How phylodynamic modelling can help with the control of human and animal diseases

Nigel French

IDReC
Infectious Disease Research Centre

One Health
AOTEAROA

Te Kunenga
ki Pūrehuroa
MASSEY UNIVERSITY

Acknowledgements

Mycoplasma

- Amy Burroughs
- Edna Gias
- Jonathan Foxwell
- Kate Sawford
- Barbara Binney
- Mary van Andel
- Patrick Biggs
- Samuel Bloomfield
- Simon Firestone

COVID-19

- Jemma Geoghegan
- James Hadfield
- Jordan Douglas
- David Welch
- Joep de Ligt
- Una Ren
- David Winter
- Mike Bunce

Ministry for Primary Industries
Manatū Ahu Matua
NEW ZEALAND

MINISTRY OF HEALTH
MANATŪ HAUORA

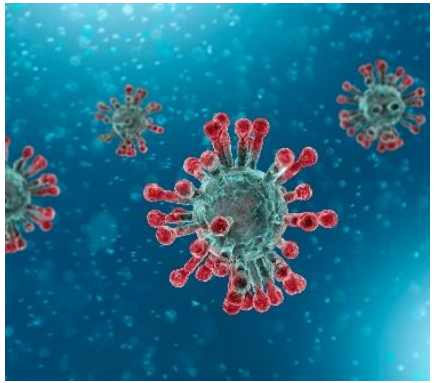
hrc nz

MINISTRY OF BUSINESS, INNOVATION & EMPLOYMENT
HĪKINA WHAKATUTUKI

Core teams

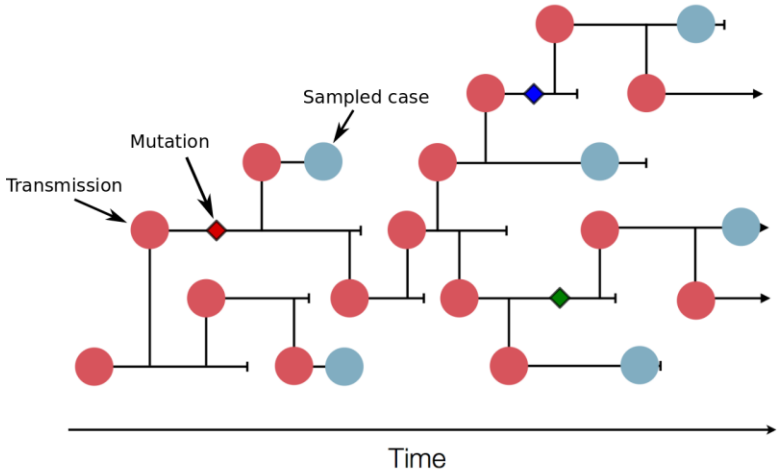
Outline

- What is phylodynamics?
- How is it used to inform decision making?
- Visualisation tools
- Phylodynamics for the control of *Mycoplasma bovis* in NZ
- Phylodynamics for the control of COVID-19 in NZ
- Application for other pathogens



Phylodynamic modelling

- Phylodynamic models - consider evolution and transmission of pathogens within a population
  - Use epi and genetic data to reconstruct transmission histories
  - Individual cases and transmission networks
  - Who acquired infection from whom?
  - R_{eff} (R_t)
- Typically use genetic sequence data and epidemiological information
  - e.g. likely period of infectiousness/exposure window



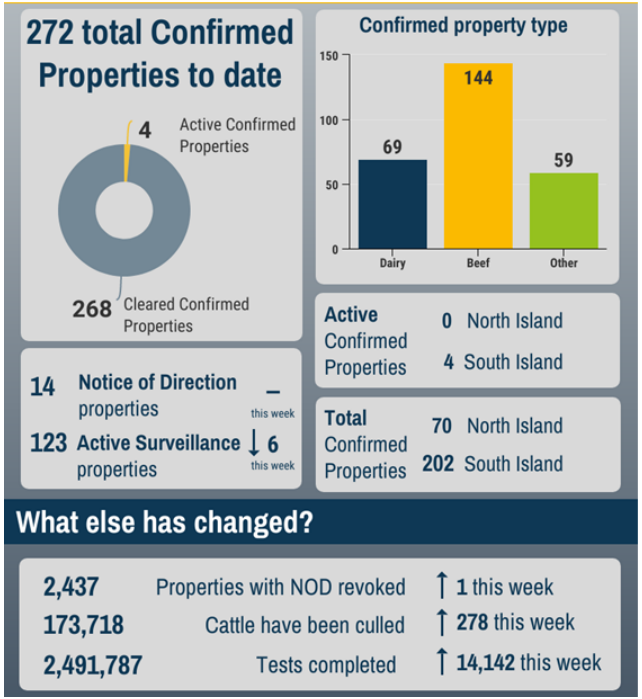
Source:
https://docs.nextstrain.org/

Applications

	Mycoplasma	COVID-19
Estimated time of common ancestor – date of incursion?	✓	(✓)
Who acquired infection from whom?	✓	✓
Effective reproduction number?	(✓)	✓
International origin?	(✓)	✓
Interactive visualisation	✓	✓
Publicly available narratives		✓

Mycoplasma bovis

- Causes mastitis, lameness and septicaemia in cattle
- Endemic in most countries
- Never detected in New Zealand... until
- First detected July 2017
- S. Canterbury
- MPI 2018
  - *Mycoplasma bovis* disease eradication programme
  - Te hōtaka whakakore i te mate mycoplasma bovis



Biosecurity New Zealand
Tiakitanga Pūtaiao Aotearoa

Mycoplasma bovis – – role of phylodynamics in informing the elimination programme

- Genome sequencing of 800+ isolates from ~45% farms
  - Isolation, culture, and sequencing carried out by MPI
- How long had it been in New Zealand?
  - Inform decision to eliminate
- How is it being spread between farms?
  - Who is infecting whom?
  - When did transmission occur?

Genomic Epidemiology outputs for MPI

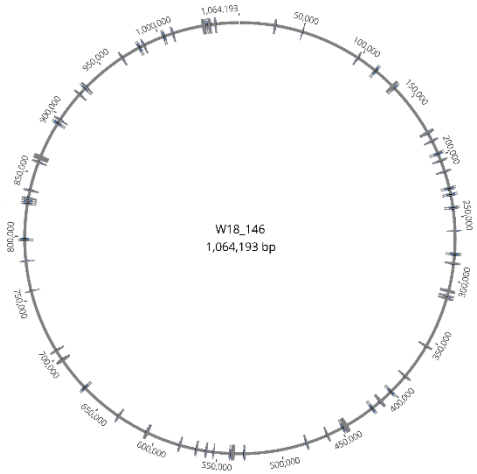
<u>For every new sequence batch</u>:

- Microreact and Nextstrain builds
- Estimated date of common ancestor
- Mutation/SNP distributions
- Circular dendrogram showing ‘clades’
- For each clade and subset:
  - Estimate between farm transmission probabilities
  - Transmission network diagram
  - Inferred ancestral states (farms)

Microreact

Mycoplasma bovis genome

- Short read sequencing of 800+ isolates
- Long read sequencing of two NZ isolates
- 1.06Mb genome
- Not universal genetic code (TGA encodes tryptophan)
- Multiple Insertion Sequence (IS) elements all around the genome, 5.7% of genome
- Associated with genome rearrangements and modulating gene expression (growth)



AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology

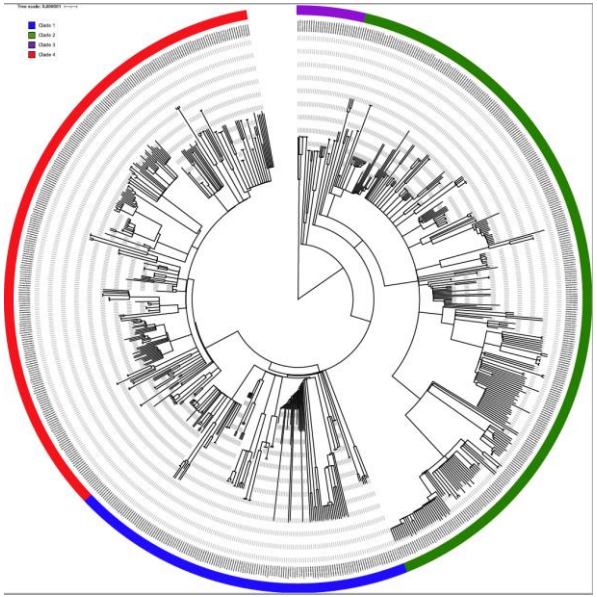
High Prevalence of Diverse Insertion Sequences within the rRNA Operons of *Mycoplasma bovis*

Eytan Amram,[a,b] Ilya Borovok,[c] Yaarit Nachum-Biala,[b] Roger Ayling,[d] Uri Lerner,[a,e] Shimon Harrus,[b] Inna Lysnyansky[a]

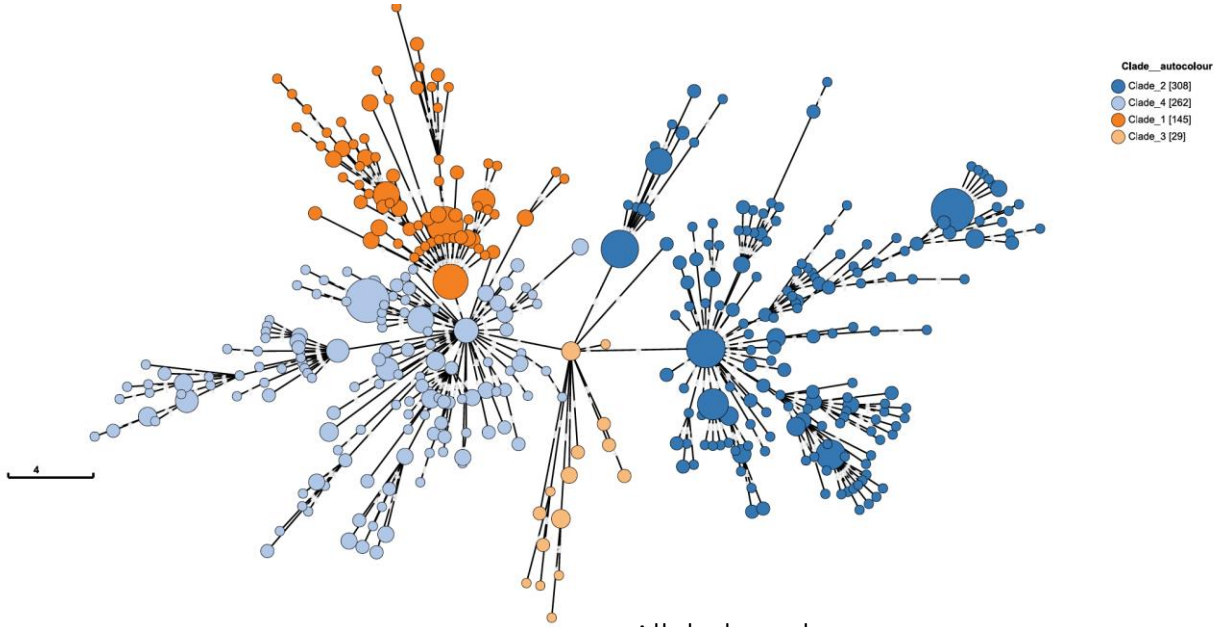
Population structure of *M. bovis* in New Zealand

