Genomic epidemiology: what is it? and how will it inform public health policy?

Nigel French

Infectious Disease Research Centre, Massey University



Infectious Disease mini-symposium September 2014



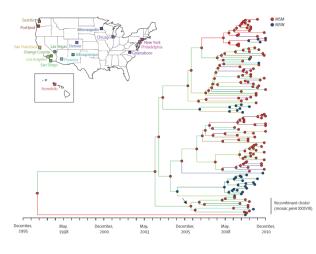
Te Kunenga ki Pūrehuroa





Genomic epidemiology

- Epidemiology informed by whole genome sequencing
 - NGS technology
 - Host and/or pathogen
- Used for communicable and noncommunicable diseases
 - GWAS
- Number of studies increased in last 5-6 years
- Lot of work in progress



Recent articles

Volume 20, Number 9-September 2014

Research

EMERGING INFECTIOUS DISEASES®

Genomic Epidemiology of *Salmonella enterica* Serotype Enteritidis based on Population Structure of Prevalent Lineages

Xiangyu Deng⊠, Prerak T. Desai, Henk C. den Bakker, Matthew Mikoleit, Beth Tolar, Eija Trees, Rene S. Hendriksen, Jonathan G. Frye, Steffen Porwollik, Bart C. Weimer, Martin Wiedmann, George M. Weinstock, Patricia I. Fields¹, and Michael McClelland¹

OPEN ACCESS Freely available online



Whole Genome Sequencing Reveals Local Transmission Patterns of *Mycobacterium bovis* in Sympatric Cattle and Badger Populations

Roman Biek¹⁹, Anthony O'Hare¹⁹, David Wright^{2,3}, Tom Mallon², Carl McCormick², Richard J. Orton¹, Stanley McDowell², Hannah Trewby¹, Robin A. Skuce², Rowland R. Kao¹*

Population genomics of post-vaccine changes in pneumococcal epidemiology

Nicholas J Croucher^{1,2}, Jonathan A Finkelstein^{3,4}, Stephen I Pelton⁵, Patrick K Mitchell¹, Grace M Lee^{3,6,7}, Julian Parkhill², Stephen D Bentley^{2,8,10}, William P Hanage^{1,10} & Marc Lipsitch^{1,9,10}



Recent articles



Genomic epidemiology of the *Escherichia coli* O104:H4 outbreaks in Europe, 2011 www.pnas.org/cgi/doi/10.1073/pnas.1121491109

Yonatan H. Grad^{a,b}, Marc Lipsitch^{b,c}, Michael Feldgarden^d, Harindra M. Arachchi^d, Gustavo C. Cerqueira^d, Michael FitzGerald^d, Paul Godfrey^d, Brian J. Haas^d, Cheryl I. Murphy^d, Carsten Russ^d, Sean Sykes^d, Bruce J. Walker^d, Jennifer R. Wortman^d, Sarah Young^d, Qiandong Zeng^d, Amr Abouelleil^d, James Bochicchio^d, Sara Chauvin^d, Timothy DeSmet^d, Sharvari Gujja^d, Caryn McCowan^d, Anna Montmayeur^d, Scott Steelman^d, Jakob Frimodt-Møller^{e,f}, Andreas M. Petersen^{f,g}, Carsten Struve^f, Karen A. Krogfelt^f, Edouard Bingen^{h,i}, François-Xavier Weill^j, Eric S. Lander^{d,k,l,1}, Chad Nusbaum^d, Bruce W. Birren^d, Deborah T. Hung^{a,d,m,n,2}, and William P. Hanage^{b,1,2}

Genomic epidemiology of *Neisseria gonorrhoeae* with reduced susceptibility to cefixime in the USA: a retrospective observational study

Lancet Infect Dis 2014;

14: 220-26

Yonatan H Grad, Robert D Kirkcaldy, David Trees, Janina Dordel, Simon R Harris, Edward Goldstein, Hillard Weinstock, Julian Parkhill, William P Hanage, Stephen Bentley, Marc Lipsitch

Distinguishable Epidemics of Multidrug-Resistant *Salmonella* Typhimurium DT104 in Different Hosts

27 SEPTEMBER 2013 VOL 341 SCIENCE

A. E. Mather, ¹ S. W. J. Reid, ² ‡ D. J. Maskell, ³ J. Parkhill, ¹ M. C. Fookes, ¹ S. R. Harris, ¹ D. J. Brown, ⁴ J. E. Coia, ⁴ M. R. Mulvey, ⁵ M. W. Gilmour, ⁵* L. Petrovska, ⁶ E. de Pinna, ⁷ M. Kuroda, ⁸ M. Akiba, ⁹ H. Izumiya, ¹⁰ T. R. Connor, ¹† M. A. Suchard, ¹¹ P. Lemey, ¹² D. J. Mellor, ¹³ D. T. Haydon, ¹³ N. R. Thomson ¹‡

MASSEY UNIVERSITY

Recent articles: review and commentary

Evolutionary epidemiology: preparing for an age of genomic plenty

Phil Trans R Soc B 368: 20120193.

http://dx.doi.org/10.1098/rstb.2012.0193

O. G. Pybus¹, C. Fraser² and A. Rambaut³



Bacterial genomes in epidemiology—present and future

Nicholas J. Croucher¹, Simon R. Harris², Yonatan H. Grad^{1,3} and William P. Hanage¹

Phil Trans R Soc B 368: 20120202. http://dx.doi.org/10.1098/rstb.2012.0202



Applications

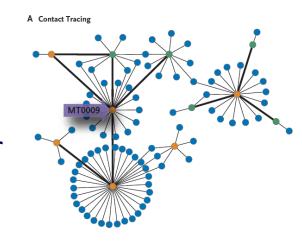
- Transmission
 - Estimating Ro
 - Transmission in social networks
 - Anthroponoses, zoonoses (incl reverse zoonoses)
- Host switching
 - Incl. Zoonoses
- Phenotype:genotype relationships
 - Broad definition of phenotype (host association, seasonality, carbon source utilisation)
- Geographical movements and incursion
 - Number and origin of introductions
- Evolution
 - Impact of interventions (vaccination)
 - Changes in virulence
 - Antimicrobial resistance (Debbie Williamson)



Whole-Genome Sequencing and Social-Network Analysis of a Tuberculosis Outbreak

Gardy et al, 2011 New Engl J Med, 364, 730-739

- 3 year outbreak of TB in British Colombia
 - In community struggling with alcoholism, drug abuse and transient housing status
- Typed isolates using mycobacterial interspersed repetitive unit—variable-number tandem-repeat (MIRU-VNTR)
 - Evidence that outbreak was due to a single clone
- Traditional contact tracing could not identify a source





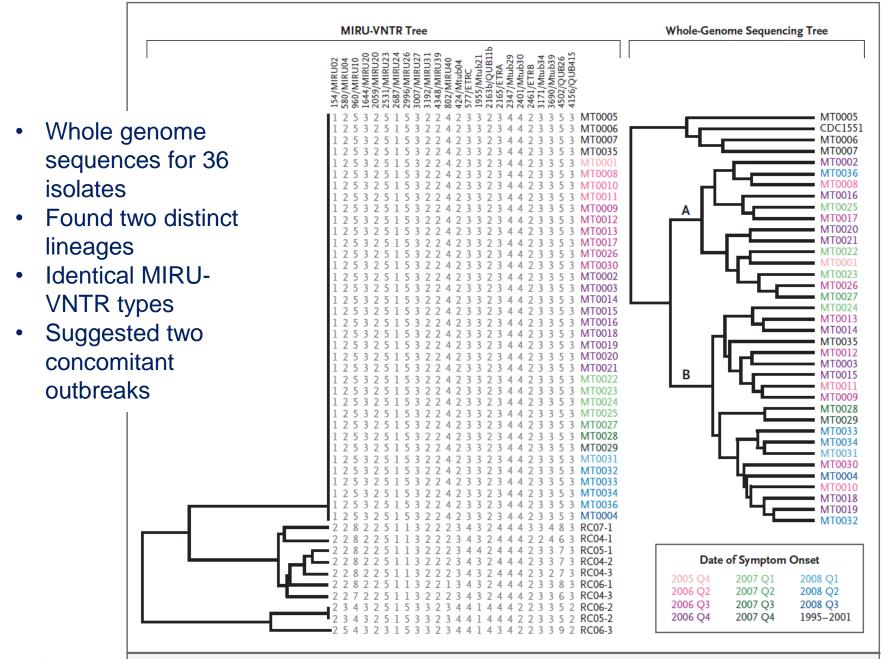


Figure 3. Hierarchical Clustering of Patterns of Mycobacterial Interspersed Repetitive Unit-Variable Number Tandem Repeats (MIRU-VNTRs) and Whole-Genome-Derived Genetic Distances of 36 Mycobacterium tuberculosis Isolates.

