



Application of Four Genotyping Methods to *Mycoplasma bovis* Isolates Derived from Western Canadian Feedlot Cattle

Andrea Kinnear,^a Matthew Waldner,^a Tim A. McAllister,^b Rahat Zaheer,^b Karen Register,^c  Murray Jelinski^a

^aWestern College of Veterinary Medicine, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

^bLethbridge Research and Development Centre, Agriculture and Agri-Food Canada, Lethbridge, Alberta, Canada

^cRuminant Diseases and Immunology Research Unit, USDA Agricultural Research Service, National Animal Disease Center, Ames, Iowa, USA

ABSTRACT *Mycoplasma bovis* is a significant pathogen of feedlot cattle, responsible for chronic pneumonia and polyarthritis syndrome (CPPS). *M. bovis* isolates ($n = 129$) were used to compare four methods of phylogenetic analysis and to determine if the isolates' genotypes were associated with phenotypes. Metadata included the health status of the animal from which an isolate was derived (healthy, diseased, or dead), anatomical location (nasopharynx, lung, or joint), feedlot, and production year (2006 to 2018). Four *in silico* phylogenetic typing methods were used: multilocus sequence typing (MLST), core genome MLST (cgMLST), core genome single nucleotide variant (cgSNV) analysis, and whole-genome SNV (wgSNV) analysis. Using Simpson's diversity index (D) as a proxy for resolution, MLST had the lowest resolution ($D = 0.932$); cgSNV ($D = 0.984$) and cgMLST ($D = 0.987$) generated comparable results; and wgSNV ($D = 1.000$) provided the highest resolution. Visual inspection of the minimum spanning trees found that the memberships of the clonal complexes and clades had similar structural appearances. Although MLST had the lowest resolution, this methodology was intuitive and easy to apply, and the PubMLST database facilitates the comparison of sequence types across studies. The cg methods had higher resolution than MLST, and the graphical interface software was user-friendly for nonbioinformaticians, but the proprietary software is relatively expensive. The wgSNV approach was the most robust for processing poor-quality sequence data while offering the highest resolution; however, application of its software requires specialized training. None of the four methods could associate genotypes with phenotypes.

KEYWORDS beef, feedlot, genotyping, *Mycoplasma*

M*ycoplasma bovis* is associated with a plethora of diseases in cattle. Reviews on this subject commonly emphasize its role in chronic pneumonia, mastitis, and arthritis (1–3). While the incidence of *M. bovis* mastitis in Canada is relatively low (4, 5), it is considered an emerging disease of dairy cattle (6). The incursion of *M. bovis* into New Zealand and that country's considerable efforts to eradicate this organism underscore its potential to cause significant economic losses to the dairy industry. In the feedlot industry, *M. bovis* is associated with bovine respiratory disease (BRD) and chronic pneumonia and polyarthritis syndrome (CPPS) (7). However, its role in the pathogenesis of BRD has been equivocal. In the early 1990s, it was hypothesized that *M. bovis* was unlikely to be a primary pathogen but potentiated the pathogenesis of other bacterial BRD agents (8). It is now considered a primary pathogen of pneumonia in preweaned calves (9, 10) and an opportunistic pathogen of BRD in feedlot cattle (1, 2, 7). Cattle are frequently asymptomatic carriers; the organism is recovered from the nasal passages of healthy cattle (11–13). It is hypothesized that stressors such as transport, commingling, and adverse climatic conditions trigger a stress response, resulting in elevated

Citation Kinnear A, Waldner M, McAllister TA, Zaheer R, Register K, Jelinski M. 2021. Application of four genotyping methods to *Mycoplasma bovis* isolates derived from western Canadian feedlot cattle. J Clin Microbiol 59:e00044-21. <https://doi.org/10.1128/JCM.00044-21>.

Editor Brad Fenwick, University of Tennessee at Knoxville

© Crown copyright 2021. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Murray Jelinski, murray.jelinski@usask.ca.