3-4 distinct genomic groups or ‘clades’ evolved prior to 2018

7 gene MLST – ST 21

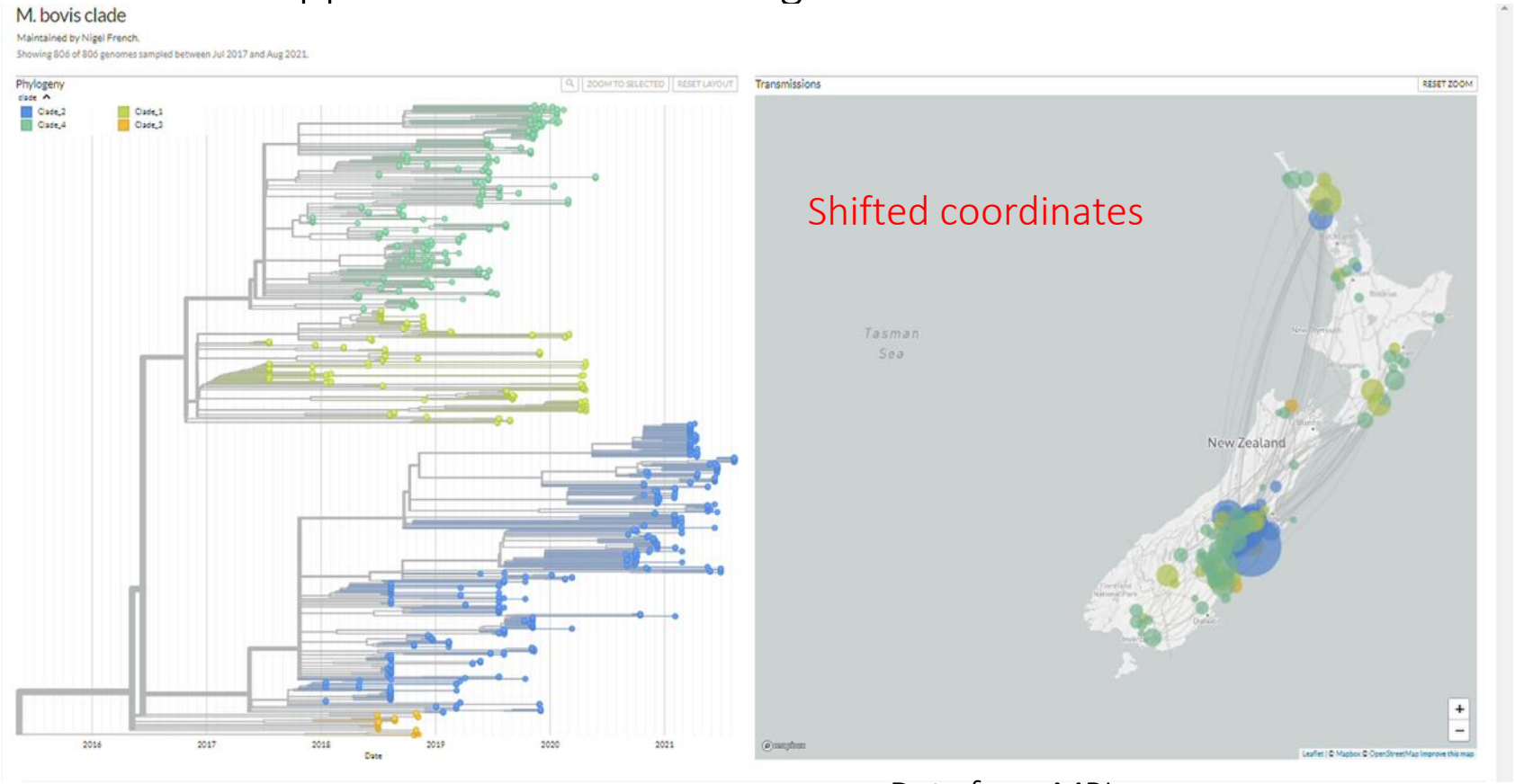


SNP-based



Allele-based

Nextstrain application and the emergence and extinction of clades

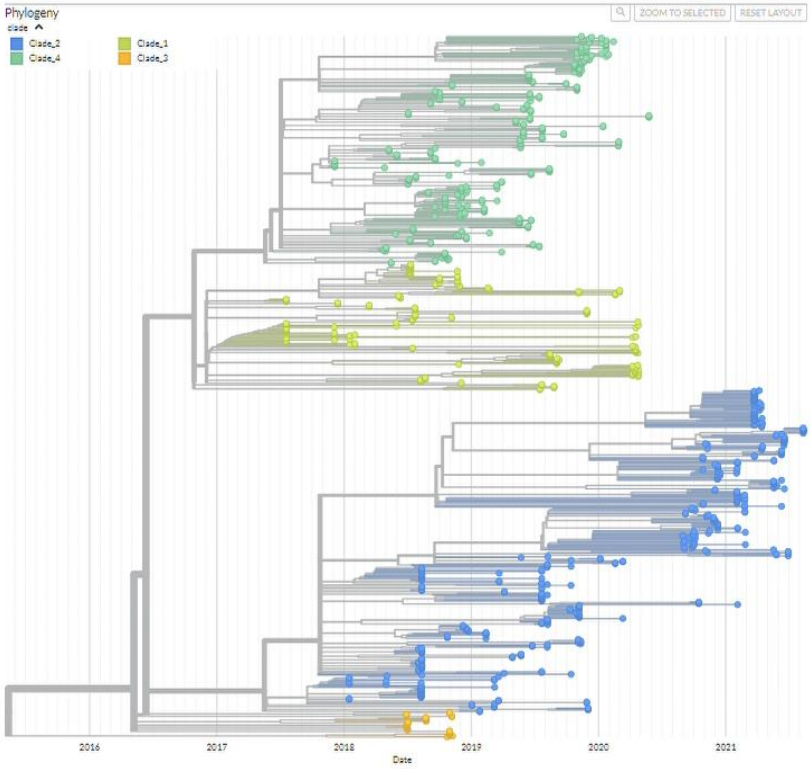


Data from MPI
Nextstrain from James Hadfield

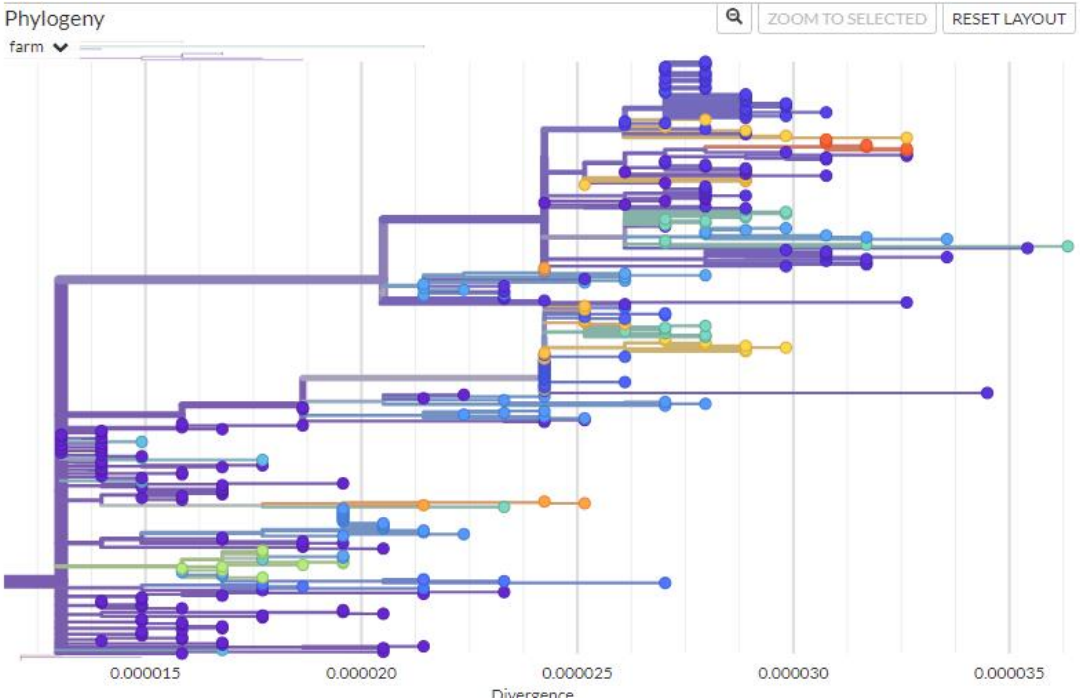
M. bovis clade

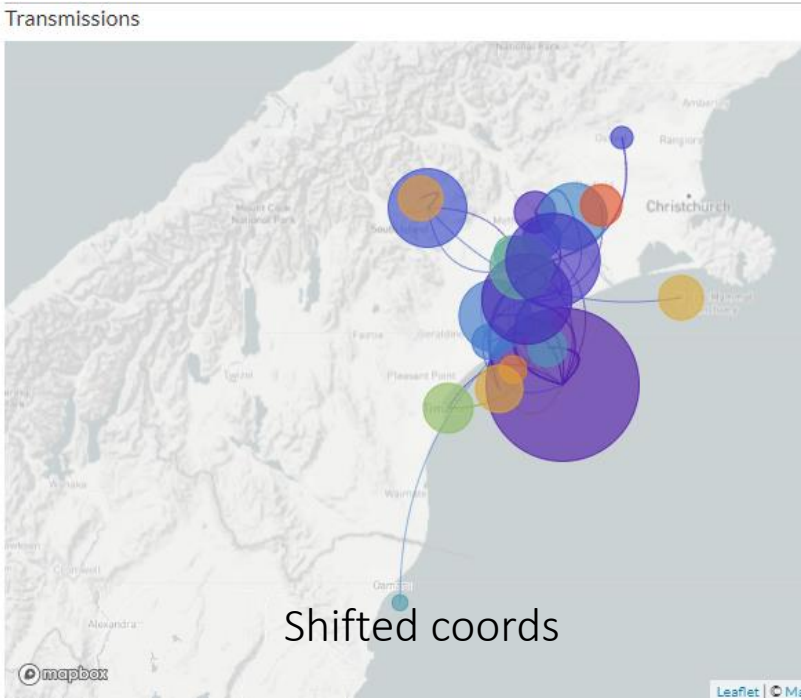
Maintained by Nigel French.

Showing 806 of 806 genomes sampled between Jul 2017 and Aug 2021.