Social network analysis

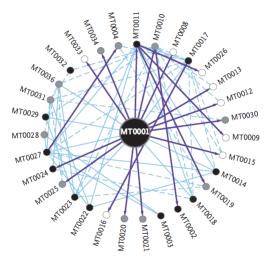
- Interviewed cases and constructed a social network to look for key:
 - Persons
 - Places
 - Behaviours
- Highlighted the most probable source case



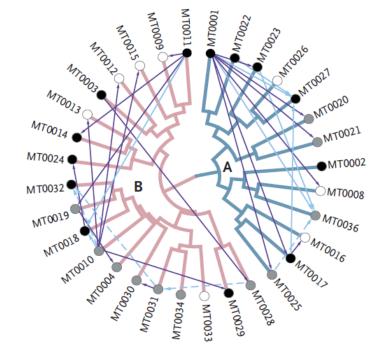
Combining phylogeny and SNA

- Brought phylogeny and social network together
- MT001 = most probable source for first outbreak
 - "superspreader"
 - Most cases
 unconnected with
 each other, but had
 contact with MT001





B Whole-Genome Sequencing and Social-Network Analysis



Truth and error (Croucher et al)

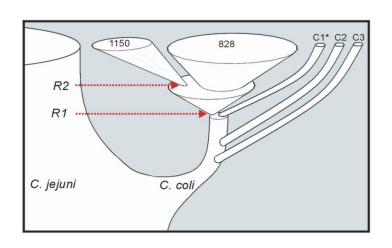
Table 1. Summary of the processes that generate variation in genomic alignments in epidemiology. It should be noted that of these, point mutations inherited clonally should form the basis of any phylogenetic analysis. See Tim Vaughan's talk

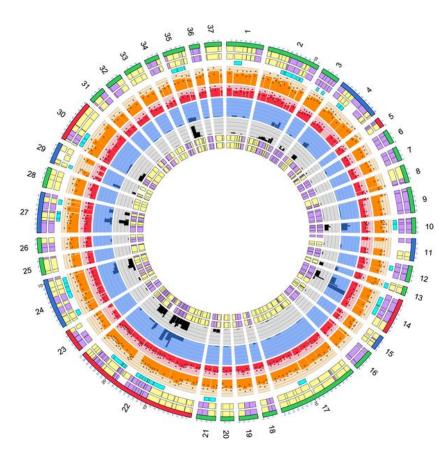
	processes
true variation	point mutation
	insertions and deletions (tandem repeats/homopolymers, etc)
	homologous recombination
	mobile genetic elements
false-positives	SNPs generated <i>in vitro</i>
	mapping errors (frequently associated with repeat sequences)
	systematic sequencing errors
	'ectopic' mapping—reads mapped to paralogous rather than homologous loci in reference.
	poor filters
false-negatives	failure to assemble repetitive regions
	systematic sequencing errors
	poor filters
	distantly related reference sequence, such that variants cannot be called throughout the genome

Phil Trans R Soc B 368: 20120202. http://dx.doi.org/10.1098/rstb.2012.0202



Genomic epidemiology of Campylobacter spp. in New Zealand





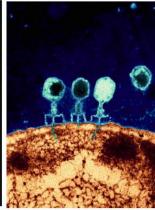
~ 400 genomes in NZ collection

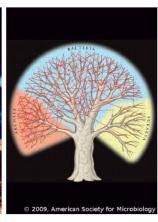


Campylobacter: Mechanisms of evolution

- NZs dominant enteric bacterial pathogen (8,000 cases)
- Mutation and recombination (HGT)
- Evidence from multiple studies indicate recombination <u>much</u> more important than mutation for *C. jejuni*
 - Natural transformation
 - Transduction
- See Tim Vaughan's talk







Conjugation

Transduction

Network of Life

Estimating the Relative Roles of Recombination and Point **Mutation in the Generation of Single Locus Variants** in Campylobacter jejuni and Campylobacter coli

Shoukai Yu · Paul Fearnhead · Barbara R. Holland · Patrick Biggs · Martin Maiden · Nigel French

J Mol Evol (2012) 74:273-280 DOI 10.1007/s00239-012-9505-4

Analysis of Recombination in Campylobacter jejuni from MLST Population Data

Paul Fearnhead, Nick G.C. Smith, Mishele Barrigas, Andrew Fox, Nigel French DOI: 10.1007/s00239-004-0316-0

J Mol Evol (2005) 61:333-340

ESTIMATING THE RELATIVE RATE OF RECOMBINATION TO MUTATION IN BACTERIA FROM SINGLE-LOCUS VARIANTS USING COMPOSITE LIKELIHOOD METHODS

By Paul Fearnhead* Shoukai Yu[†] Patrick Biggs[†] Barbara Holland[‡] and Nigel French[†]

AOAS, resubmitted



Ratio of recombination to mutation (events)

Bacteria	Estimate of λ	95% CI	
H. influenzae	4.9	(3.3,7.4)	
S. aureus	1.4	(0.92,2.1)	
S. uberis	11	(4.8,180)	
C. jejuni	3.4	(2.9,4.1)	←
C. coli	0.43	(0.21, 0.88)	

Table 6

Estimate of common λ across MLST loci, together with putative 95% confidence intervals.

From Fearnhead et al 2014
Based on Single locus MLST variants – i.e. varied at one of 7 housekeeping genes

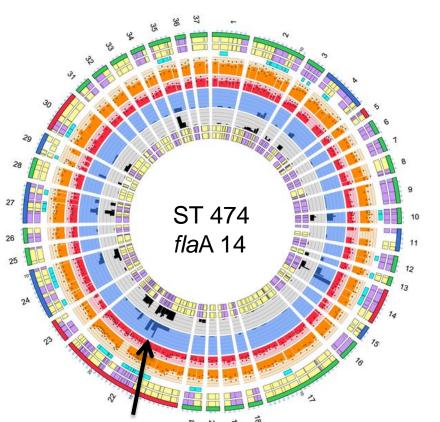




Whole-Genome Comparison of Two *Campylobacter jejuni* Isolates of the Same Sequence Type Reveals Multiple Loci of Different Ancestral Lineage

Patrick J. Biggs¹*, Paul Fearnhead², Grant Hotter³, Vathsala Mohan¹, Julie Collins-Emerson¹, Errol Kwan¹, Thomas E. Besser⁴, Adrian Cookson⁵, Philip E. Carter⁶, Nigel P. French¹





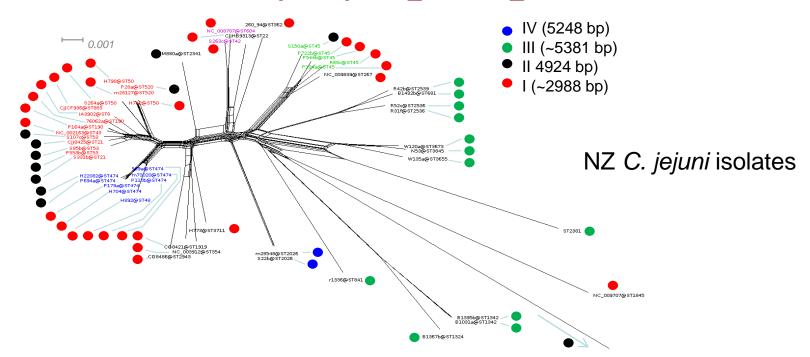
Non-homologous recombination, ykgC

ST 474 evolution

- Compared just 2 isolates: Human and poultry isolates, same genotype, time and space.
- Epidemic strain
- 83 genes differed, 55 with amino acid differences.
- 96.7% were imported via recombination.
- Recombination much more important than mutation for generating divergence.
- Imported from multiple lineages since common ancestor
- MRCA ~ 120-130 years ago, recent population growth



Host association and reservoir attribution which animals are people getting infected from?



