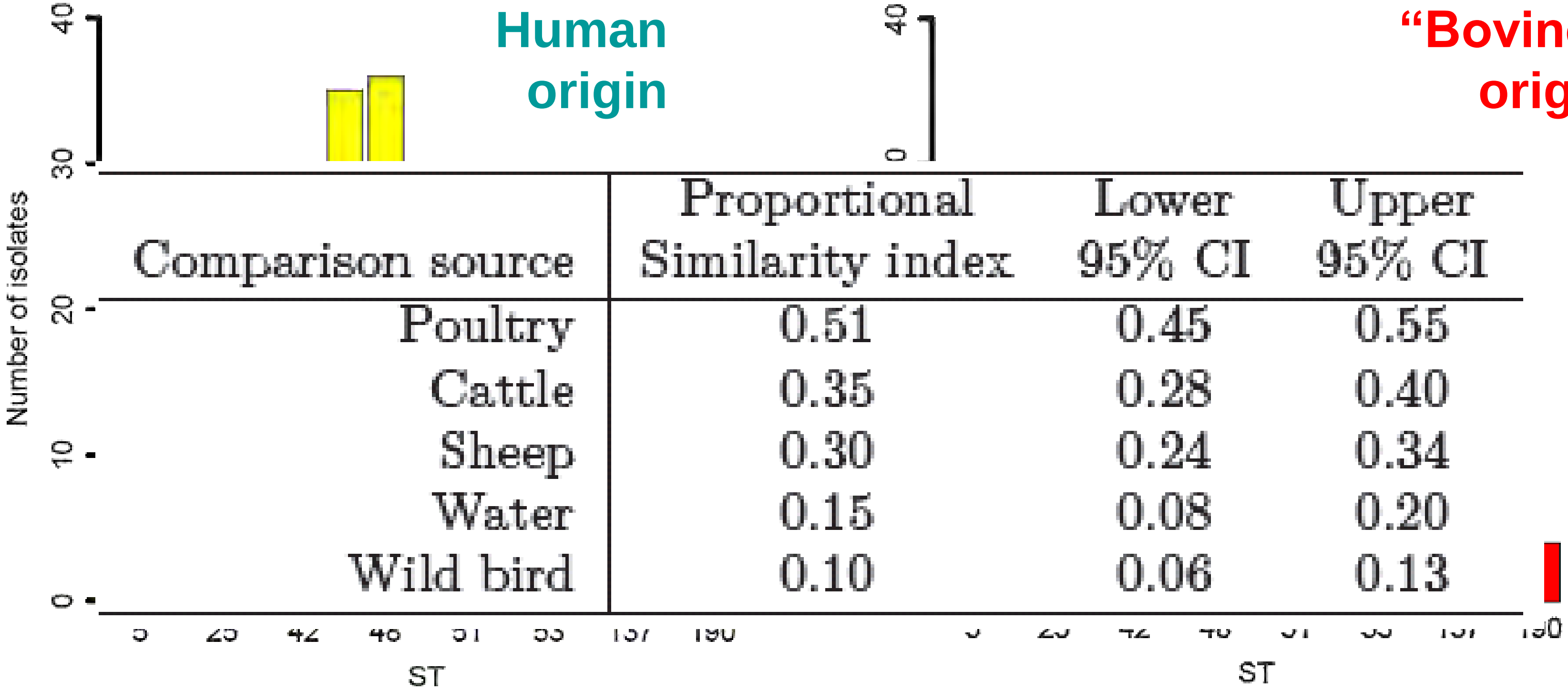
Application of MLST and modelling

<u>Policy based surveillance</u>

- Source attribution
  - Proportional similarity
    - Area of overlap
  - Dutch model
    - Simple deterministic assignment
  - Hald model
    - Bayesian model assignment with uncertainty
  - Island model
    - Population genetics approach

Proportional Similarity Index (PS)

The PS estimates the area of overlap between the frequency distributions of e.g. bacterial sub types from different sources.