Received 7 January 2021

Returned for modification 12 February 2021

Accepted 19 April 2021

Accepted manuscript posted online

5 May 2021

Published 18 June 2021

levels of glucocorticoids that may impair the immune response, allowing for increased bacterial shedding and clinical disease (14).

A number of molecular techniques have been used for typing *M. bovis* isolates. The formative techniques included pulsed-field gel electrophoresis (PFGE) (15–17), random amplified polymorphic DNA (RAPD) (15, 18), and amplified fragment length polymorphism (AFLP) (15, 19). These methods have been largely replaced by PCR-based methods such as multilocus sequence typing (MLST) and multiple-locus variable-number tandem-repeat analysis (MLVA) (20–22). While MLVA has greater discriminatory power than MLST (21), the latter method is an unambiguous, reproducible, and scalable procedure for characterizing isolates of bacterial species by using a universally acceptable schema (23) (<https://pubmlst.org/>). These attributes make MLST well suited for epidemiological studies spanning multiple research laboratories and for the comparison of isolates over time (21, 24–26).

Core genome MLST (cgMLST) is an extension of MLST but provides higher resolution. Whereas a typical MLST scheme uses the alleles of seven housekeeping genes, cgMLST analyzes alleles from hundreds of genes, an approach facilitated by whole-genome sequencing (WGS). cgMLST has been used for typing poultry mycoplasmas (27), investigating outbreaks of *M. bovis* in dairy cattle (28, 29), and examining the genetic relatedness and evolution of isolates obtained from cattle in Denmark and neighboring countries (30). The authors of the Denmark study noted that the cgMLST and WGS typing techniques offered greater discriminatory power than MLST and hence may become the new standard in phylogenetic typing. However, these methods do have disadvantages, namely, cost, time, and the need for technical expertise for conducting the analyses and interpreting the results.

A further progression into higher-resolution strain typing involves genome-wide comparisons of single nucleotide variants (SNVs). This method can be applied to SNVs in the core genome (cgSNVs) or the whole genome (wgSNVs). Using WGS, Australian researchers concluded that 75 *M. bovis* isolates collected between 2006 and 2015 were of the same lineage, suggesting few, if any, incursions of new strains over the study duration. Similarly, an Israeli study used cgSNV analysis to evaluate the genomic diversity of *M. bovis* isolates from mastitis cases between 1994 and 2017 and compared these to BRD isolates from local feedlot cattle and from calves imported from Europe and Australia (31). There was a clear genetic distinction between the isolates from Europe and Australia, and a dominant genotype was associated with mastitis. wgSNV analysis has also been used to compare the relatedness of 250 *M. bovis* isolates originating from seven countries (32). These isolates formed six clades, with U.S. isolates exhibiting the greatest genetic diversity but also clustering with Canadian isolates.

The objectives of this study were, first, to assess the level of concordance between four different molecular genotyping methods (*in silico* MLST, cgMLST, cgSNV, and wgSNV) by using a data set of 129 *M. bovis* isolates and, second, to determine if one or more methods could resolve genotypic differences among isolates derived from cattle of different health statuses (healthy, sick, and dead cattle), from different anatomical locations (nasopharynxes, lungs, and joints), from different feedlots, and over a 12-year period (2006 to 2018).

MATERIALS AND METHODS

Sample collection. A series of cross-sectional studies, spanning the years 2006 to 2018, provided 129 *M. bovis* isolates, which were recovered from the nasopharynxes, lungs, and joints of feedlot cattle, as described previously (33). Five isolates were recovered from cattle imported from Idaho, USA, while all others were from cattle resident in western Canada. The deep nasopharyngeal (DNP) swabs were obtained from healthy and morbid cattle. Each animal's health status was determined by the timing of the disease and the presentation of clinical signs. BRD is the most common disease of feedlot cattle, and cases peak within 21 days after arrival at the feedlot. Thus, a putative BRD diagnosis was based on the timing of the disease and a constellation of the clinical signs consistent with this disease (i.e., fever, depression, nasal discharge, dyspnea). Health status was determined by trained feedlot personnel. The DNP swabs from live cattle were obtained in accordance with Animal Use Protocols (no. 20070023 and 20170021) approved by the University of Saskatchewan's Animal Research Ethics Board and Lethbridge Research Centre's Animal Care Committee (no. 1641). Lung and joint samples were obtained at the time

of postmortem examination from cattle with gross pathological findings consistent with caseonecrotic bronchopneumonia and, in some instances, concurrent septic arthritis.