- Whole genome sequencing to identify markers associated with host
 - ykgC oxidoreductase ruminant associated?
 - Cj1069-rpsF region
 - Sheppard et al vit B5 synthesis, GWAS study (PNAS 2013)



Mutation and the short-term evolution of

C. jejuni in vivo (Anja Friedrich)



- Chicken feed trial experiment
 - Caprylic acid
- Inoculated with C. jejuni
 - ST-474 and ST-45 strains

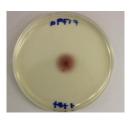


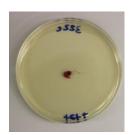
- Isolates recovered on days 33, 39 and 43
- 27 genome sequenced, + 2 parents

Mutation and the short-term evolution of *C. jejuni in vivo* (Anja Friedrich)

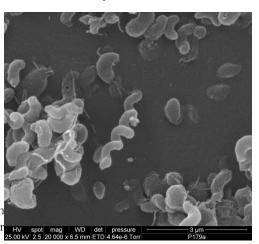
- Only ST-474 recovered
 - 15 core SNPs/point mutations (kSNP*) confirmed
- 6 independent mutations in one gene
 - mreB, actin homologue involved in cell shape
 - No obvious cell shape changes?
- Other mutations in motility genes
 - ptmG flagellar glycosylation gene
 - Motility accessory factor

*Gardner and Hall 2013

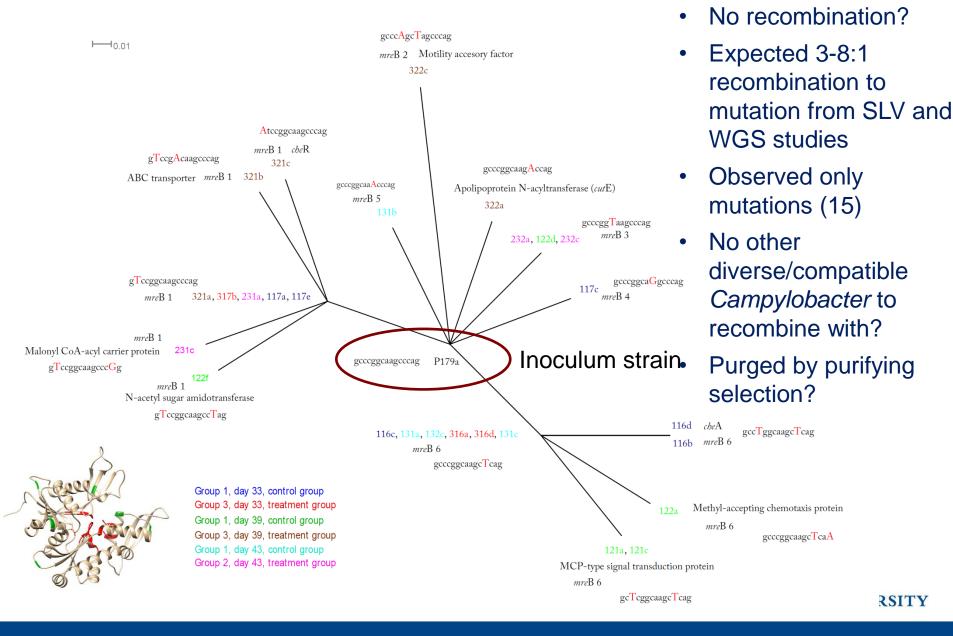








Multiple mutations in mreB



Parallel evolution? (same SNPs in Oxford, UK human cases)





Chicken

Oxford

cases

trial

Global view: Campylobacter niche specialism

- Evidence of generalists and specialists
 - Agricultural strains
 - Wildlife strains
- C. jejuni just 4,000 years old? Traced back to common ancestor
- Species convergence / introgression
 - Associated with livestock strains
 - Sheppard et al *Science* 2008:
 Vol. 320 pp. 237-239



Hybrid speciation in agricultural Campylobacter

OLECULAR ECOLOGY

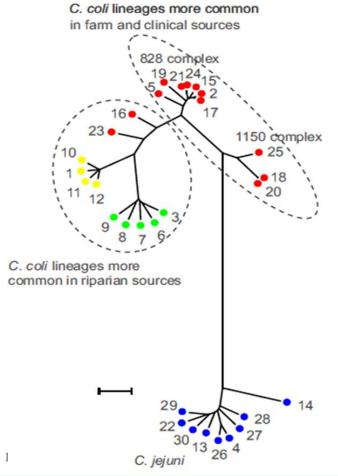
Molecular Ecology (2013) 22, 1051-1064

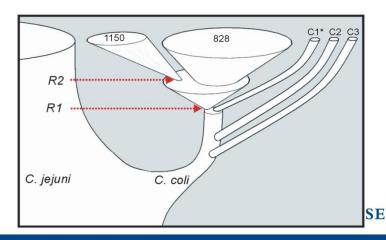
doi: 10.1111/mec.12162

Sheppard et al 2011

Progressive genome-wide introgression in agricultural Campylobacter coli

- Hybridization facilitates niche adaptation in diverse taxa.
- Introgression mainly jejuni coli





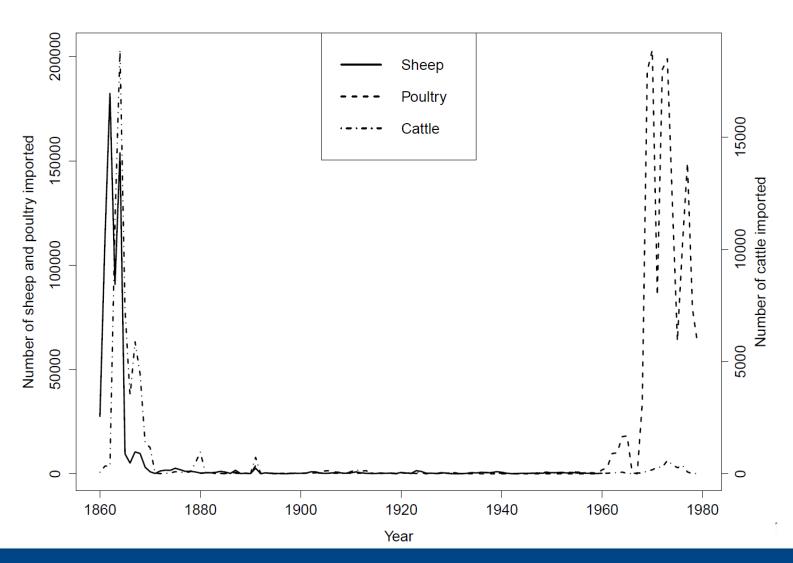


Livestock populations in NZ

Barbara Binney

Binney et al, New Zealand Vet Journal 2014





Forever discovering

Tracing genealogical histories...

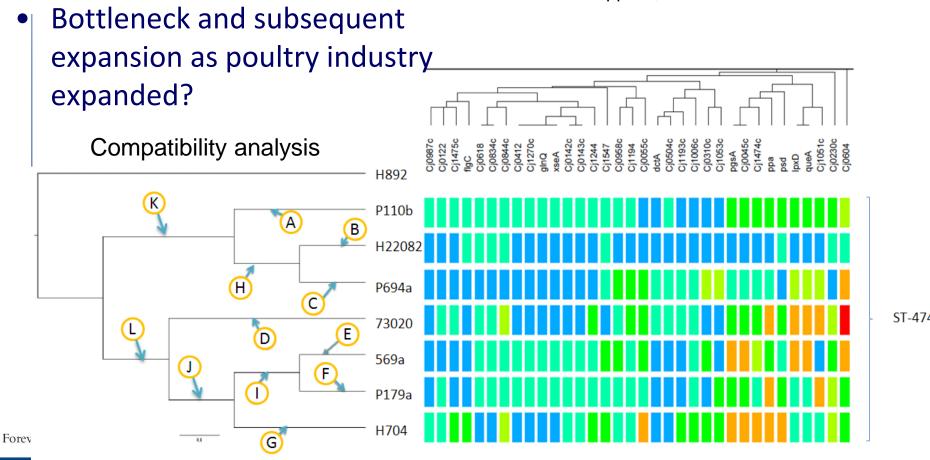
 Epidemic Campylobacter jeuni strain ST 474 - common ancestor introduced in ruminants early 1800s?