Can be used to explore local transmission

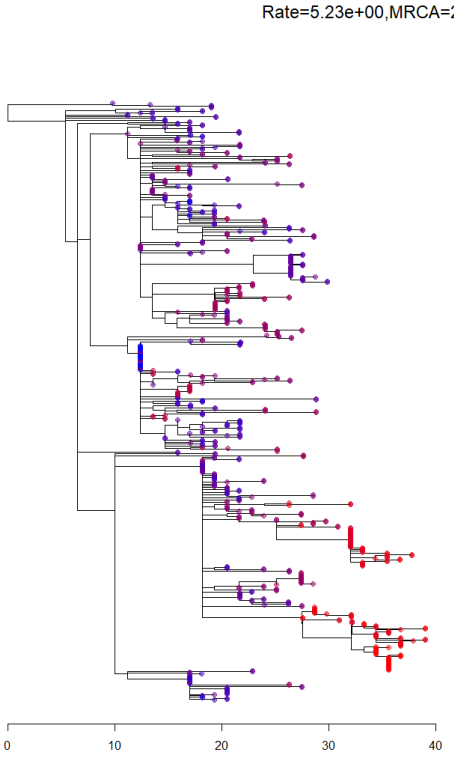


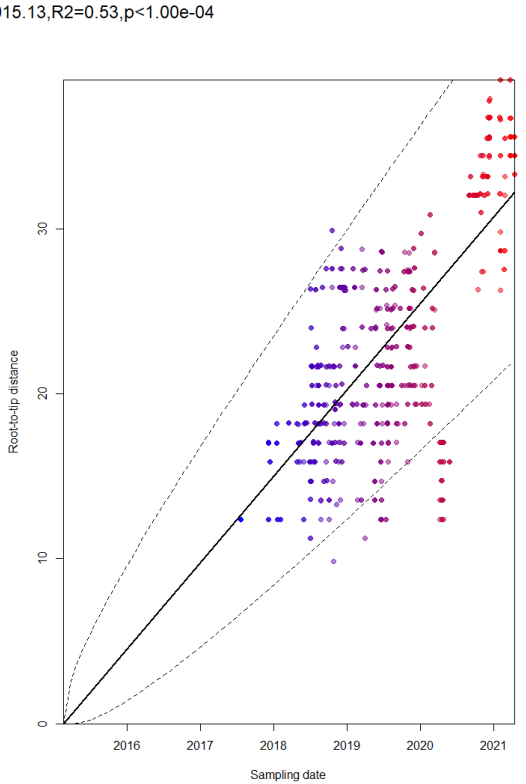


Ancestral dating: approach

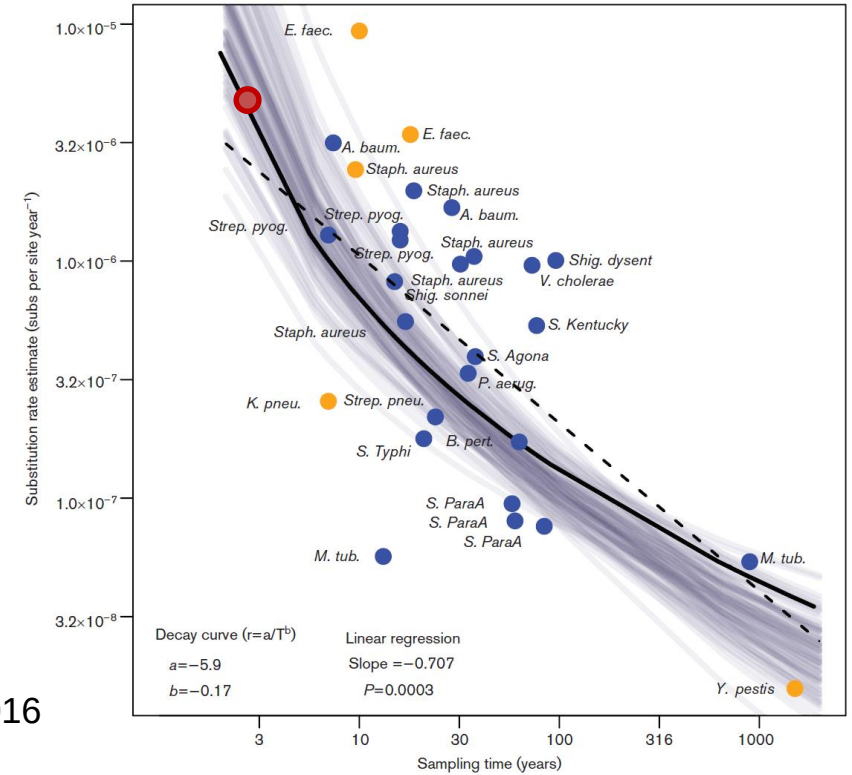
- Is there a time signal?
- Ancestral state reconstruction
  - Use multiple methods
  - Optimisation of base substitution model, tree and clock models
- Estimate time to Most Recent Common Ancestor

Didelot et al *Nucleic Acids Research*, 2018
Bouckaert R, et al. *PLoS Comput Biol* , 2014





Molecular clock and substitution rate:
Estimated rates vary between species and by time period of investigation

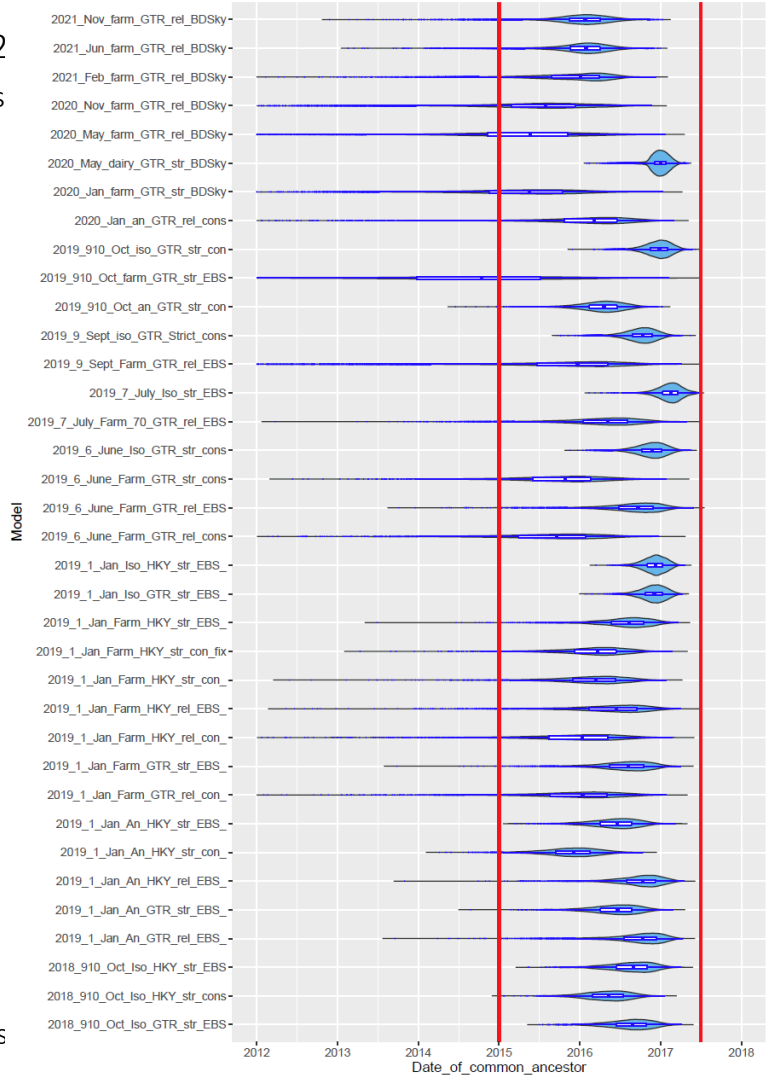


M. bovis ~ 5-8x10^{-6} per site per year

From Duchene et al 2016

November 202

806 isolates 115 farms



Oct 2018

230 isolates 74 farms

Substitution model: (GTR, HKY etc..)

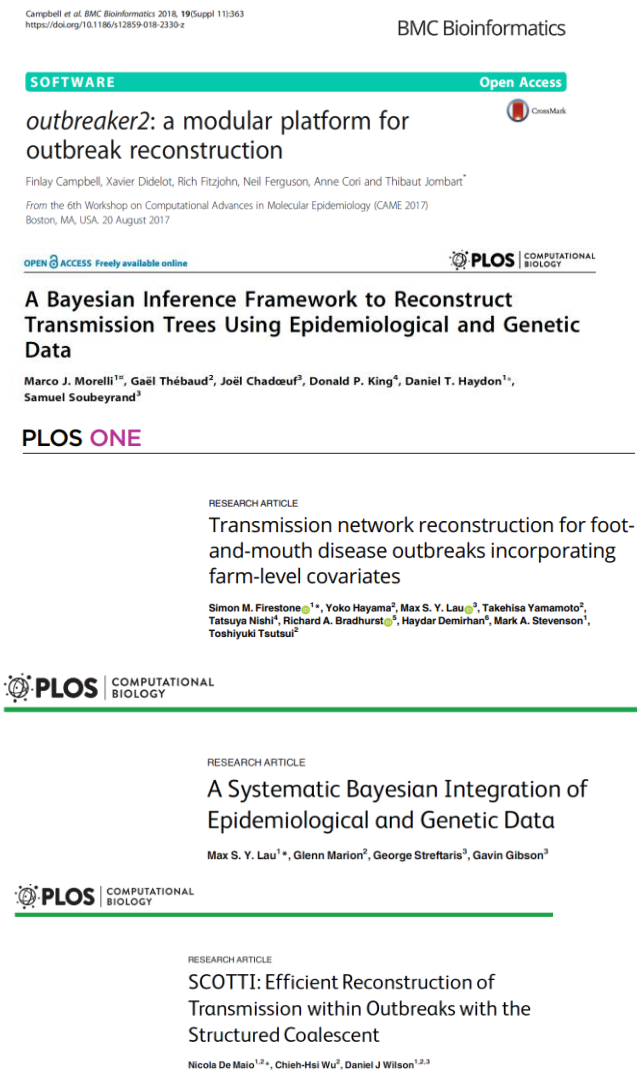
Clock model: (Strict, relaxed, random etc..)

Tree model: (constant, exponential, Birth-Death skyline)

(Reasonably) consistent evidence of recent introduction.
(MRCA consistent with tracing, surveillance)

Who Infected Whom (WIW) models

- Many models – ongoing research
- Different models, different underlying assumptions and use of data
- SCOTTI (Structured COalescent Transmission Tree Inference)
  - Model each host (farm) as a separate pathogen population (deme), and transmission as ‘migration’ between hosts
  - Accounts for within host evolution but not movement data.
  - Combines genetic information from samples with epidemiological information (farm exposure to infection and movement restrictions)
  - Accounts for indirect transmission involving unsampled hosts



Campbell et al. BMC Bioinformatics 2018, 19(Suppl 11):363
https://doi.org/10.1186/s12859-018-2330-z

BMC Bioinformatics

SOFTWARE Open Access

outbreaker2: a modular platform for outbreak reconstruction

Finlay Campbell, Xavier Didelot, Rich Fitzjohn, Neil Ferguson, Anne Cori and Thibaut Jombart*

From the 6th Workshop on Computational Advances in Molecular Epidemiology (CAME 2017)
Boston, MA, USA. 20 August 2017

OPEN ACCESS Freely available online

PLOS | COMPUTATIONAL BIOLOGY

A Bayesian Inference Framework to Reconstruct Transmission Trees Using Epidemiological and Genetic Data

Marco J. Morelli[1*], Gaël Thébaud[2], Joël Chadœuf[3], Donald P. King[4], Daniel T. Haydon[1], Samuel Soubeyrand[3]

PLOS ONE

RESEARCH ARTICLE

Transmission network reconstruction for foot-and-mouth disease outbreaks incorporating farm-level covariates

Simon M. Firestone[1*], Yoko Hayama[2], Max S. Y. Lau[3], Takehisa Yamamoto[2], Tatsuya Nishi[4], Richard A. Bradhurst[5], Haydar Demirhan[6], Mark A. Stevenson[1], Toshiyuki Tsutsui[2]