Culture and isolation. Small changes in media and isolation methods occurred over the 12-year period; the isolates recovered in 2006 to 2008 were cultured in Hayflick medium (prepared in-house) (34), and all subsequent isolates were cultured with pleuropneumonia-like organism (PPLO) broth and agar (BD Difco, Fisher Scientific, Waltham, MA, USA). The PPLO media were supplemented with 10 g/liter yeast extract (BD Diagnostic Systems, Fisher Scientific) and 20% horse serum (Invitrogen, Fisher Scientific), as described previously (33). Supplemented media also contained 0.05% thallium(I) acetate, 500 U/ml penicillin G, and/or 0.5% sodium pyruvate (Sigma-Aldrich, St. Louis, MO, USA).

DNP samples and swabs of fresh-cut tissues were used to inoculate PPLO starter cultures. Cultures were serially filtered through 0.45- and 0.20- μ m-pore-size filters (Basix; VWR International, Radnor, PA, USA) to remove other bacteria, such as coinfecting BRD pathogens (0.45 μ m), and to select for *Mycoplasma* spp. (0.20 μ m). Filtrates were inoculated into supplemented PPLO broth and were grown under a 5% CO₂ atmosphere with 75% humidity at 37°C. Culture growth was visualized by agitating the culture tube to elicit a perceptible mass of organisms at the bottom of the tube. Cultures with visible growth were subcultured onto PPLO agar and incubated for 3 to 6 days. Single colonies exhibiting a "fried-egg" morphology were picked, plated on PPLO agar, and incubated for 72 h. One to three individual colonies per culture were used to inoculate separate aliquots of PPLO broth. After 48 h, each culture was stored in PPLO medium with 20% (vol/vol) glycerol at -80°C. A single culture was chosen to inoculate the PPLO medium for DNA extraction.

DNA extraction and identification. Isolates were grown in PPLO medium for 48 h, and genomic DNA was extracted using the GenElute Bacterial Genomic DNA kit (Sigma-Aldrich) according to the manufacturer's instructions, except that the final elution buffer was replaced with 10 mM Tris (pH 8.5). The extracted genomic DNA was assessed for quality using gel electrophoresis and was quantified fluorometrically using a Qubit analyzer (Thermo Fisher Scientific). The isolation of high-molecular-weight DNA with a yield of ≥ 1 ng/ μ l was sufficient to proceed. Cultures were confirmed as *M. bovis* by using a species-specific PCR assay targeting the *uvrC* gene (35) and by sequencing the V3-V4 16S rRNA gene (36). The 16S rRNA amplicon was purified using a QIAquick PCR kit (Qiagen, Venlo, Netherlands) and was submitted for Sanger sequencing (Macrogen, Seoul, South Korea). Forward and reverse sequences were assembled and edited using the Staden package (version 1.6-r; <http://staden.sourceforge.net/>) and were compared to the National Center for Biotechnology Information (NCBI) nonredundant (nr) nucleotide database using BLASTn run with default settings. The species was assigned based on the highest match identity (37).

Whole-genome sequencing and assembly. Sequencing libraries of genomic DNA were prepared using an Illumina Nextera XT DNA library preparation kit (Illumina Inc., San Diego, CA, USA) and were sequenced on an Illumina MiSeq platform using the MiSeq v2 reagent kit to generate 250-bp paired-end reads. Genomes were assembled for the MLST, cgMLST, and cgSNV methods using Ridom SeqSphere+ in pipeline mode (38). Raw paired-end reads were imported and preprocessed by down-sampling to 180 \times coverage and trimming at the 5' and 3' ends until an average quality of 30 in a window of 20 bases was achieved. Reads were assembled using SKESA (39). Genome assembly for the wgSNV method was performed using Trimmomatic, v0.39 (40), for read trimming and SPAdes, v3.14.1 (41), for *de novo* assembly of the contigs. Trimmomatic settings were as follows: sliding window, 5:15; leading, 5; trailing, 5; minlen, 25. SPAdes was run with settings -careful and -k 127. Contigs with <1,000 nucleotides were removed from the analysis.