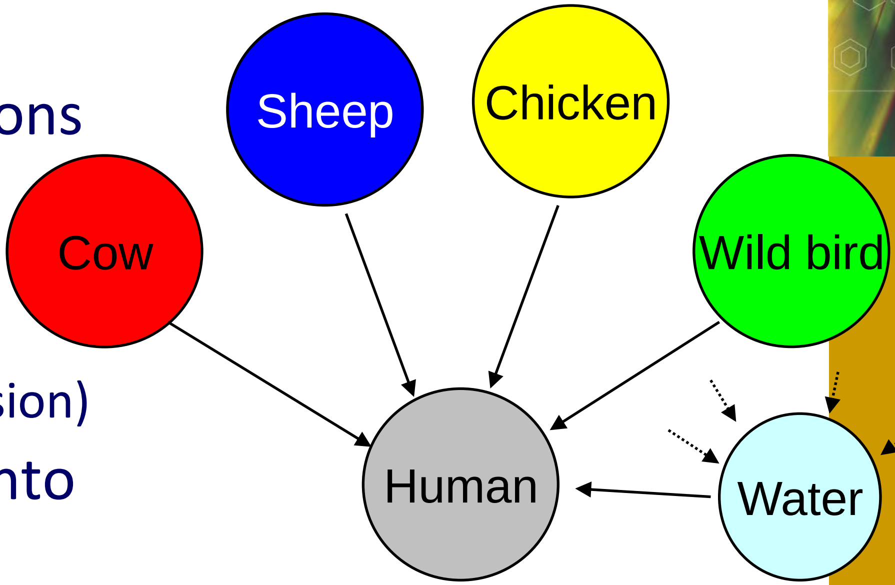
Human origin

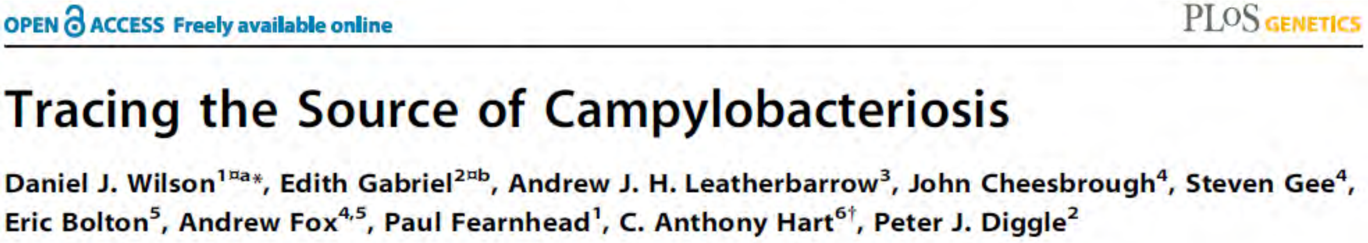
"Bovine" origin

Comparison source	Proportional Similarity index	Lower 95% CI	Upper 95% CI
Poultry	0.51	0.45	0.55
Cattle	0.35	0.28	0.40
Sheep	0.30	0.24	0.34
Water	0.15	0.08	0.20
Wild bird	0.10	0.06	0.13

Island model (Wilson et al 2008)

- Population genetics approach
- Genealogical method based on 'coalescent'
    - Cross-validation
- Use MLST data in animal populations ("islands") to estimate:
    - Mutation rates
    - Recombination rates
    - Migration rates (inter-host transmission)
- From these estimate 'migration' into human population
    - Source attribution