Evolution of *Campylobacter* Species in New Zealand

17

Nigel French, Shoukai Yu, Patrick Biggs, Barbara Holland, Paul Fearnhead, Barbara Binney, Andrew Fox, Dai Grove-White, Jessica W. Leigh, William Miller, Petra Muellner and Philip Carter

In Campylobacter Ecology and Evolution. Eds Sheppard, Méric



Novel Clonal Complexes with an Unknown Animal Reservoir Dominate *Campylobacter jejuni* Isolates from River Water in New Zealand[∇]

P. E. Carter, ¹* S. M. McTavish, ¹† H. J. L. Brooks, ² D. Campbell, ³ J. M. Collins-Emerson, ⁴ A. C. Midwinter, ⁴ and N. P. French ⁴

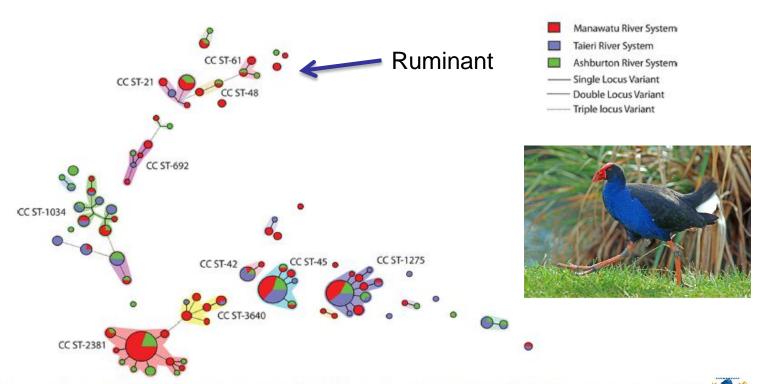
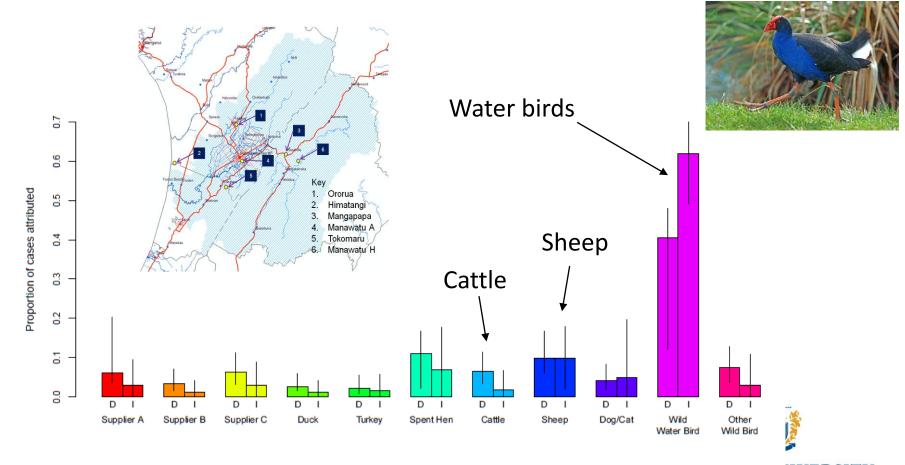


FIG. 1. Minimum spanning tree of New Zealand water isolates based on sequence types. Clusters are identified by the clonal complex number sociated with the sequence types in the cluster.

Te Kunenga ki Pūrehuroa

Campylobacter jejuni sources in water

Most Campylobacter in water in Manawatu from wildlife – even in dairy catchments



Pathogen prediction (Barbara Binney)

Whole genome sequencing ~ 400 NZ isolates



	MLST	Pathogenic Families Matched	prediction score	predicted to be human pathogen	
W194b	ST2381	40	79.701	yes	Water rail associated
N3d	ST2381	41	84.042	yes	Water rail associated
W83a	ST2381	44	97.693	yes	Water rail associated
N191	ST42	428	858.871	yes	Cattle associated
S263a	ST42	430	865.367	yes	Cattle associated
N31	ST42	442	889.374	yes	Cattle associated

Some strains of *C. jejuni* less pathogenic? Consistent with epidemiology

October 2013 | Volume 8 | Issue 10 | e77302



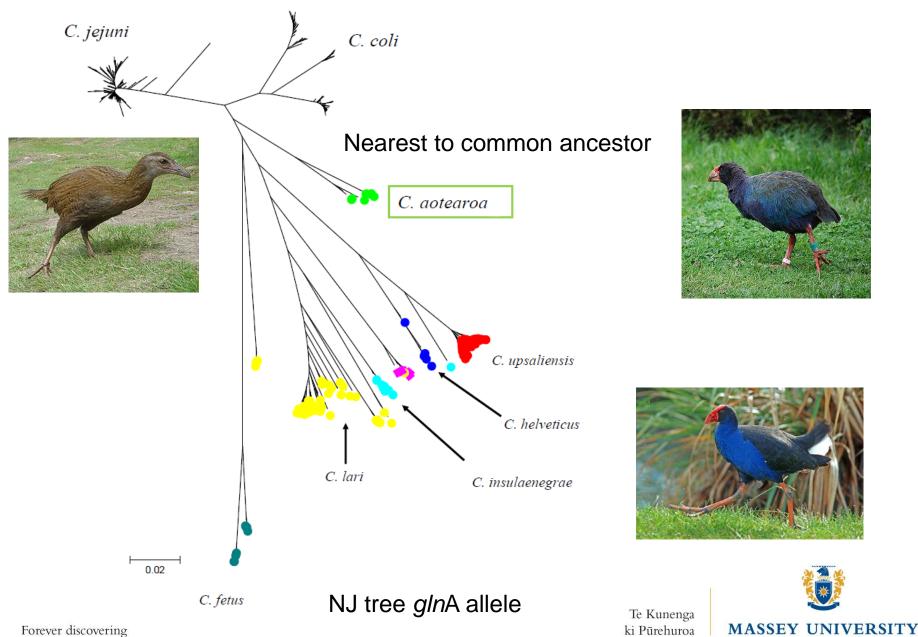


PathogenFinder - Distinguishing Friend from Foe Using Bacterial Whole Genome Sequence Data

Salvatore Cosentino^{1*}, Mette Voldby Larsen¹, Frank Møller Aarestrup², Ole Lund¹

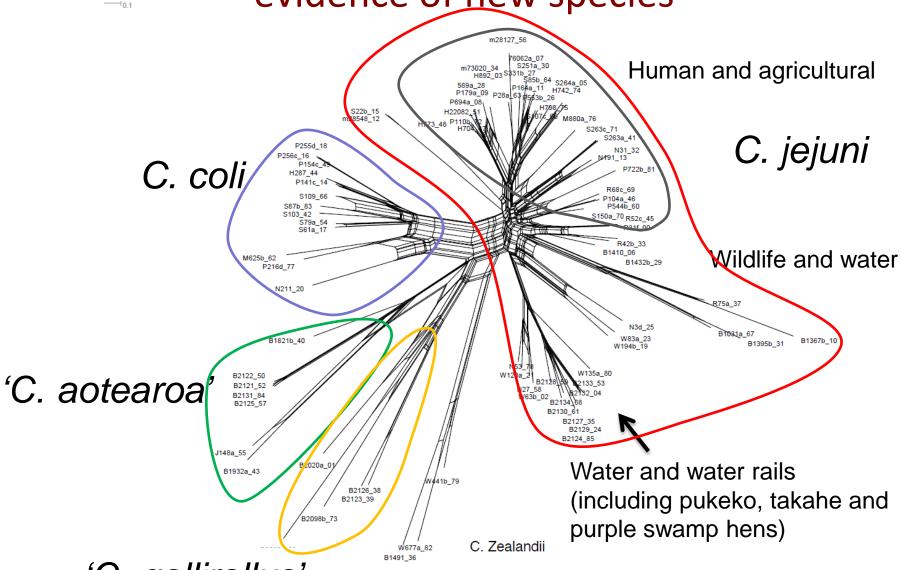
1 Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Kgs. Lyngby, Denmark, 2 National Food Institute, Technical University of Denmark, Kgs. Lyngby, Denmark