Genotyping methods. *De novo* assemblies were used to assign allelic profiles and sequence types (STs) as per the PubMLST reference method (<https://pubmlst.org/organisms/mycoplasma-bovis/>). The *M. bovis* PG45 reference genome (GenBank accession no. NC_014760.1) was included in each genotyping method with the 129 isolates. The MLST scheme included alleles of the following genes: *dnaA*, *gltX*, *gpsA*, *gyrB*, *pta-2*, *tdk*, and *tkl* (25; version update, 15 March 2021).

The same *de novo* assemblies were used to develop an *ad hoc* cgMLST scheme using Ridom SeqSphere+ (version 6.0.2) (38). The reference strain *M. bovis* PG45 (GenBank accession no. NC_014760.1) was used as the seed genome with the following criteria: minimum length, ≥ 50 bases; start and end codons on each end of the gene; no multiple copies of genes with a BLAST overlap of ≥ 100 bp with an identity of $\geq 90\%$; and no overlap with genes of >4 bases. Genes identified in the seed genome were queried against the following 10 penetration genomes in order to identify genes for inclusion in the final *ad hoc* cgMLST scheme: *M. bovis* Hubei-1 (GenBank accession no. NC_015725.1), HB0801 (NC_018077.1), CQ-W70 (NZ_CP005933.1), NM2012 (NZ_CP011348.1), HB0801-P115 (NZ_CP007589.1), 08M (NZ_CP019639.1), Ningxia-1 (NZ_CP023663.1), JF4278 (NZ_LT578453.1), 16M (NZ_CP038861.1), and XBY01 (NZ_CP045797.1). Penetration genomes were queried using BLAST (version 2.2.12) and were required to have equivalent targets that met the BLAST hit overlap of 100% with an identity of $\geq 90\%$ in all query genomes. The following criteria were used: word size, 11; mismatch penalty, -1; match reward, 1; gap open costs, 5; and gap extension costs, 2. Targets were also required to have a single stop codon at the end of the gene in $>80\%$ of penetration query genomes. The resulting cgMLST scheme consisted of 506 genes (loci) and covered 55.1% of the *M. bovis* PG45 genome.

The distance matrix used for cgMLST phylogenetic analysis omitted genomes missing $>10\%$ of distance columns and removed columns with missing values. This resulted in isolates being typed based on the alleles of 296 loci.

SNVs for the cgSNV method were determined from 283 gene targets (loci) of the *M. bovis* PG45 reference genome (NC_014760.1). Comparison of all genomes to these targets yielded 6,408 SNV positions, which were filtered to 3,925 SNVs in 283 loci by including only substituted SNV positions (hiding

insertions/deletions) and having no neighboring SNV positions in a window of 10 bases. SNVs for the wgSNV method were identified within the 130 genomes by kSNP, v3.1 (42), which yielded an SNV matrix file. The settings used for kSNP were `-core` and `-k 31`. The SNV matrix contained 14,383 SNVs across all genomes.

Simpson's diversity index (43) compared the discriminatory power of each strain-typing method, based on the clustering of isolates within the minimum spanning trees (MSTs) for the MLST, cgMLST, and cgSNV methods. The wgSNV method did not cluster any two of the isolates into a single type. Therefore, each isolate was classified as a unique type, as defined by a unique genotype for each isolate.