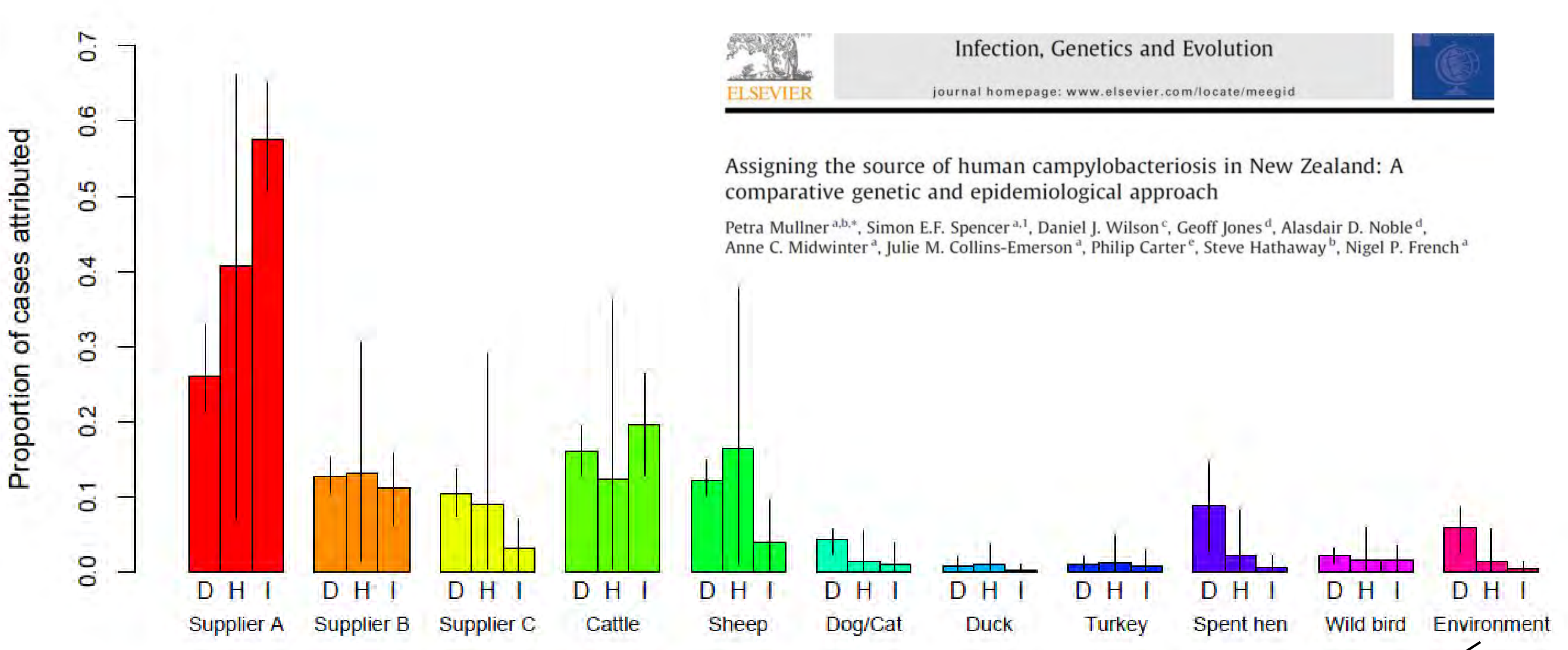

OPEN ACCESS Freely available online

PLoS GENETICS

Tracing the Source of Campylobacteriosis

Daniel J. Wilson[1¤a*], Edith Gabriel[2¤b], Andrew J. H. Leatherbarrow[3], John Cheesbrough[4], Steven Gee[4], Eric Bolton[5], Andrew Fox[4,5], Paul Fearnhead[1], C. Anthony Hart[6†], Peter J. Diggle[2]