PLOS | COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

A Systematic Bayesian Integration of Epidemiological and Genetic Data

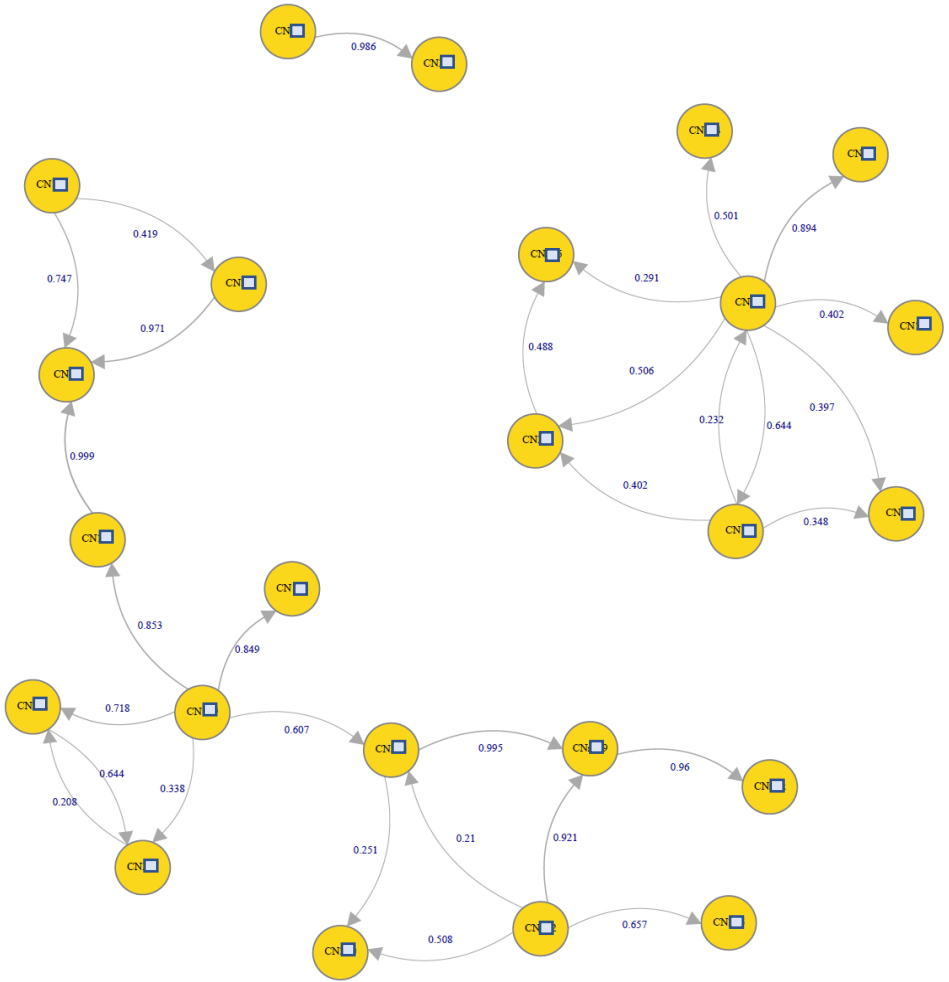
Max S. Y. Lau[1]*, Glenn Marion[2], George Streftaris[3], Gavin Gibson[3]

PLOS | COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

SCOTTI: Efficient Reconstruction of Transmission within Outbreaks with the Structured Coalescent

Nicola De Maio[1,2]*, Chieh-Hsi Wu[2], Daniel J Wilson[1,2,3]



SCOTTI models within host (farm) variation

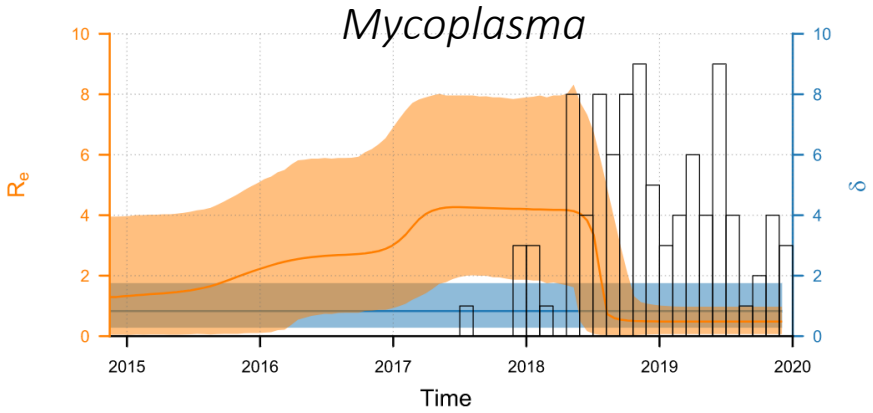
Allows for unsampled hosts

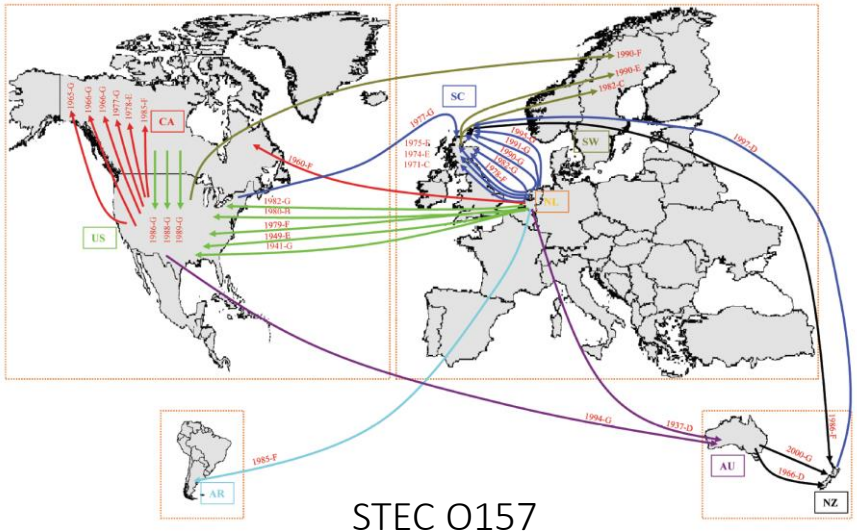
Needs to be considered alongside epi data

to	from	probability
CNXXX	CNYYY	0.73
	Unsampled	0.09
	CNZZZ	0.07

Other applications of phylodynamics

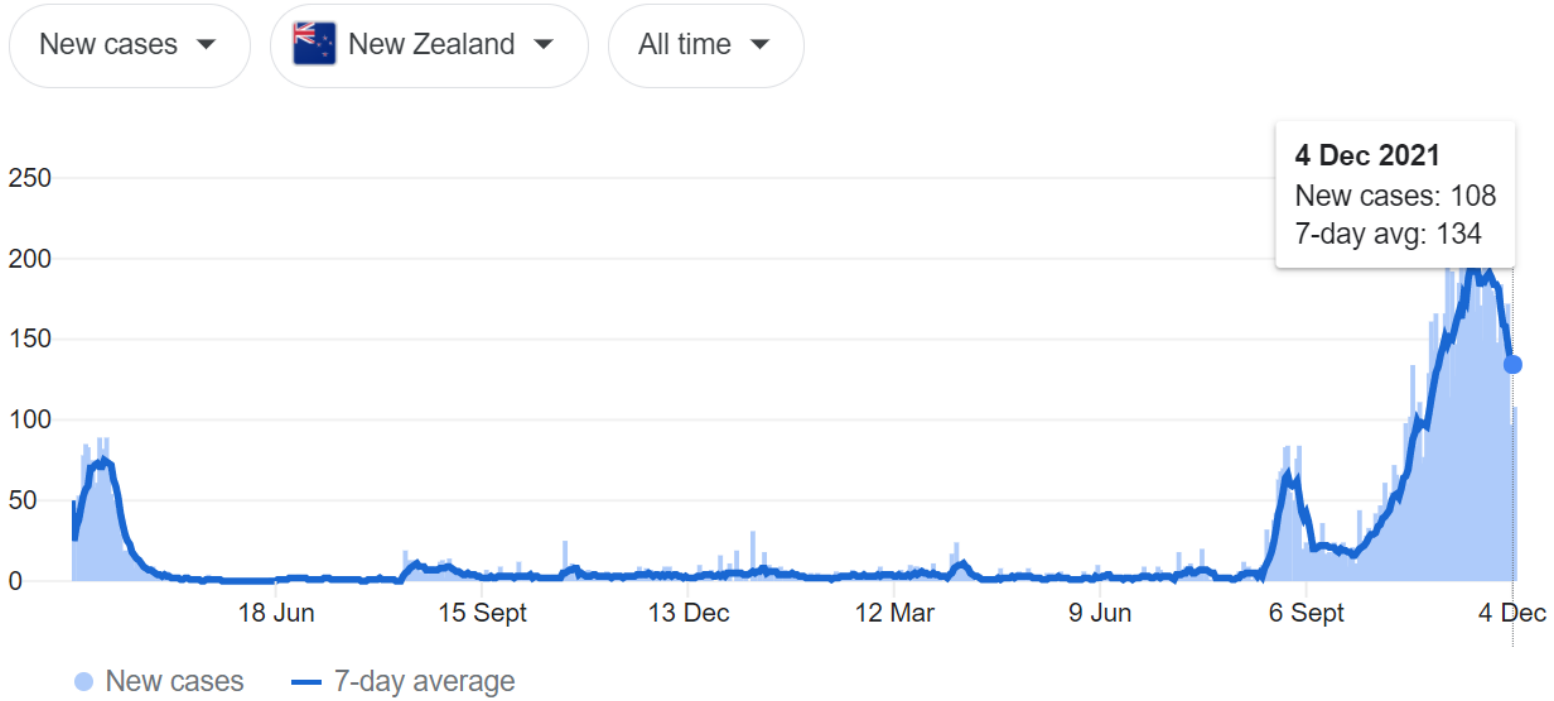
- Reproduction number estimation
- Global origin





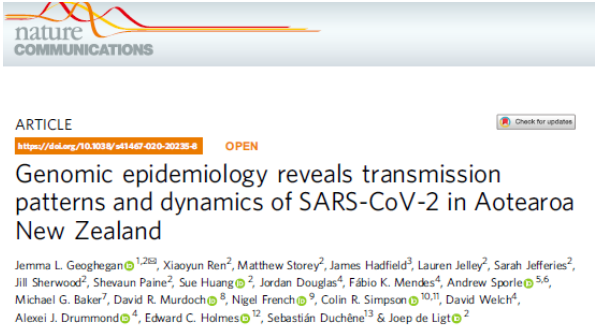
STEC O157

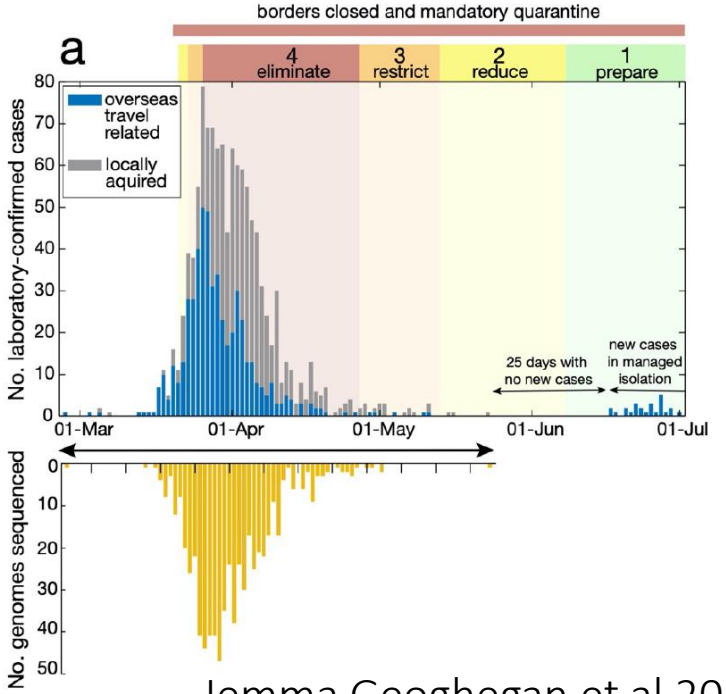
COVID-19 and role of phylodynamics



COVID-19 phylodynamics
Example from first wave

- 649 SARS-CoV-2 genome sequences from the ‘first wave’ between 26 February and 22 May 2020,
- 56% of all confirmed cases in this time period.

