Whole genome sequencing for determining the source of *Mycobacterium bovis* in livestock herds and wildlife in New Zealand

M PRICE-CARTER*, D COLLINS, G DE LISLE
AgResearch, Hopkirk Research Institute, Palmerston North, New Zealand

Abstract
The ability to DNA fingerprint *Mycobacterium bovis* (*M. bovis*) isolates helped to define the role of wildlife in the spread of bovine tuberculosis in New Zealand. DNA fingerprinting results currently help to guide wildlife control measures and also aid in tracing the source of infections that result from movement of livestock. During the last 4 years we have developed the ability to distinguish New Zealand (NZ) *M. bovis* isolates by comparing the sequences of whole *M. bovis* genomes (WGS) and demonstrated that WGS provides much higher resolution than our other established typing methods and thus improved the definition of the regional localisation of NZ *M. bovis* types. WGS analysis has led to the confirmation of epidemiological sourcing of infection, to better defining the source of new infections by ruling out possible sources, and has revealed probable wildlife infection in an area considered to be vector free. The routine use of WGS analyses for sourcing new *M. bovis* infections will be an important component of the strategy employed to eradicate bovine TB from NZ livestock and wildlife.

Keywords: *Bovine tuberculosis, Mycobacterium bovis, strain typing, whole genome sequencing, New Zealand*

Introduction
Control of bovine tuberculosis in domestic livestock in New Zealand (NZ) is driven by the zoonotic risk of the causative agent *Mycobacterium bovis*, and its possible impacts on international trade. (1) Although in many countries bovine TB has been controlled successfully with test and slaughter strategies and movement restriction, control is particularly challenging in countries such as NZ in which there is a wildlife reservoir of infection (1,2). Despite these challenging circumstances, the control of bovine TB in NZ has recently been re-evaluated and there are now ambitious goals of achieving TB free livestock and wildlife by 2026 and 2040 respectively.

DNA fingerprinting has played an important role in the control of *M. bovis* in NZ. Studies that employed an early DNA fingerprinting assay that compared the restriction pattern of DNA digests (REA typing) of *M. bovis* isolates, demonstrated that livestock and wildlife in the same regions tended to share the same types and thus helped to define the role of wildlife in the spread of bovine tuberculosis in New Zealand (1,2,3). REA typing was used routinely for over 20 years to efficiently guide wildlife control measures and to aid in tracing the source of infections that resulted from movement of livestock (4). DNA fingerprinting by an easier to perform and interpret but less sensitive VNTR assay currently aids *M. bovis* control efforts (5).

Recent advances in DNA sequencing have facilitated the routine comparisons of entire bacterial genomes (whole genome sequencing (WGS)) for determining the source Mycobacterial infections and this technology shows promise in aiding bovine TB control (6). During the last four years we have developed the ability to distinguish NZ *M. bovis* isolates by WGS and demonstrated that WGS provides much higher resolution than our other established typing methods and thus has improved the definition of the regional localisation of NZ *M. bovis* types (7). Through partnership and contracted work with TBfree and collaborations with Welcome TrustSanger, the University of Glasgow, USDA, AHVLA, Landcare and Massey, we have developed a database with over 500 entries of important *M. bovis* types, a data processing method and an indication of how quickly *M. bovis* types acquire mutations in different NZ bovine TB cycles (8). This information has helped to precisely define the lineage of NZ *M. bovis* types and has facilitated accurate determination of the source of isolates by identifying single nucleotide polymorphisms (SNPs) that differ from a reference genome and by comparing these to SNPs to those detected in other isolates. Examples will be presented of infections in Mt. Cargill, Waiuku, and Harihari in NZ to illustrate how the resulting SNP linages are superior to REA and VNTR typing for determining the source of new *M. bovis* infections.

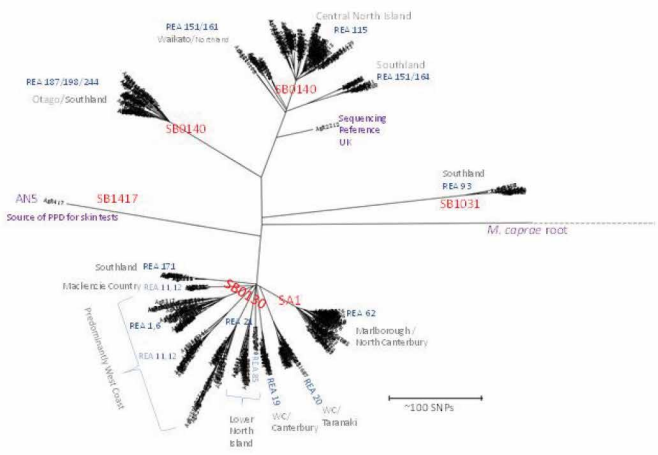
Methodology
The AgResearch *M. bovis* archive has over 8000 NZ isolates that were cultured between 1985 and 2016, from livestock and wildlife suspected of *M. bovis* infection during the post-mortem examination performed as part of routine surveillance. Conventional microbiological tests (described in *de* Lisle *et al*. 2008) were used to positively identify *M. bovis* infection. The WGS database has been assembled by characterising isolates from the archive selected to provide a representative sample of the *M. bovis* population circulating in cattle and wildlife across New Zealand between 1985 and 2016. Most isolates that were characterised by WGS were previously either REA typed (4) or VNTR typed (5). Selected isolates were cultured in Tween albumin (TAB) media from frozen stock and DNA was prepared by CTAB extraction. DNA library preparation and genome sequencing was performed either at the NZGL facility at Massey in New Zealand, or at the Welcome Trust Sanger Institute or at The