Source attribution



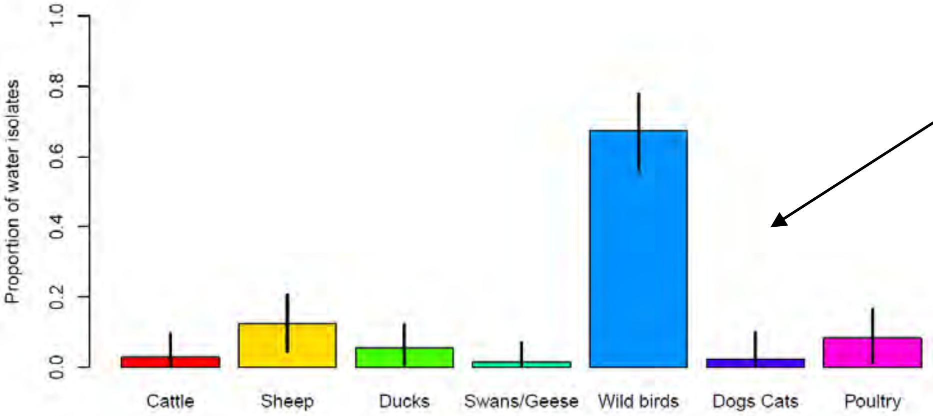
ELSEVIER

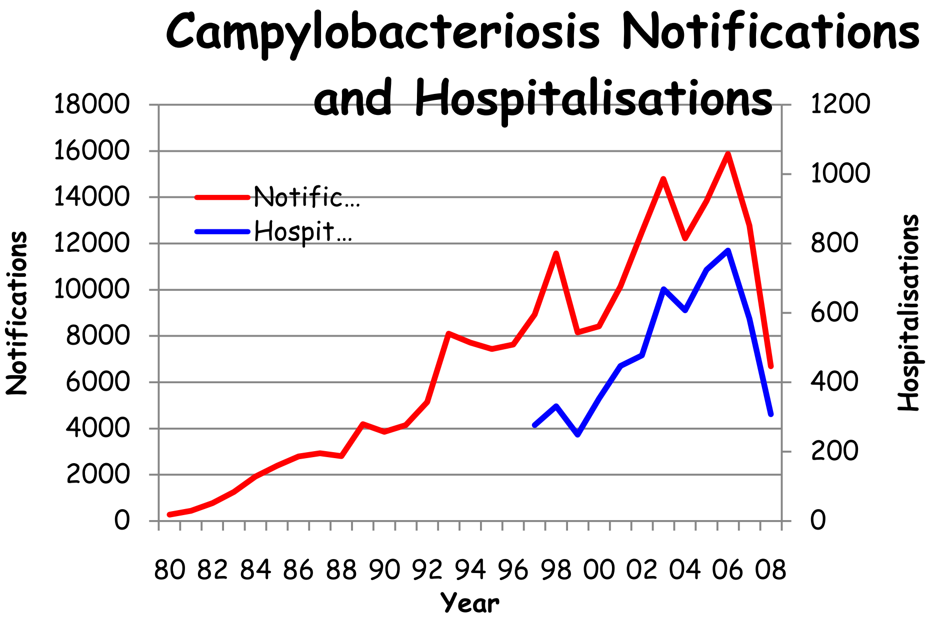
Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Assigning the source of human campylobacteriosis in New Zealand: A comparative genetic and epidemiological approach

Petra Mullner [a,b,*], Simon E.F. Spencer [a,1], Daniel J. Wilson [c], Geoff Jones [d], Alasdair D. Noble [d], Anne C. Midwinter [a], Julie M. Collins-Emerson [a], Philip Carter [e], Steve Hathaway [b], Nigel P. French [a]