Multiple water rail associated lineages in New Zealand



Pan-genome analysis NZ Campylobacter:





'C. gallirallus'

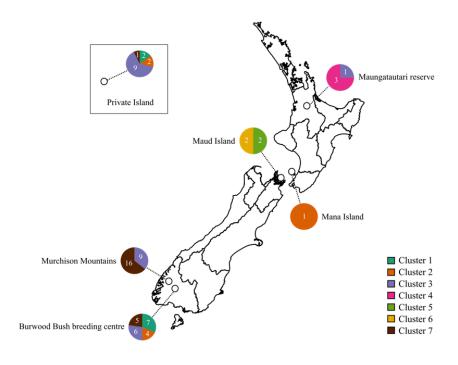
ki Pūrehuroa

MASSEY UNIVERSITY

Takahē Campylobacter study

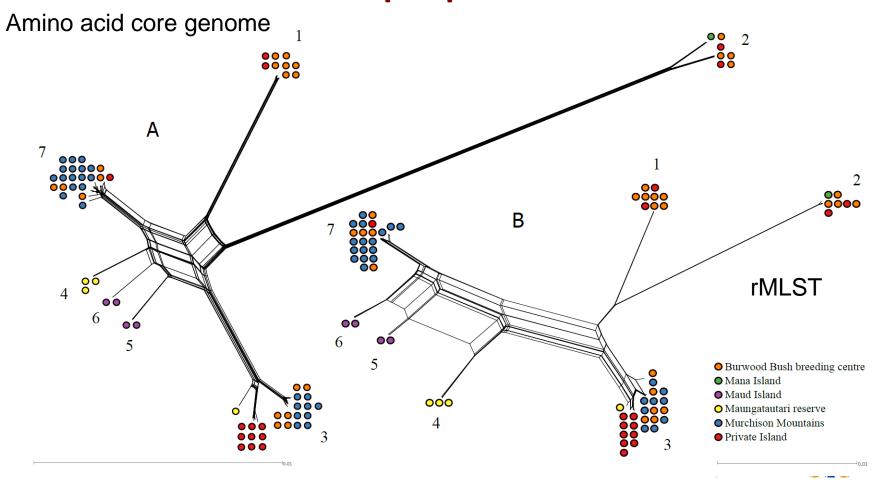


- Aim: Investigate the influence of geographic isolation and translocation on the molecular differentiation of commensal *Campylobacter*
- Faecal samples and swabs
 118 takahē
 - Nov 2011 Apr 2013
- 6 locations in NZ





C. Aotearoa 70 genome: location and population structure



Strong evidence of spatial structuring of genotypes according to translocation site Implications for translocation as a conservation tool

Salmonellosis: multiple serotypes, complex epidemiology

• In NZ salmonellosis is the second most notified foodborne disease after campylobacteriosis.

Table 3.1: Total number of cases from January 2000 to December 2009 for the ten most prevalent serotypes

	-	
Serotype	Number of Cases	
S. Typhimurium DT160	2592	
S. Typhimurium DT1	1010	
S. Typhimurium DT135	844	
S. Brandenburg	734	
S. Typhimurium DT156	705	Niama tumba ialah
S. Infantis	657	Non-typhoidal
S. Typhimurium DT101	570	>15,000 cases
S. Enteritidis phage type 9a	544	
S. Typhimurium DT42	334	
S. Saint Paul	310	_

From French et al Development and application of new tools for the analysis of Salmonella surveillance data... Final report, SCIG-MAS-001 April 2011

Te Kunenga ki Pūrehuroa



Typhimurium DT 160 most prevalent serotype until 2011

- In 1998, Salmonella DT160 identified as a human pathogen in NZ
- Initial outbreak in Christchurch in humans and sparrows (2000)
- Subsequently isolated from farmed animals and wild birds.
- After 2000, increased significantly progressing from the South to the North Island.
- Now 2nd behind Typhimurium RDNC-May 06.





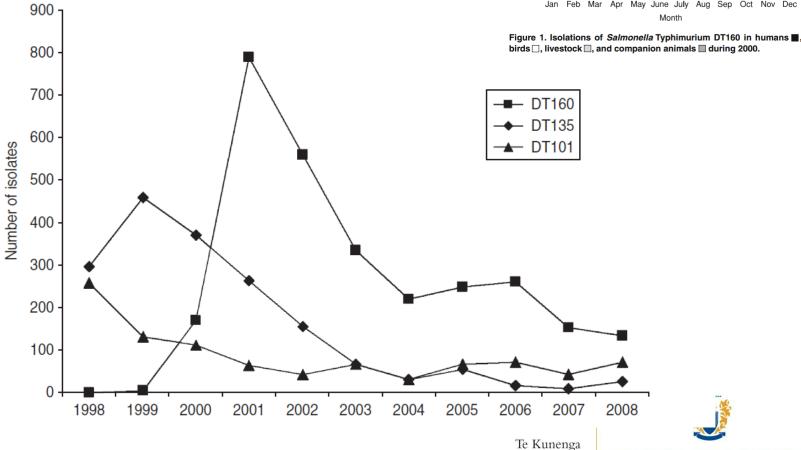
DT160 epidemiology

The evolution and distribution of phage ST160 within Salmonella enterica serotype Typhimurium

M. Price-Carter and others

Forever discovering

Epidemiology and Infection, 139(8):1262-71, 2011



60

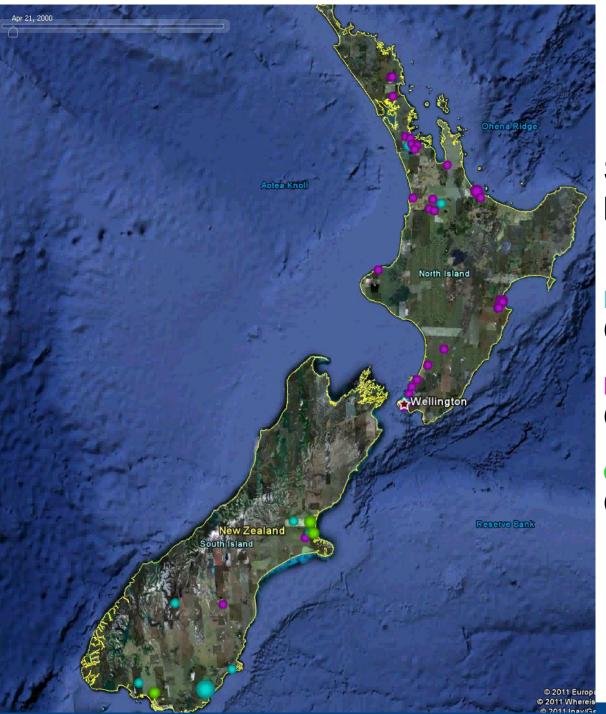
20 10

ki Pūrehuroa

Month

MASSEY UNIVERSITY

of. 50 40



Cases2kml (Chris Jewell)

Salmonellosis in humans 2000-2009

Blue=Brandenburg (sheep, direct contact)

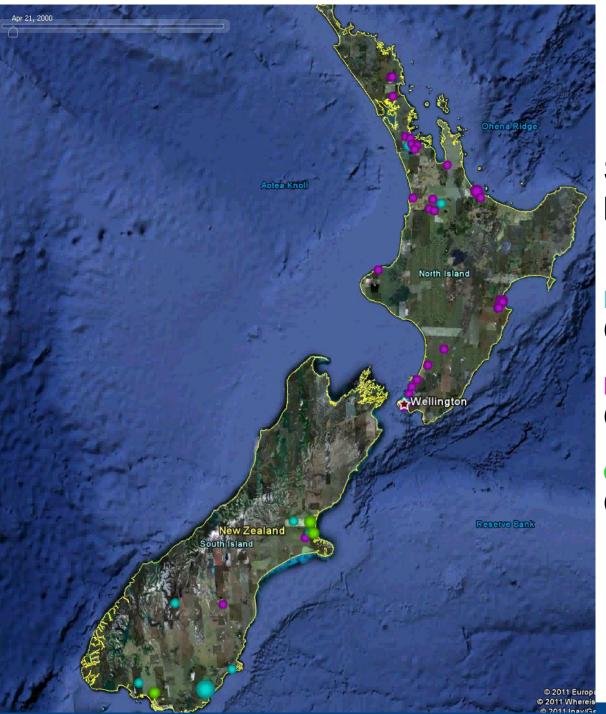
Purple=Typhimurium 156 (food, cattle)