Welcome Trust Glasgow facility in the UK, on an IIlumina MiSeq instrument, with 2 X 250bp paired end reads. Genomic data were trimmed and mapped to the UK M. bovis reference genome (A2122). The resulting alignments were processed to generate a list of SNPs for consideration. SNPs detected in regions that are not well characterised by this methodology (11), such as PE PGRS regions, insertion (IS) elements or in regions that were poorly covered were excluded from the analysis. The remaining SNPs were processed together in order to devise concatenated lists, which were compared in order to define the phylogenetic relationship of the isolates. The relationship between isolates are illustrated with Bio-neighbour joining distance plots that were generated using a Juke Cantor model and SNP tables.

Results

The relationship determined by WGS for the *M. bovis* types that are prevalent in NZ is illustrated by the SNP lineage in Figure 1. Most are genotypes that were common in the UK when cattle were imported into NZ in late 1860s (12). WGS results corroborate findings from previous REA and VNTR typing studies which revealed that distinct types predominate in different parts of NZ (3,4,5).

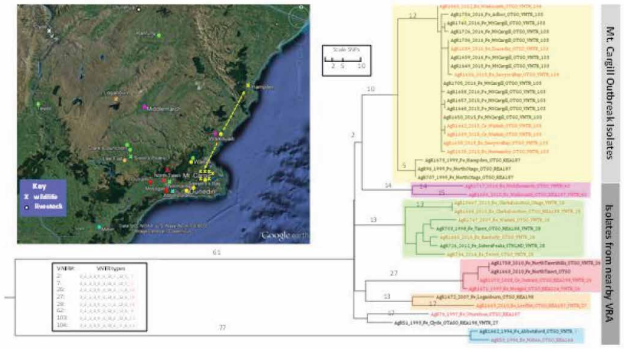
Figure 1. Radial phylograph illustrating the relationship of NZ *M. bovis* isolates. The scale bar indicates the distance in SNPs between isolates. SB numbers labelled in red are internationally recognised spoligotypes based on differences in the DR/Crispr region. The REA types listed are the predominant REA type(s) in the indicated cluster.



Mt. Cargill

An investigation of isolates from the Mt. Cargill region of the South Island was carried out to aid in determining the source of this very recent widespread infection. For this investigation, in addition to isolates from recently infected cattle, farmed deer and wildlife, a selection of isolates from the AgResearch strain archive from sources near the outbreak and of similar types from further away were characterised by WGS in order to determine if this outbreak was from local wildlife reinfection, or from introduction of a different type into the region. Results of this investigation indicate that Mt. Cargill outbreak isolates share recent common ancestors with livestock and wildlife isolates from north of Mt Cargill and are more distantly related to the other examined types that are prevalent in the Mt Cargill region, suggesting that this infection is more likely to have moved into Mt. Cargill from wildlife or livestock from the north than to have come from local wildlife. Although in most cases VNTR types of these isolates correlated well with SNP sub-clusters, since the closest relative, a 2012 Waikouaiti isolate (AgR1664) had a slightly variant VNTR type (Figure 2), if the Mt. Cargill outbreak investigation was based solely on VNTR results the relevance of this isolate to the outbreak would be much less evident than it is from the SNP lineage.

Figure 2. Genetic and spatial relationship of *M. bovis* isolates from southeast Otago. The scale bar in the phylograph indicates the distance in SNPs between isolates, and numbers above the branches of the phylograph indicate the number SNPs common to the indicated cluster. Livestock metadata in the phylograph is coloured red and wildlife metadata is coloured black. Symbols in the map indicate the approximate regional source of the isolate and are coloured to match the genetic cluster of the isolate.

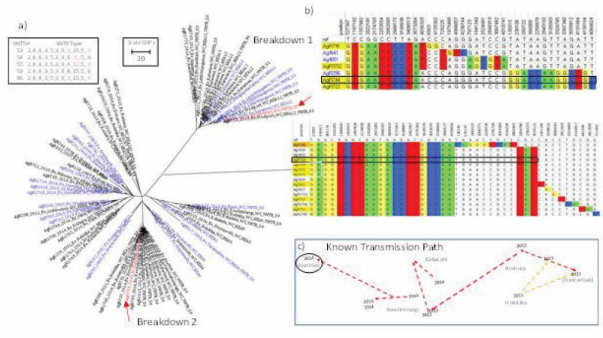


HariHari farm

The investigation of two concurrent breakdowns on the same farm in HariHari clearly demonstrated the close relationship between the breakdown isolates and isolates identified as probable first links in these transmission paths and to wildlife isolates from nearby these probable sources of infection. Although the two types that were detected in this herd could be distinguished from each other by VNTR assay, it was impossible without WGS analysis to determine the source of the two different infections because all of the possible different suspects had identical VNTR types.