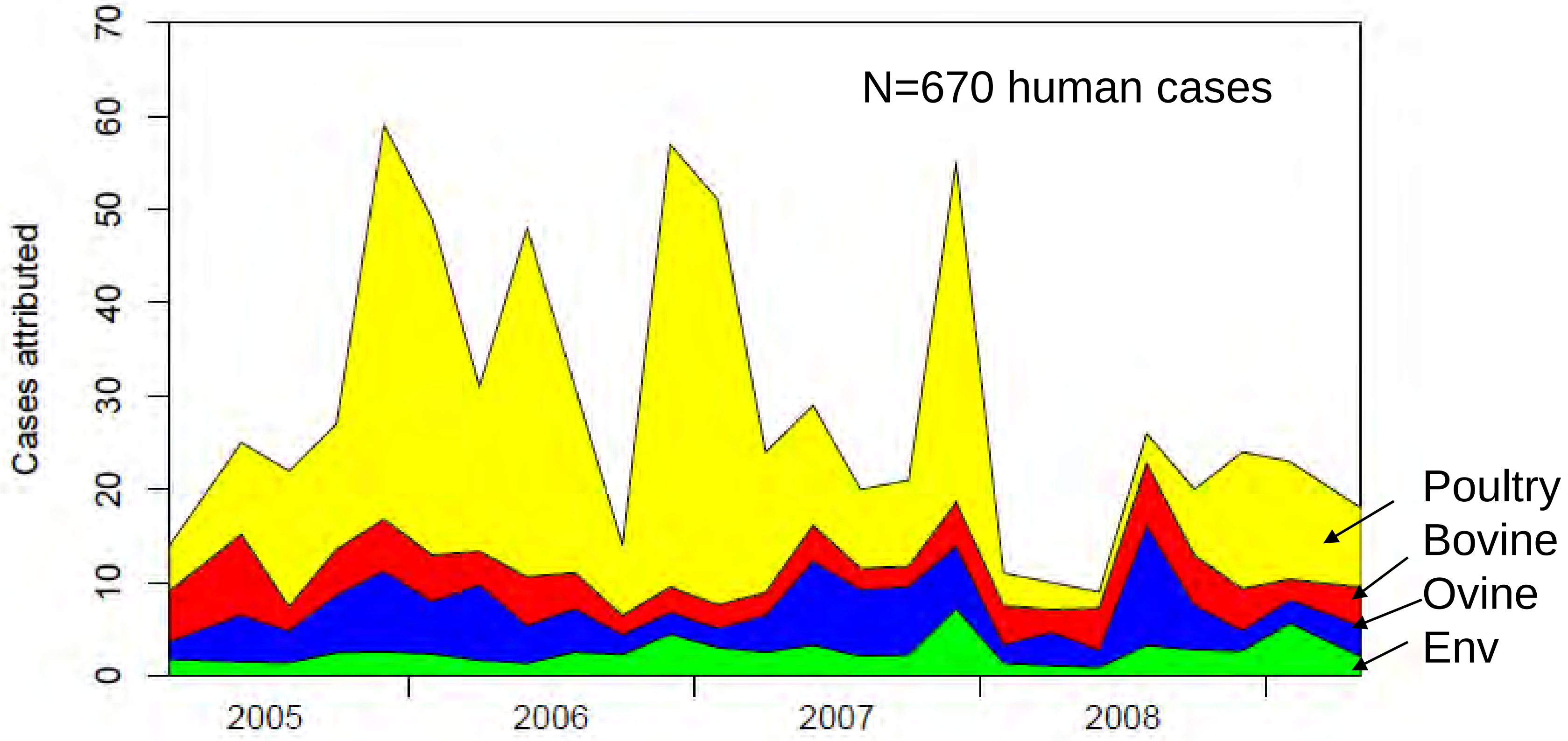
Campylobacteriosis Notifications and Hospitalisations

NEW ZEALAND
nzfsa
FOOD SAFETY AUTHORITY

Campylobacter in Poultry – Risk Management Strategy 2007 - 2010

…2006-2009, now 2008-2011

Modelling post intervention change in attribution



Dynamic Hald model

Conclusions

- Molecular tools advanced in recent years
  - From RFLP - Next generation sequencing
  - Need to consider many factors (discrimination, typeability, portability….)
- Modelling tools – recent advances and applications
- Important element of infectious disease surveillance
  - Event/control and policy/strategy based
  - Many applications
- Multilocus sequence typing played important role in local and national control of campylobacteriosis