Green=Typhimurium 160 (wild birds, food)



Te Kunenga ki Pūrehuroa





Salmonellosis in humans 2000-2009

Blue=Brandenburg (sheep, direct contact)

Purple=Typhimurium 156 (food, cattle)

Green=Typhimurium 160 (wild birds, food)

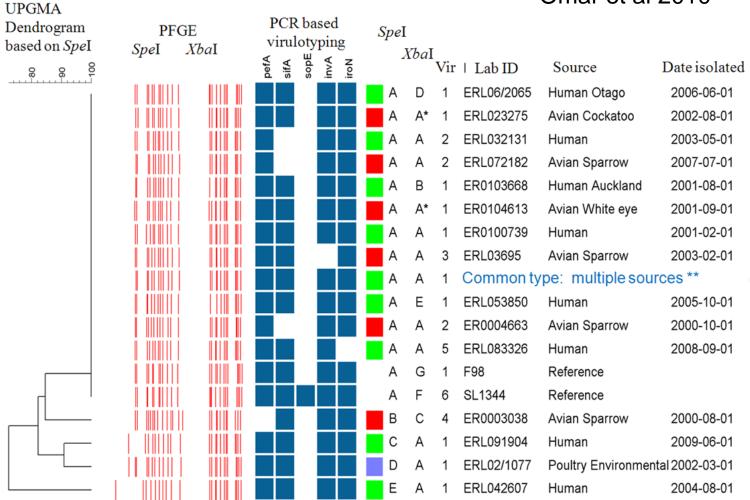




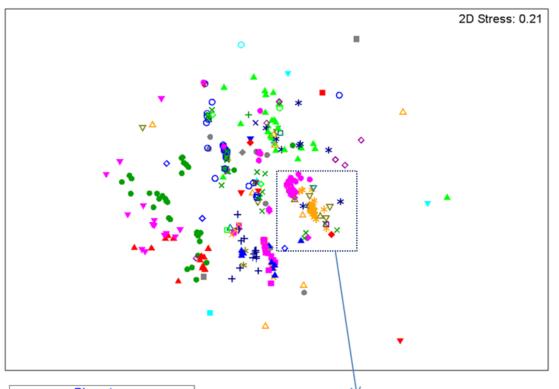


S. Tyhimurium DT 160: initial genotyping study

Omar et al 2010



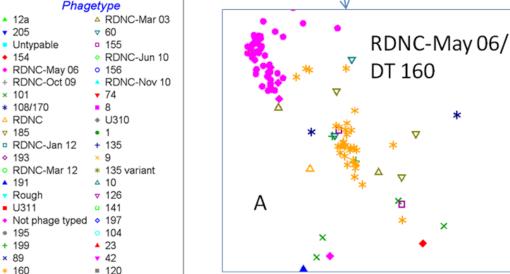
^{**}Total of 76 isolates comprising 23 human, 24 avian, 29 poultry environment.



Pleydell et al 2013

Multidimensional scaling plot of **PFGE** distance matrix: Salmonella **Typhimurium**

DT160 evolved into RDNC-May 06?



* 160 • 120 Te Kunenga Forever discovering ki Pūrehuroa



Genome sequencing: preliminary SNP analysis



Sam Bloomfield

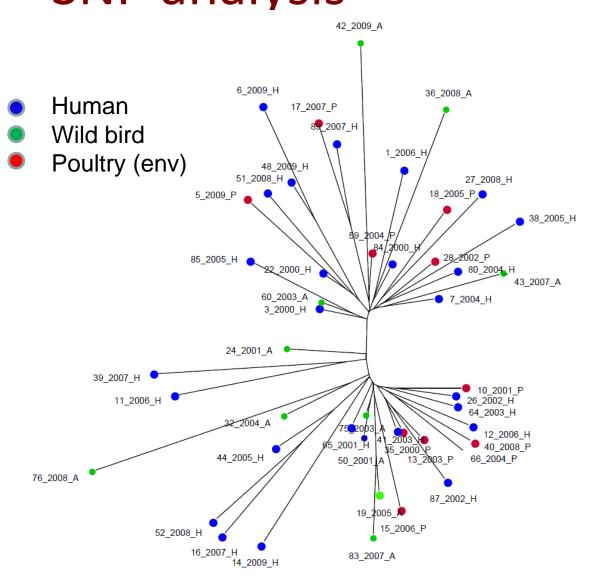
50/120 isolates sequenced

~400 SNPs (needs full QA)

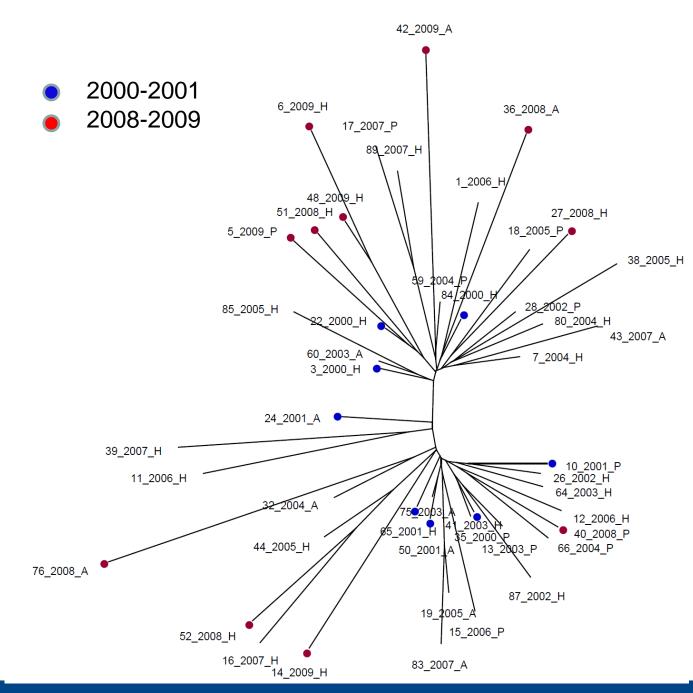
Resolves phylogeny – all unique profiles