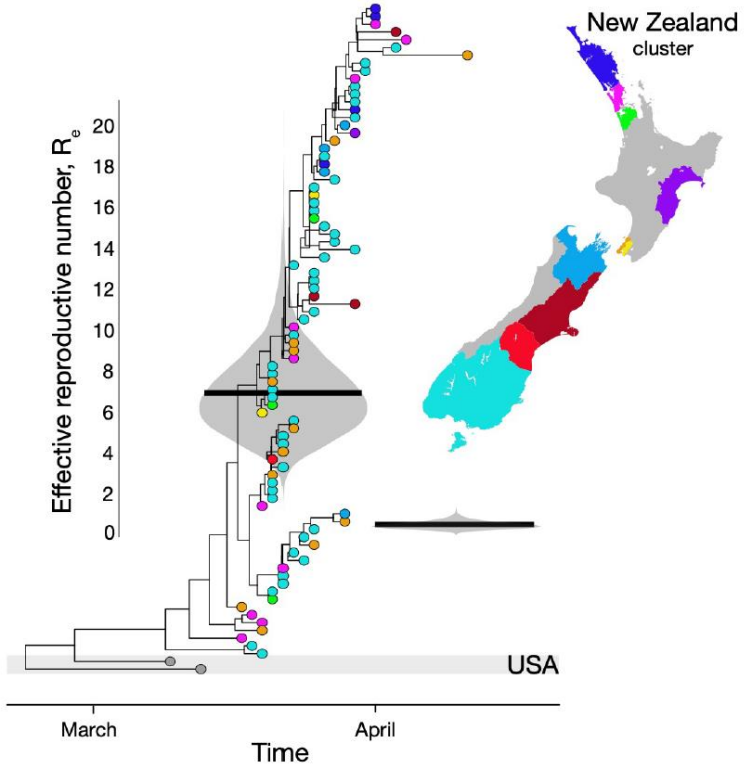




Jemma Geoghegan et al 2020

Phylodynamic modelling of outbreaks

- Linked cases to clusters
- Bluff wedding (98 cases)
- Excluded others
- Superspreader event
- Modelling showed effect of control measures on effective R at two time-points
- Delicined from 7 to 0.2 within 2 weeks of lockdown
- Origin USA



Real-Time Genomics for Tracking Severe Acute Respiratory Syndrome Coronavirus 2 Border Incursions after Virus Elimination, New Zealand

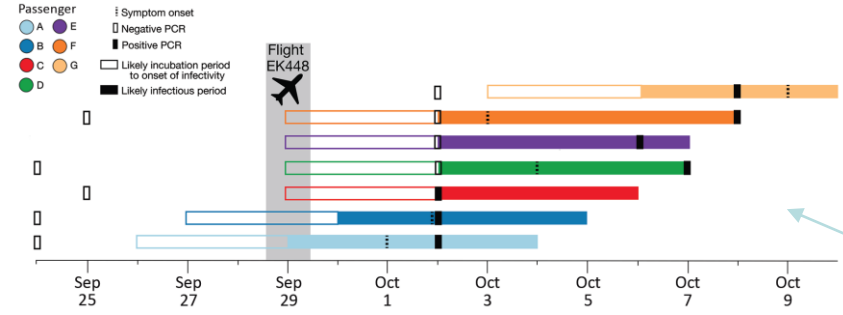
Jordan Douglas, Jemma L. Geoghegan, James Hadfield, Remco Bouckaert, Matthew Storey, Xiaoyun Ren, Joep de Ligt, Nigel French, David Welch

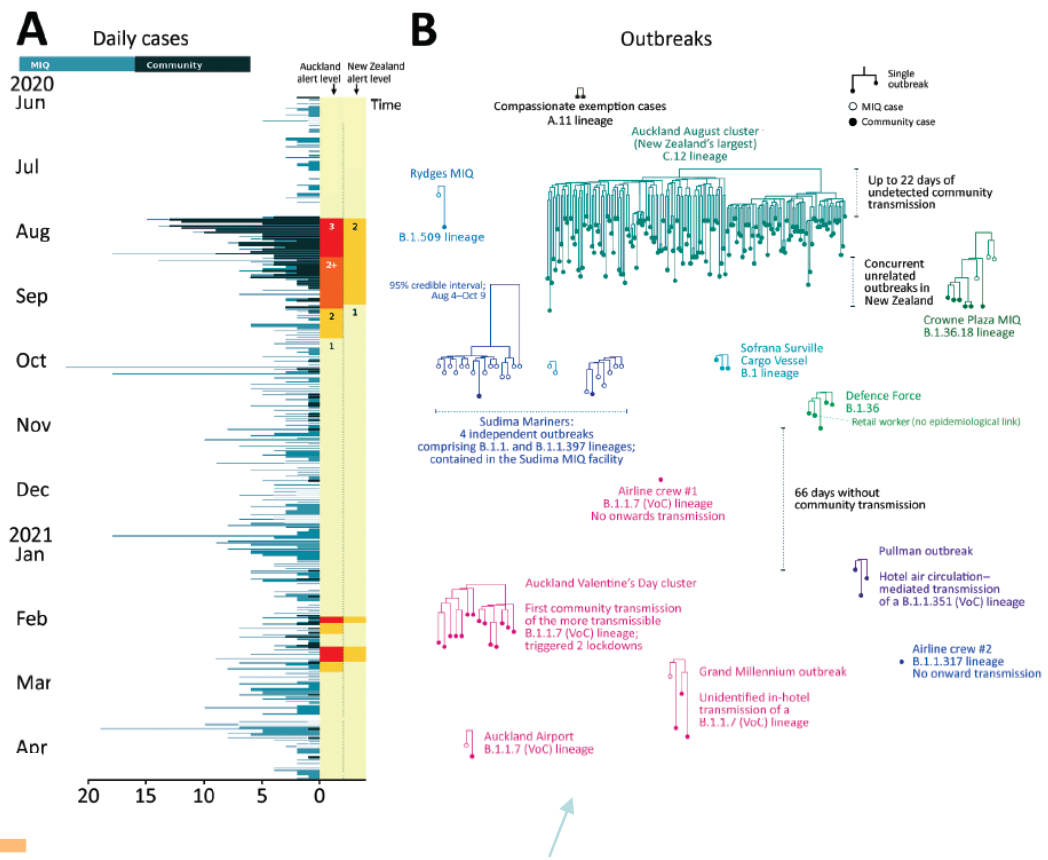
Use of Genomics to Track Coronavirus Disease Outbreaks, New Zealand

Jemma L. Geoghegan,[1] Jordan Douglas,[1] Xiaoyun Ren, Matthew Storey, James Hadfield, Olin K. Silander, Nikki E. Freed, Lauren Jelley, Sarah Jefferies, Jillian Sherwood, Shevaun Paine, Sue Huang, Andrew Sporle, Michael G. Baker, David R. Murdoch, Alexei J. Drummond, David Welch, Colin R. Simpson, Nigel French, Edward C. Holmes, Joep de Ligt

Genomic Evidence of In-Flight Transmission of SARS-CoV-2 Despite Predeparture Testing

Tara Swadi,[1] Jemma L. Geoghegan,[1] Tom Devine, Caroline McElnay, Jillian Sherwood, Phil Shoemack, Xiaoyun Ren, Matt Storey, Sarah Jefferies, Erasmus Smit, James Hadfield, Aoife Kenny, Lauren Jelley, Andrew Sporle, Andrea McNeill, G. Edwin Reynolds, Kip Moulded, Lindsay Lowe, Gerard Sonder, Alexei J. Drummond, Sue Huang, David Welch, Edward C. Holmes, Nigel French, Colin R. Simpson, Joep de Ligt





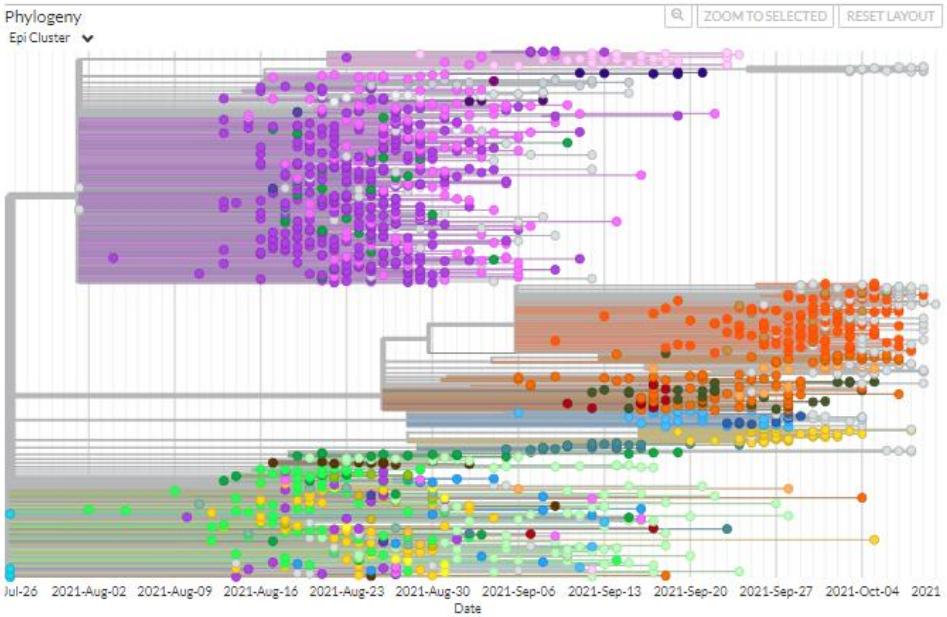
Emerging Infectious Diseases • www.cdc.gov/eid • Vol. 27, No. 9, September 2021

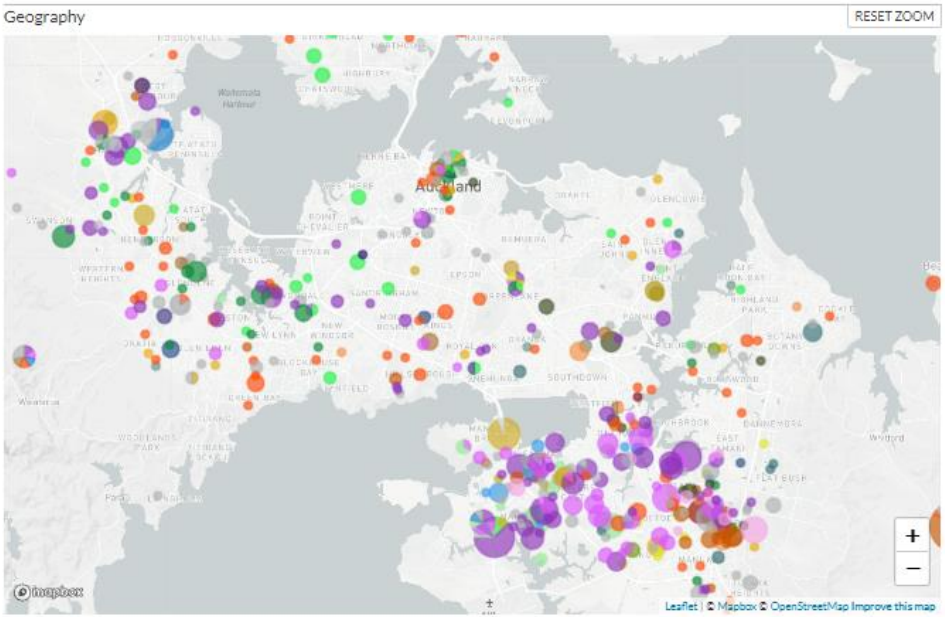
Emerging Infectious Diseases • www.cdc.gov/eid • Vol. 27, No. 5, May 2021