Acknowledgements

- Staff – lecturers, RAs
  - Dr Eve Pleydell, Dr Deb Prattley,
  - Dr Julie Collins-Emerson, Dr Anne Midwinter,,
  - Dr Simon Spencer, Dr Jonathan Marshall, Dr Patrick Biggs
- Lab team:
  - Rebecca Pattison, Rukhshana Akhter, Errol Kwan, Lynn Rogers, Isabel Li, Jim Learmonth, Anthony Pita, Sarah Moore, Angie Reynolds
- PhD students
  - Petra Mullner, Vathsala Mohan, Patricia Jaros, Hamid Irshad
- Masters students
  - Particularly Tui Shadbolt, Ann Sears
- MidCentral Public Health
  - Dr Jill McKenzie and team
- MedLab Central
- ESR - Phil Carter, Sharla McTavish, Ruth Pirie
- AgResearch – Grant Hotter, Adrian Cookson
- CDRP team
  - NIWA – Graham McBride
  - ESR – Rob Lake
  - NZFSA – Peter van de Logt, Dr Donald Campbell
- Michael Baker
- Massey – IMBS, IFNHH, IFS (Dr Barbara Holland, Dr Geoff Jones, Dr Alasdair Noble, Prof Martin Hazelton)
  - Allan Wilson Centre
- Universities of Liverpool, Lancaster, Oxford, Chicago
- Undergraduate student
  - Jeffrey Ling
- Industry

NZFSA, MoH, FRST, Royal Society Marsden and Industry funding

Publications

1. Mullner, P, Shadbolt, T, Collins-Emerson, JM, Midwinter, AC, Spencer, S, Marshall, J, Carter,P, Campbell, D, Wilson, DJ, Hathaway, S, Pirie, R, and French, NP. Molecular and spatial epidemiology of human campylobacteriosis - source association and genotype-related risk factors, ***Epidemiology and Infection* 2010** in press .
2. Mullner, P, Collins-Emerson, JM, Midwinter, Philip Carter[3], Simon E. F. Spencer, Peter van der Logt[2], Steve Hathaway[2], Nigel French[1]Molecular epidemiology of *Campylobacter jejuni* in a geographically isolated country with a uniquely structured poultry industry. ***Applied and Environmental Microbiology*, 2010** in press
3. Hotter, G. S. Li I. H., and French N. P.. Binary genomotyping using lipooligosaccharide biosynthesis genes distinguishes between *Campylobacter jejuni* isolates within poultry-associated multilocus sequence type. ***Epidemiology and Infection***, **2010** in press available on-line.
4. Mullner P, Spencer, SEF, Daniel J. Wilson, Geoff Jones, Alasdair D. Noble, Anne C. Midwinter, Julie M. Collins-Emerson, Philip Carter, Steve Hathaway, Nigel P. French. Assigning the source of human campylobacteriosis in New Zealand: A comparative genetic and epidemiological approach. **2009 *Infection, Genetics and Evolution*** 9: 1311-9
5. Carter PE, McTavish SM, Brooks HJ, Campbell D, Collins-Emerson JM, Midwinter AC, French NP. Novel clonal complexes with unknown animal reservoir dominate *Campylobacter jejuni* isolated from river water in New Zealand. ***Applied and Environmental Microbiology***. **2009** 75: 6038-46
6. McTavish SM, Pope CE, Nicol C, Campbell, D, French N and Carter PE Multilocus Sequence Typing of *Campylobacter jejuni*, and the Correlation between Clonal Complex and Pulsed-field Gel Electrophoresis Macrorestriction Profile. **2009 *FEMS Microbiology Letters*** 298(2):149-56
7. French, N.P., Midwinter, A., Holland, B., Collins-Emerson, J., Pattison, R., Colles, F.M., and Carter, P., Molecular epidemiology of *Campylobacter jejuni* isolated from wild bird faecal material in children's playgrounds. ***Applied and Environmental Microbiology***, **2009**. **75**(3): p. 779-83.
8. Mullner, P., Jones, G., Noble, A., Spencer, S., Hathaway, S., and French, N.P., Source Attribution of Food Borne Zoonosis in New Zealand: A Modified Hald Model**. *Risk Analysis***, **2009**. 29(7):970-84**.**
9. Grove-White, D.H. Leatherbarrow,A.J.H., Cripps, P.J. Diggle, P.J. and French N.P. Temporal and farm managment associated variation in the faecal pat prevalence of *Campylobacter jejuni* in ruminants in three zones in Lancashire. ***Epidemiology and Infection***, 2010 in press

- + NZFSA website – science for number of reports