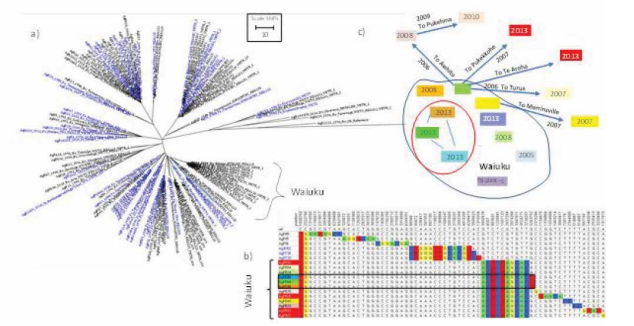
Figure 3. a) Genetic relationship of *M. bovis* isolates from concurrent HariHari herd breakdowns to other related *M. bovis* isolates in the database. Breakdown isolate metadata are coloured red. Livestock metadata are coloured black and wildlife metadata blue. b) The relationship to closest relatives are illustrated in SNP tables with DNA bases coloured to indicate differences from the reference genome. Breakdown isolates are boxed. c) The known transmission path for the source of Breakdown 2 is illustrated with arrows. Links that are supported by the genomic data coloured red.

The isolate from Breakdown 1 was identical to a livestock isolate from Westport (over 170km from HariHari) and to share a recent common ancestor with a wildlife isolate from Karamea (over 250km from HariHari), indicating that this infection resulted from movement of infected wildlife or livestock. The isolate from breakdown two was identical by WGS to isolates from a recent outbreak further up the coast in the Kowhitirangi and Arahura regions. These isolates share a recent common ancestor with a 2011 isolate from a Hokitika farm, but this Hokitika isolate is not the source of the outbreak since it is missing the three SNPs that are common to the outbreak and has additional SNPs not found in the outbreak isolates (Figure 3).



The investigation of the *M. bovis* outbreak that began in Waiuku clearly illustrated the close relationship of epidemiologically linked livestock isolates and demonstrated their distant relationship to other types from the Central North Island. In addition, a SNP detected in isolates from livestock that were not linked by movement but were farmed within 4km of one another (boxed in the SNP table and circled in the transmission path diagram) suggested that despite extensive surveillance, there may have been a wildlife vector for this Waiuku infection.

Figure 4. Waiuku outbreak Investigation. a) Genetic relationship of *M. bovis* isolates from livestock and wildlife in the Central North Island. Livestock metadata are coloured black and wildlife metadata blue. Waiuku isolates are indicated by the bracket. b) The relationship of isolates from the Waiuku outbreak is illustrated in the SNP table with DNA bases in the table coloured to indicate differences from the reference genome. c) The known transmission path with isolates coloured to match the genomic data shown in the SNP table. Isolates that are boxed in b) and circled in c) were not linked by movement but were from farms within 4km of one another.



Discussion

Results of our investigations demonstrated that when analysed by WGS, isolates cluster into the same groups that were determined by REA and VNTR analysis, but with the much finer resolution provided by WGS there is increased ability to rule out likely sources of infection. The regional clustering of types determined with REA and VNTR methods was corroborated since livestock and wildlife from the same region clustered. In addition, WGS analysis has led to the confirmation of epidemiological sourcing of infection, to better definition of the source of new infections by ruling out possible sources, and has indicated probable wildlife infection in an area considered to be vector free. The routine use of WGS analyses for determining the source of *M. bovis* infections will be an important component of the strategy employed to eradicate bovine TB from NZ.

References

1. **Livingstone PG *et al.*** *NZ Vet J* 63 (S1), 4-18, 2015
2. **Buddle BM *et al.*** NZ Vet J 63, p. 1-9, 2015
3. **Collins DM *et al.*** J. Hyg (Lond). 96(3), 431-8, 1986
4. **Collins DM.** *Vet Microbiol* 151(1-2), 2-7, 2011
5. **Price-Carter M *et al.*** *Vet Microbiol* 150(1-2), 107-14, 2011
6. **Kao R *et al.*** *Rev. Sci. Tech* 35, 241-268, 2016
7. **Price-Carter M, de Lisle GW, Collins D.** IDREC Conference Wellington NZ, 2014. www.idrec.ac.nz/idrec-symposium-2014.html
8. **Crispell J *et al.*** *BMC Genomics* 18:180, 2017
9. **de Lisle GW *et al.*** *Vet. Microbiol* 132, 402-7, 2008
10. **Garnier T e*t al.*** *Proc Natl Acad Sci USA*. 100(13), 7877-82, 2003
11. **Ford C *et al*.** *Tuberculosis* (Edinb), 92, 194-201, 2012
12. **Smith *NH.*** *Infect. Genet. Evol* 12, 857-65, 2012