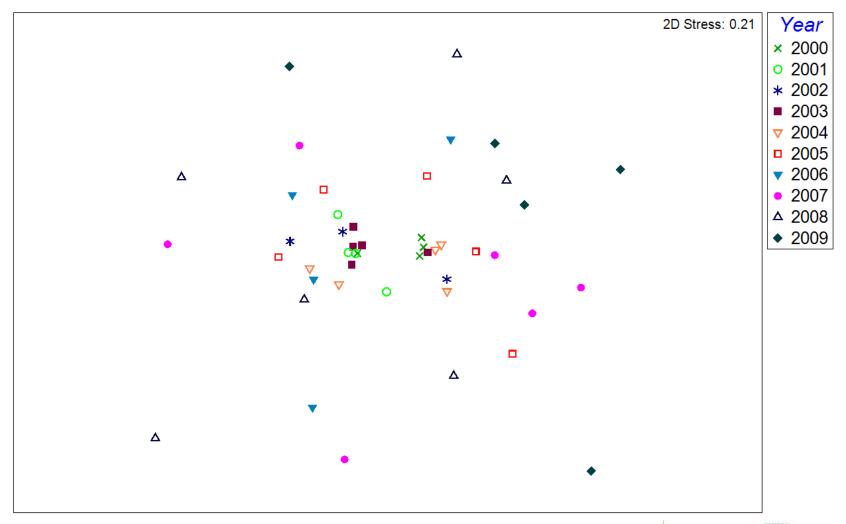
Evidence for zoonotic spread

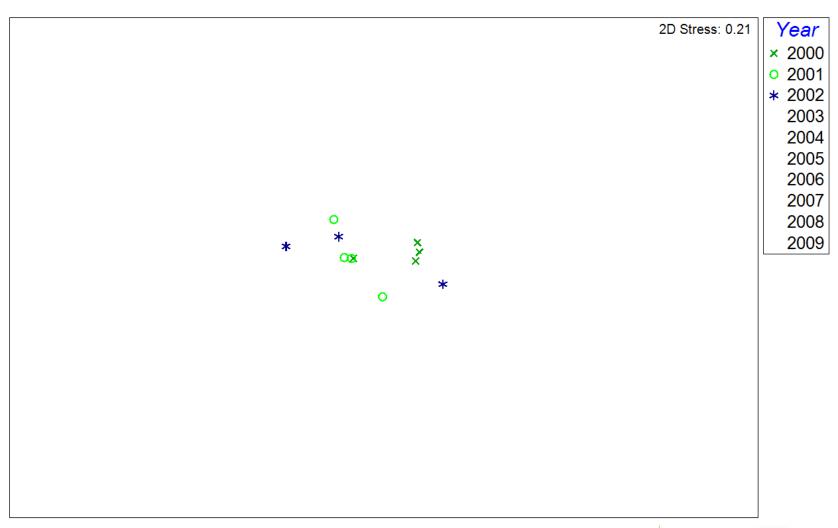
Collaboration with ESR: Phil Carter, Muriel Dufour

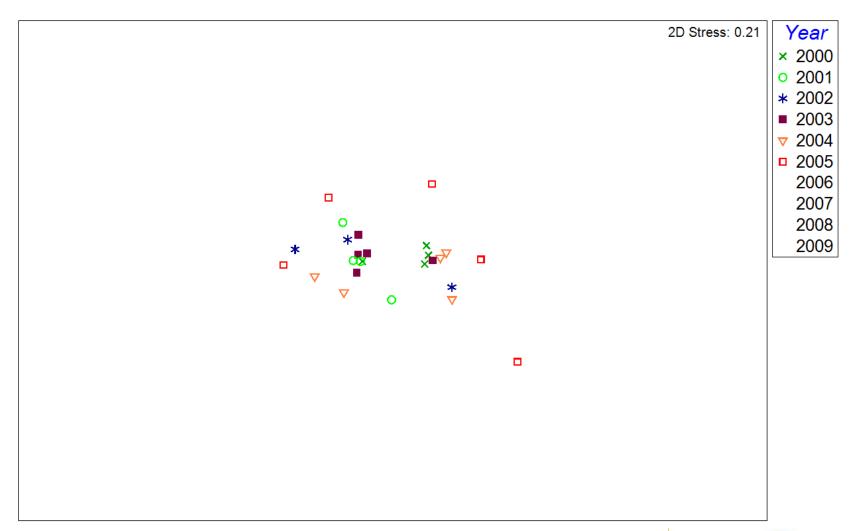


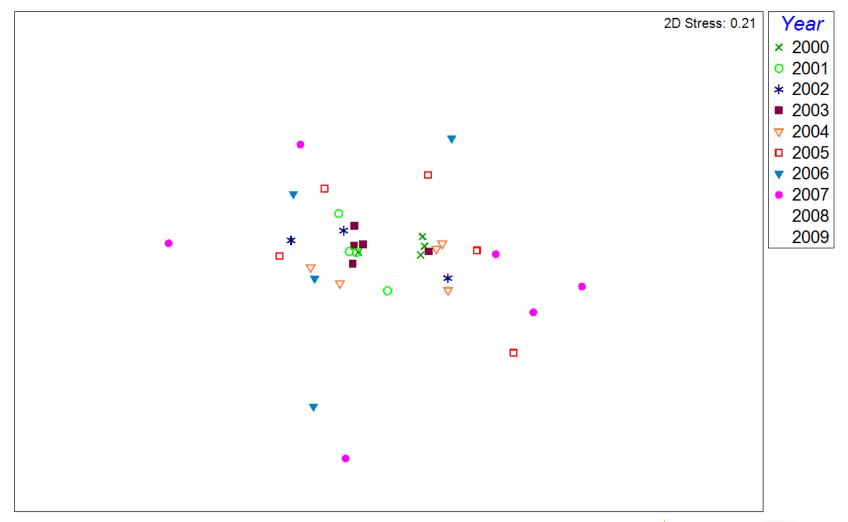
More recent isolates peripheral in tree

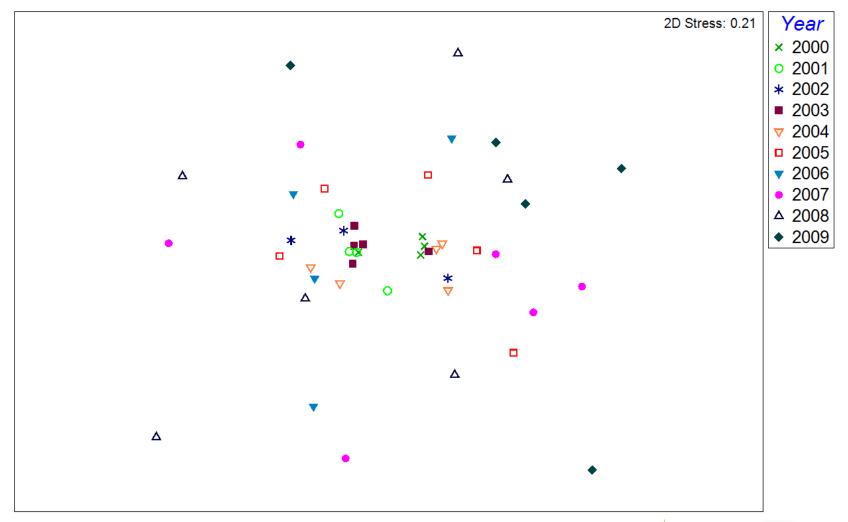




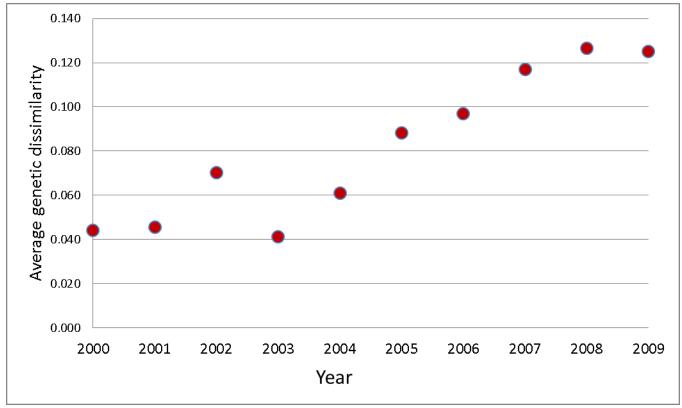








Relationship between year of 'epidemic' and genetic dissimilarity

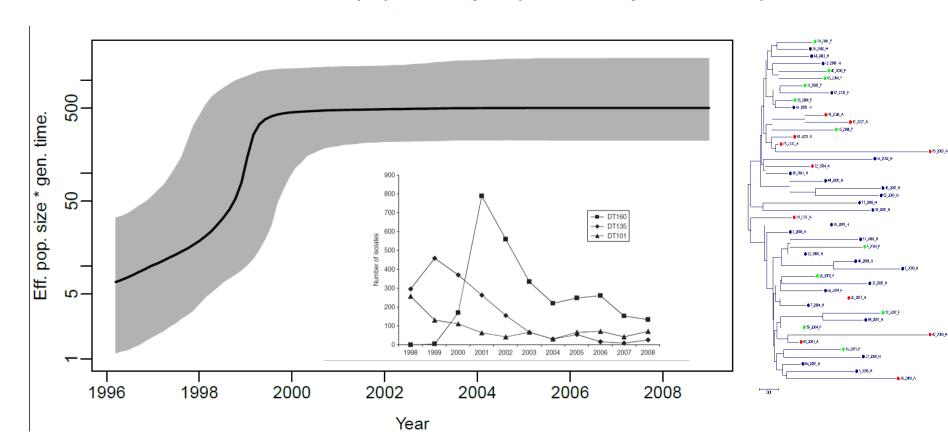


Over time course of epidemic radiating out in recent years, increasing diversity



Bayesian skyline plot (Tim Vaughan)