Emerging Infectious Diseases • www.cdc.gov/eid • Vol. 27, No. 3, March 2021

Delta outbreak in Auckland

~95% confirmed cases sequenced

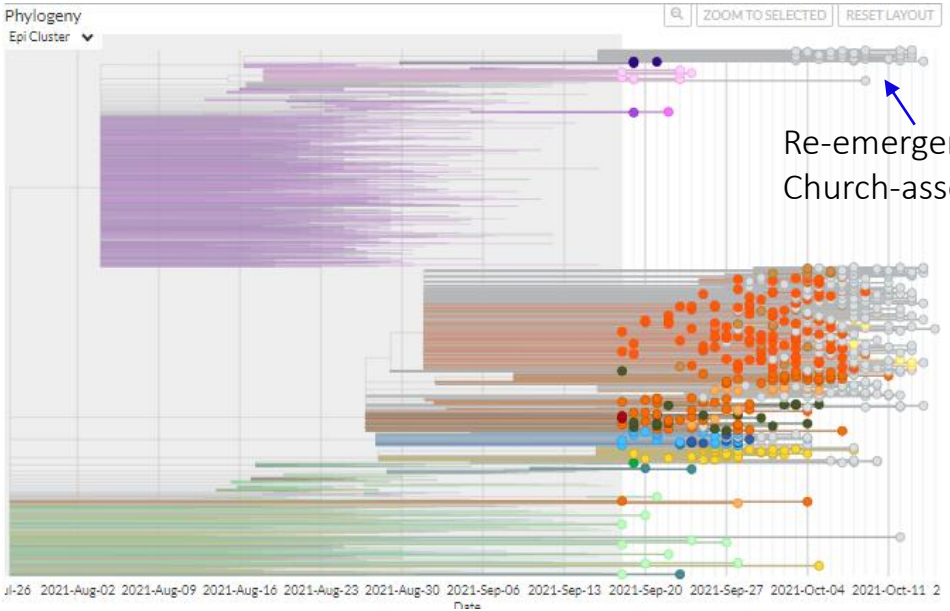




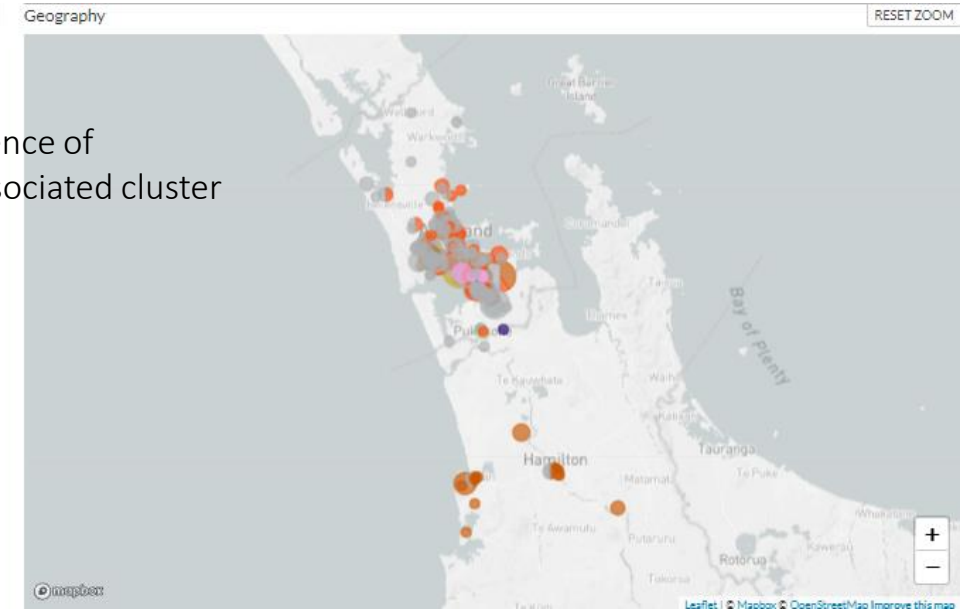
Nextstrain: James Hadfield

Data from Joep de Ligt, Una Ren, David Winter, ESR

Delta outbreak in Auckland



Re-emergence of
Church-associated cluster

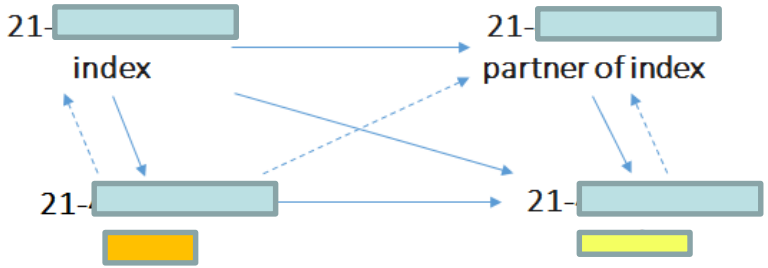


Nextstrain: James Hadfield

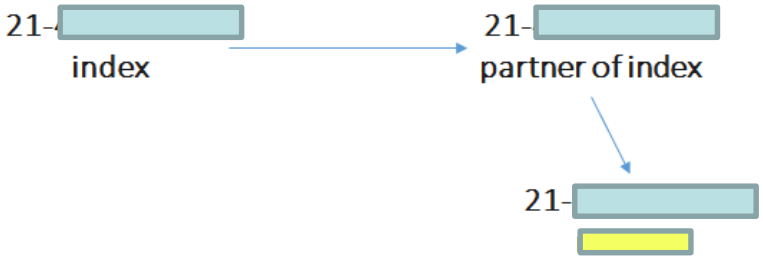
Data from Joep de Ligt, Una Ren, David Winter, ESR

Evidence to support epidemiology and decision making

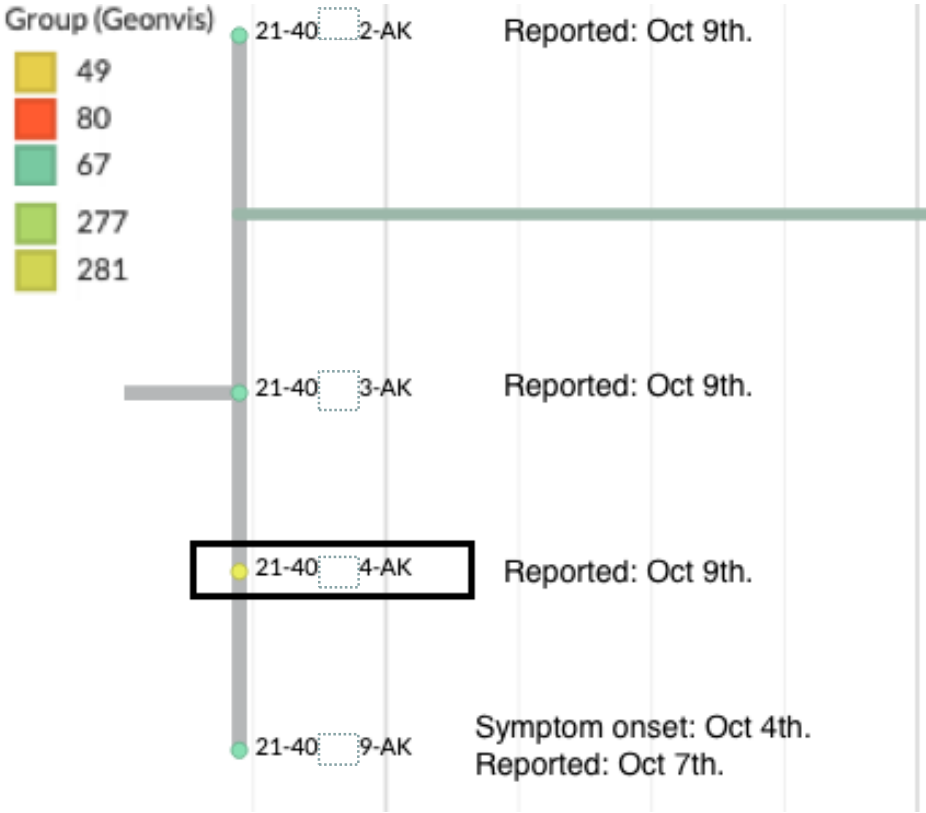
Genomics alone supports the following transmission events
(solid lines also supported by onset dates, dashed lines not consistent with onset dates):

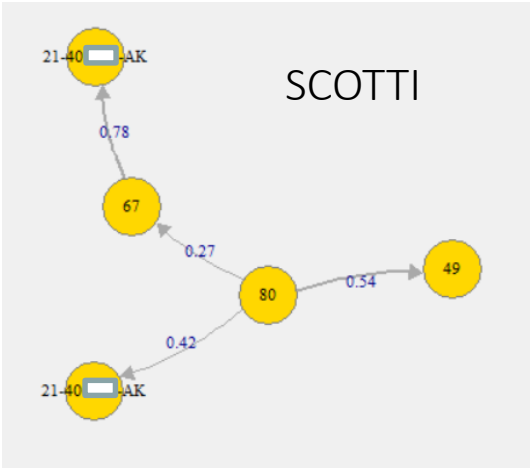


Both genomics and epidemiology support the following transmission events:



Identifying the source of ‘mystery cases’

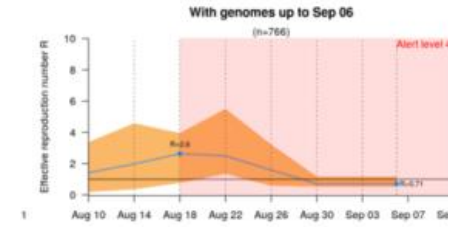


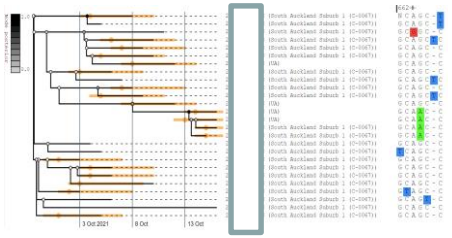


to	from	pr
21-40 4-AK	67	0.78
21-40 4-AK	Unsampled	0.16
21-40 4-AK	80	0.03

Reporting

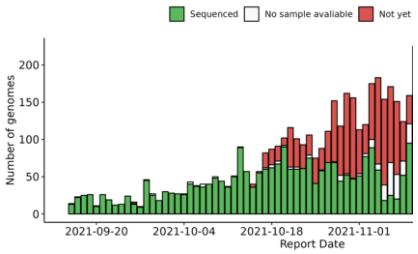
- Delta outbreak
- Daily – weekly
- >95% genomes initially
- Feedback important
- 39 reports since 11th Sept