Massive increase in bacterial population just prior to major human epidemic

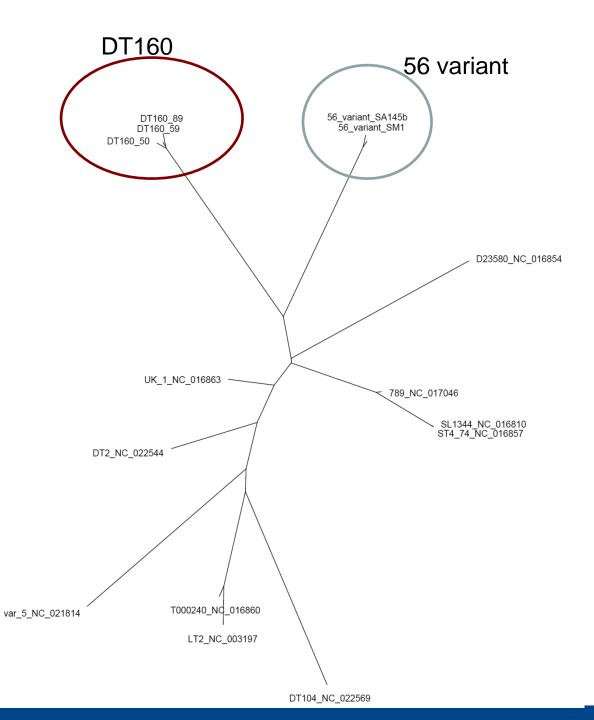


Increase in population size just prior to peak in human cases



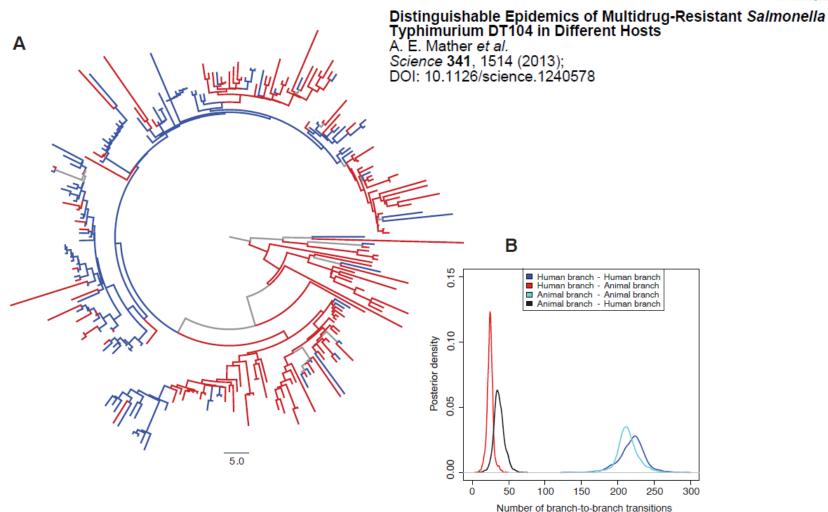
H-0.01

56 variant (RDNC May 06) evolved from DT160?



Host transitions- zoonotic or anthroponotic spread?

www.rndsystems.com



Conclusions: how will it help inform policy?

- Transmission
 - Speed of spread Ro, what vaccine coverage needed?
 - Transmission in (social) networks who's infecting whom? (people and animals, conservation studies)
 - Importance of different transmission routes (zoonotic versus person to person?)
- Host switching
 - When, where, which way?
- Phenotype:genotype relationships
 - Refining source/reservoir attribution?
 - Diagnostics
- Number and origin of introductions
 - Informing biosecurity (where come from, how often)
- Evolution
 - Impact of interventions such as vaccination
 - Effects of antimicrobial use
 - Prediction of emergent strains

Ongoing genomic epi studies in ^mEpiLab

STEC O157
Non-O157 STEC
Salmonella
Campylobacter
Leptospira
Enterococci



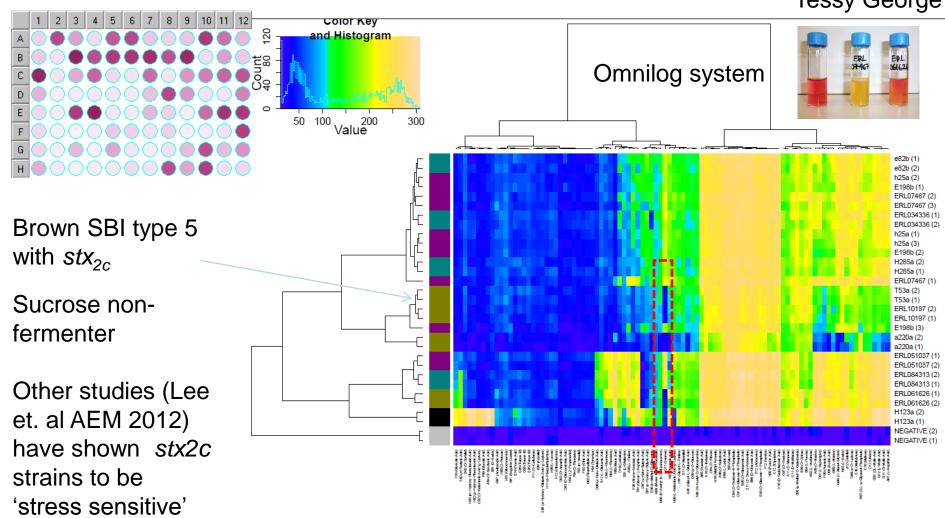
Patricia Jaros



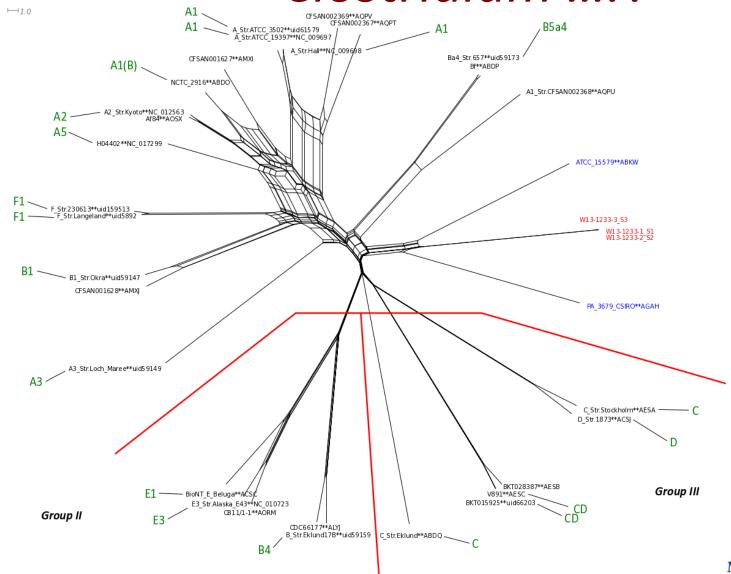
Campylobacter and STEC: genotype:phenotype studies



Tessy George



Food safety and the NZ economy? *Clostridium*?



rMLST





Acknowledgements

- ^mEpiLab team: Patrick Biggs, Jonathan Marshall, Anne Midwinter, Julie Collins-Emerson, Angie Reynolds, Rukhshana Akhter, Charlotte Bolwell, Jacki Benshop, David Hayman, Lynn Rogers, Neville Haack
- ESR Dr Phil Carter, Dr Muriel Dufour, Dr Stephen On, ERL team
- University of Auckland Tim Vaughan, Alexei Drummond
- University of Otago: Prof Michael Baker
- Prof Tim Blackmore
- Ian Bruce NZVP
- Dr Brett Gartrell, Dr Laryssa Howe
- Prof Paul Fearnhead (Lancaster)
- Prof Martin Maiden, Dr Danny Wilson, Dr Sam Sheppard
- MidCentral Public Health, MedLab Central
- AgResearch A/Prof Adrian Cookson
- MPI Peter van de Logt, Prof Steve Hathaway, Dr Donald Campbell, Dr Craig Thornley, Dr Eve Pleydell
- University of Tasmania: Dr Barbara Holland

Funding







MPI