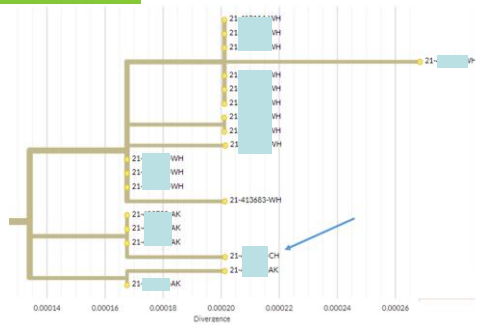


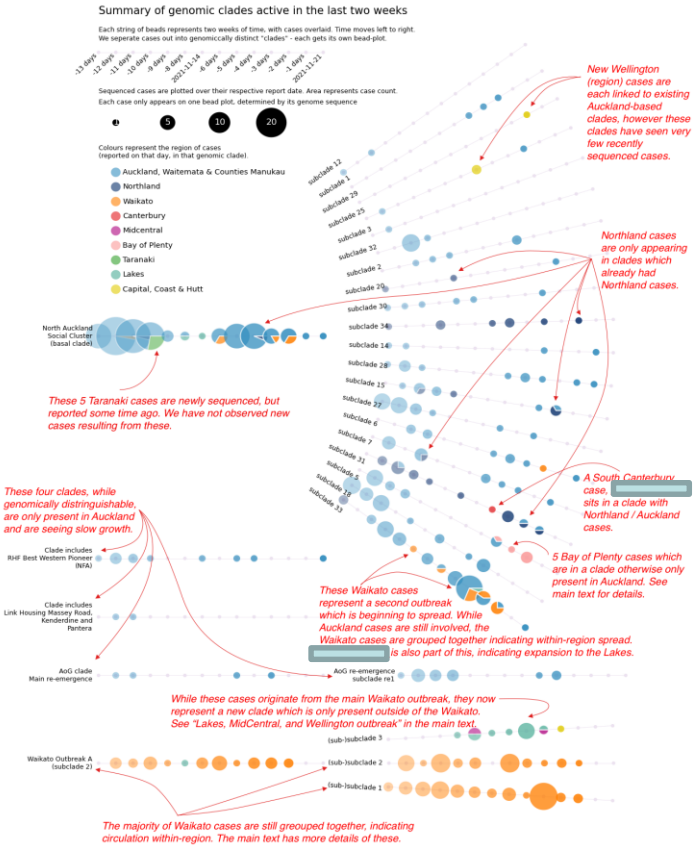


PEACH Tree

Plotting Epidemiological and Alignment CHaracters onto phylogenetic Trees







Nextstrain narrative

Real-Time Genomics for Tracking Severe Acute Respiratory Syndrome Coronavirus 2 Border Incursions after Virus Elimination, New Zealand

Jordan Douglas, Jemma L. Geoghegan, James Hadfield, Remco Bouckaert, Matthew Storey, Xiaoyun Ren, Joep de Ligt, Nigel French, David Welch

Narrative: Real-time genomics to track COVID-19 post-e...

Explore the narrative by scrolling on the left panel, or click "explore the data yourself" in the top right to interact with the data.

Real-time genomics to track COVID-19 post-elimination border incursions in Aotearoa New Zealand

Author: Jordan Douglas, PhD, Jemma L. Geoghegan, PhD, James Hadfield, PhD, Remco Bouckaert, PhD, Matthew Storey, MSc, Xiaoyun Ren, PhD, Joep de Ligt, PhD, Nigel French, PhD, David Welch, PhD [1]

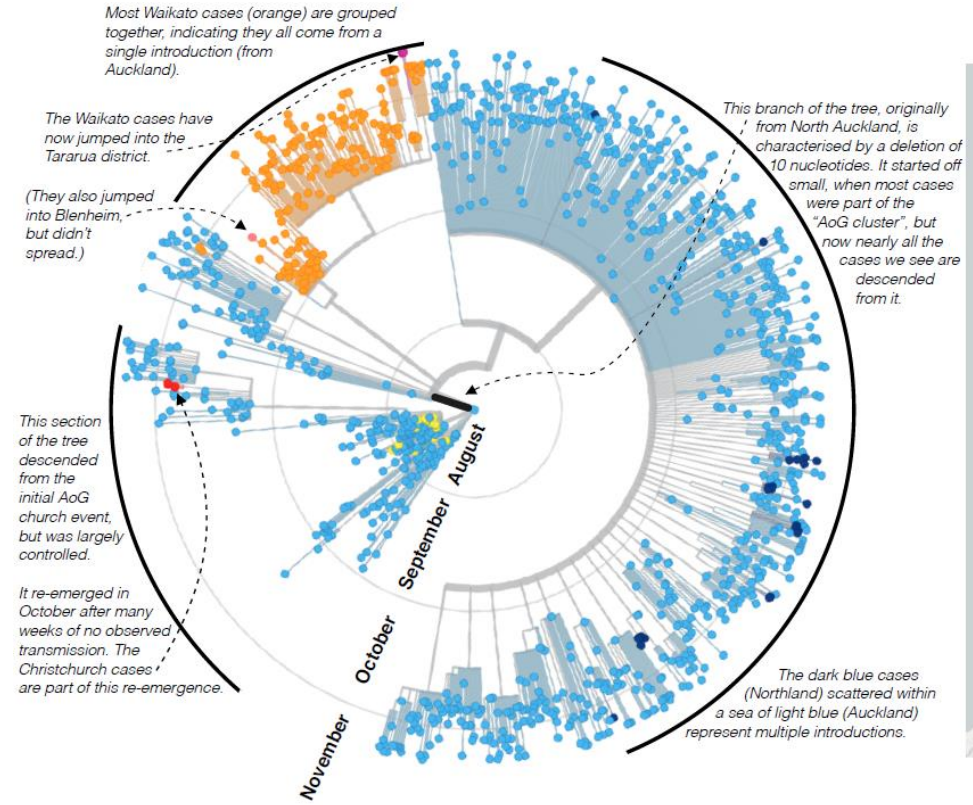
[1] Correspondence: jordan.douglas@auckland.ac.nz; The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

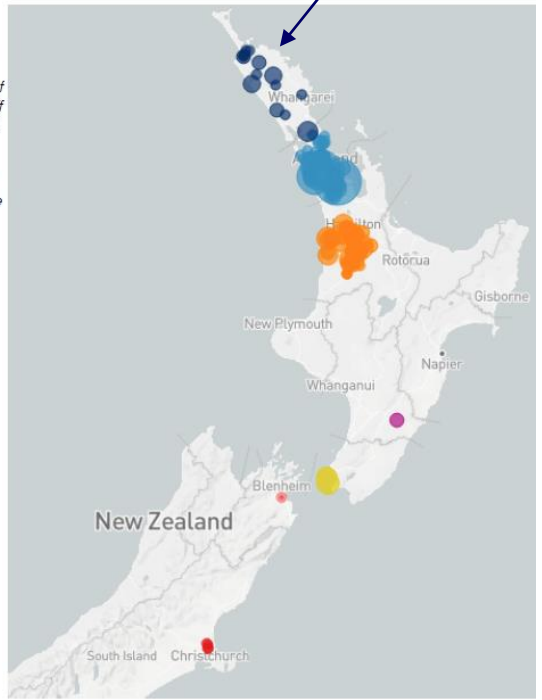
This is an interactive Nextstrain narrative which is a companion to the scientific paper Douglas et al., EID, (2021)). We recount the thirteen community outbreaks from when the virus was successfully eliminated in May 2020 to June 2021, and reflect on how genomics played a vital role in understanding, managing and containing these incursions.

Of the thirteen COVID-19 community outbreaks we look at, two led to stay-at-home orders being issued by health officials. All outbreaks originated at the border; via returnees, airline workers, and cargo vessels. With a public health system informed by real-time viral genome sequencing, which typically had complete genomes within 12 hours after a community-based positive COVID-19 test, outbreaks were well-contained. These incursions resulted in a total of 225 community cases, which unfortunately led to three deaths. Real-time genomics was a vital piece of information for establishing links between cases when epidemiological data could not be obtained or was inconclusive, genomics also assisted in the identification of different origins for concurrent outbreaks.

Nextstrain narrative: Delta outbreak

Multiple introductions into Northland



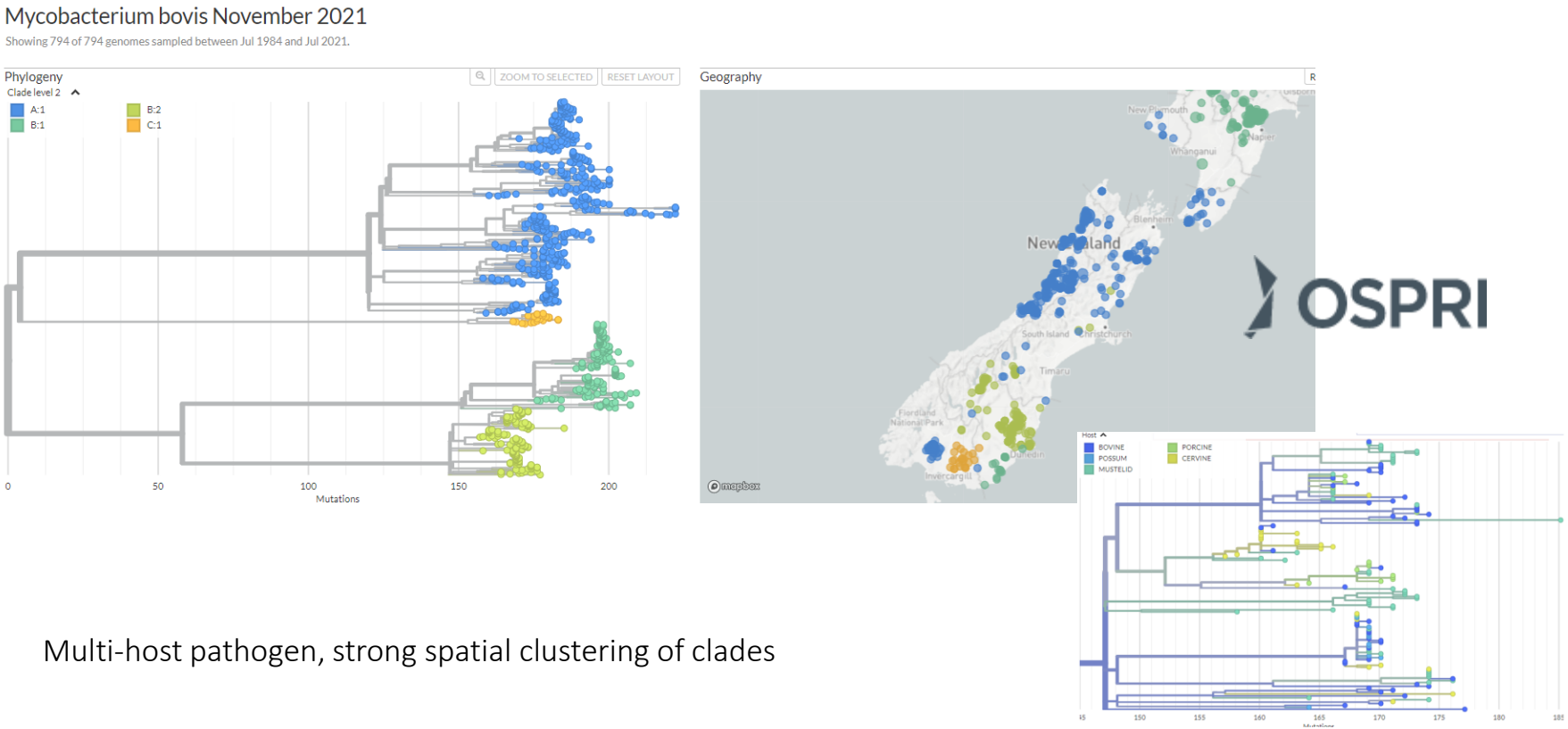


James Hadfield

Example of foodborne pathogen *Salmonella* Enteritidis: new strain emerged in NZ, 2019

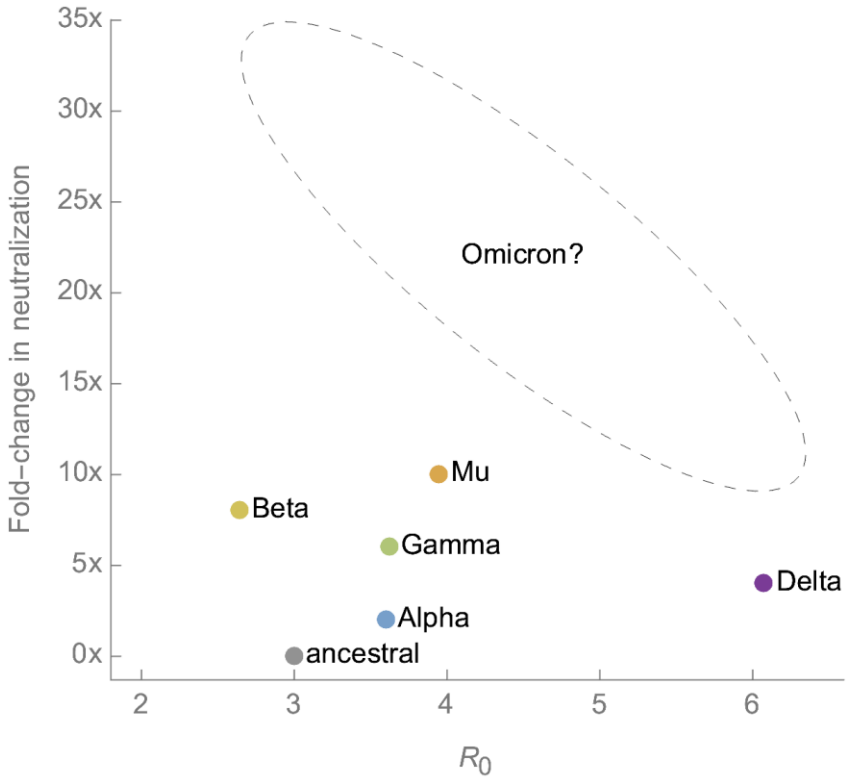
Common ancestor estimated to be early 2019

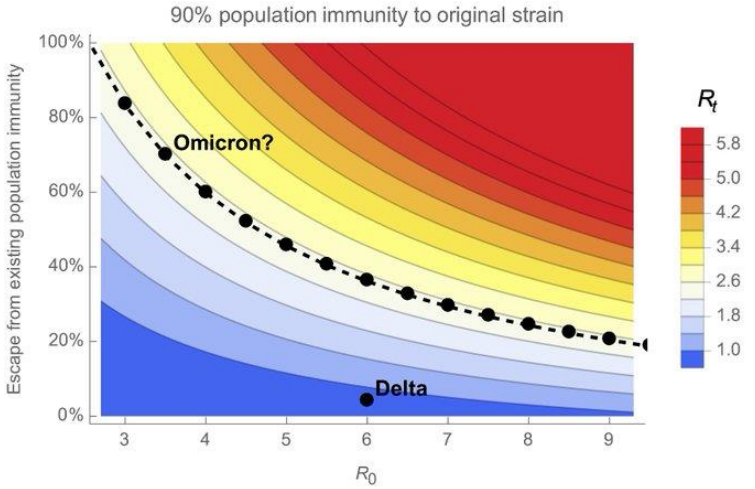
Example of multihost pathogen: Bovine Tb

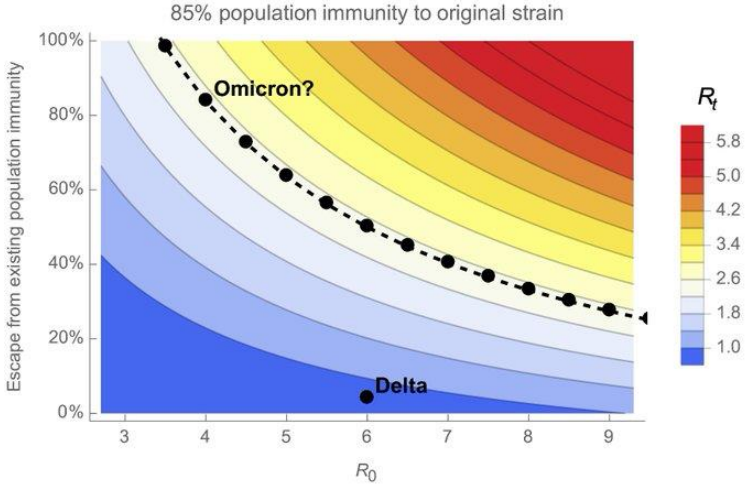


Multi-host pathogen, strong spatial clustering of clades

Omicron – estimate tMRCA then doubling time assuming generation interval to get Reff



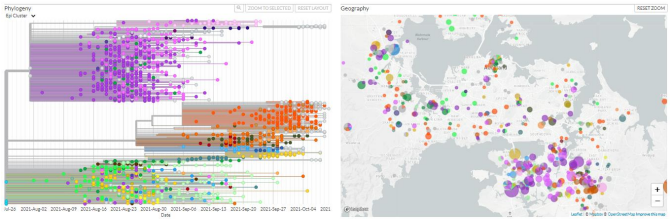




Source: Trevor Bedford, twitter

Conclusions

- Genomic epidemiology applied response in similar way in *Mycoplasma bovis* and COVID-19 responses
  - Visualisation of evolution and transmission
  - Estimation of transmission pathways
  - Date of incursion
  - Reproduction number
  - Global origin
  - Same visualisation tools, similar models
- Continue to learn and develop new approaches
  - Applied to *Mycobacterium bovis, Salmonella* Enteritidis, *Listeria* spp.
  - Challenges – a need to improve models
    - Capture more epi features
    - Run time



Acknowledgements

Mycoplasma bovis
- Amy Burroughs
- Edna Gias
- Jonathan Foxwell
- Kate Sawford
- Barbara Binney
- Mary van Andel
- Patrick Biggs
- Samuel Bloomfield
- Simon Firestone

Mycobacterium bovis
- David Wilkinson
- Kevin Crews
- Richard Curtis
- Marian Price-Carter

COVID-19
- Jemma Geoghegan
- James Hadfield
- Jordan Douglas
- David Welch
- Joep de Ligt
- Una Ren
- David Winter
- Sarah Jefferies
- Colin Simpson
- Mike Bunce
- Alexei Drummond

Salmonella Enteritidis
- Jo Kingsbury
- Jackie Wright
- Jing Wang
- Angela Cornelius
- Shevaun Paine
- Roger Cook
- Ji Zhang

Funding

Ministry for Primary Industries
Manatū Ahu Matua
NEW ZEALAND

MINISTRY OF HEALTH
MANATŪ HAUORA

hrc nz

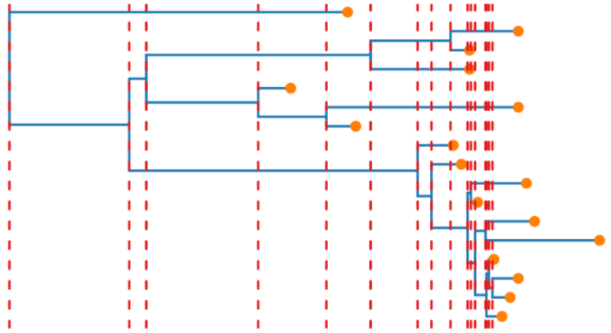
MINISTRY OF BUSINESS, INNOVATION & EMPLOYMENT
HĪKINA WHAKATUTUKI

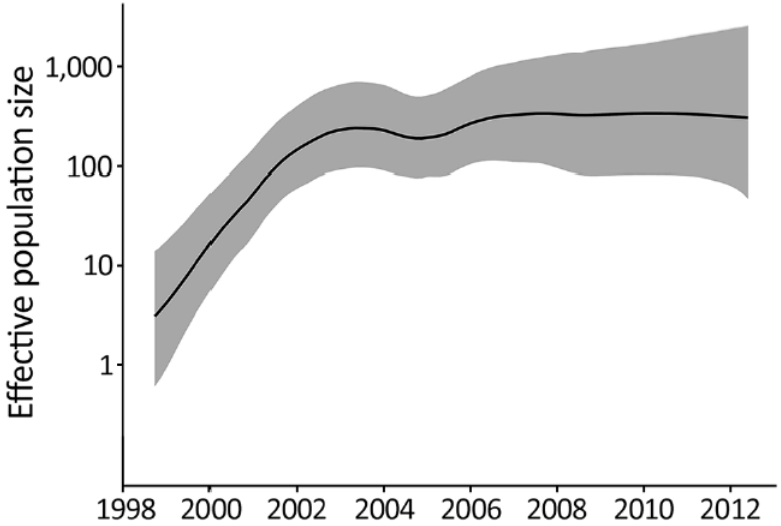
OSPRI

Extra slides

Coalescent Bayesian

- Can detect changes over time of the size of the pathogen population (and number of people infected)
  - The (effective) population size is inverse of the rate of coalescence
    - Larger population - the less likely lineages are to coalesce
    - Many branching events coincide with periods when population small
    - But only works when sampling a small fraction of the population (i.e. not in current outbreak)

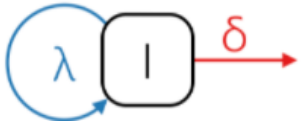




Salmonella Typhimurium

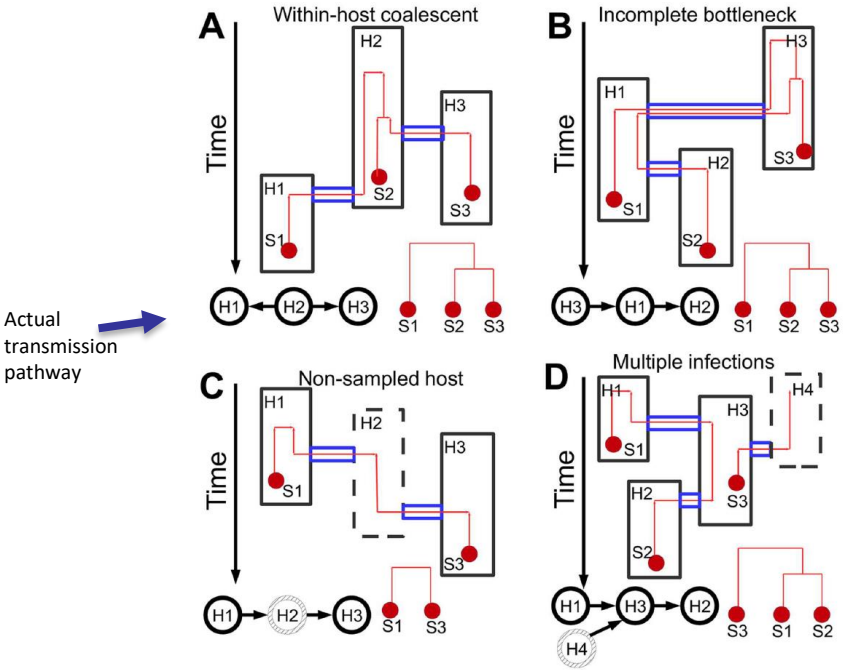
Birth-Death Skyline model

- Parameterised differently to coalescent model
- BD-Sky extends the basic coalescent model to consider contribution to number infected from both:
  - Transmission (Birth)
  - Recovery or death
- Estimates **rate new lineages** are added to the tree (transmission rate) and rate removed (recovery, death or isolation)
- Allows estimation of Reff and captures the epidemic process (and estimates sampling proportion/proportion of lineages sampled)



$$R_e = \frac{\lambda}{\delta}$$

Factors that affect interpretation of phylogenetic relationships and transmission



within-host variation can generate discordance between the phylogenetic and transmission trees

H= hosts (could be farms or individuals)

A = Incomplete lineage sorting (more ancestral lineage in H2 infected H1)

B=Incomplete bottleneck (two lineages transmitted to H1, of which 1 infect H2)

De Maio et al PLoS Comput Biol 2016