Data presentation. Neighbor joining (NJ) trees were created with iTOL (44) and MSTs with Ridom SeqSphere+. A maximum-likelihood tree of the wgSNV matrix was generated with the Tamura-Nei substitution model, using uniform rates and the nearest-neighbor-interchange heuristic method with MEGA X, v10.1.1 (45). The tree was visualized with iTOL and GrapeTree (46). Isolates were grouped into clonal complexes (CC) or clades (C), as appropriate. A CC was defined as a group of isolates with STs that differed by no more than two alleles from at least one other ST in the group. A singleton was a clonal group that differed from all other STs by at least three alleles. A clade was defined as a group of strains with a common biological ancestor. The CC were assigned by MLST analyses, while clades comprised the MSTs of the cgMLST, cgSNV, and wgSNV analyses. These MSTs were determined by visual assignment based on the root of a tree or an ST central to the tree that served as a common ancestor.

Data availability. The raw paired-end reads generated in this study are available from the Sequence Read Archive (SRA) under BioProject accession no. [PRJNA642970](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA642970) and [PRJNA708306](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA708306).

RESULTS

Isolates and assembly. *M. bovis* isolates ($n = 129$) spanning 12 production years (2006 to 2018) and 21 feedlots (A to U) were recovered from 98 individual feedlot cattle (see Table S1 in the supplemental material). All 21 feedlots were located in western Canada; 45.0% ($n = 58$) of the isolates originated from two feedlots (N, Q), each with a capacity of >20,000 head. Isolates were recovered from the nasopharynxes ($n = 49$), lungs ($n = 45$), and joints ($n = 35$). The numbers of isolates by animal health status were as follows: 82 from dead cattle, 32 from healthy cattle, and 15 from diseased cattle. DNP isolates ($n = 49$) were recovered from healthy ($n = 32$), diseased ($n = 15$), and dead ($n = 2$) cattle. All isolates ($n = 129$) underwent WGS and *de novo* assembly with the following criteria: mean N_{50} , 18,448 (range, 997 to 32,908); contig count, 186 (range, 79 to 797); coverage, 84 (range, 12 to 177); approximate completed genome size, 90% (range, 60% to 100%) relative to the PG45 reference genome.

MLST. Of 130 genomes (129 isolates plus the PG45 reference genome), 126 were assigned an existing ST. Four isolates could not be typed due to a missing allele(s) and are designated as ST "Unknown" (one also had a novel allele at the *pta-2* locus), which is perhaps a reflection of low-quality assemblies (N_{50} , 997 to 2,494; contig count, >500; coverage, 12 to 51; approximate genome size, $\leq 80\%$). The MLST scheme typed the 125 isolates into 24 known STs and 6 newly identified STs (ST149 to ST154) (Table 1). The PG45 reference genome included in the analysis was assigned ST12, as expected. Simpson's diversity index was 0.932, indicating reasonably strong separation of isolates. Two previously uploaded isolates that had been assembled using SPAdes (41) had MLST STs differing from those assigned by SKESA (39) and Ridom SeqSphere+. Two of these were within the set of four isolates that had missing alleles. These differences are likely due to low-quality isolates and alternate assembly processes.

The most prevalent STs were ST60 (23 of 130 isolates [17.69%]), ST24 (13 of 130 [10.00%]), and ST67 (12 of 130 [9.23%]). These prevalences were higher than those reported in the PubMLST database (<https://pubmlst.org/organisms/mycoplasma-bovis/>, accessed 12 March 2021): 12 of 1,139 isolates (1.05%) in ST60, 9 of 1,139 isolates (0.79%) in ST24, and 9 of 1,139 isolates (0.79%) in ST67. Among the 23 ST60 isolates, 12 had been used in another, unrelated study (25) and are represented in the PubMLST database, but under isolate identifiers (see Table S1). The STs with the highest frequencies in the PubMLST database were ST52 (127 of 1,139 isolates [11.15%]), ST62 (71 of 1,139 isolates [6.23%]), and ST21 (31 of 1,139 isolates [2.72%]), all of which are frequently identified in North America but were infrequent in the current study.

Isolates ($n = 35$) recovered from feedlot N between the years 2016 and 2018 were represented by nine STs, of which ST60 ($n = 10$ [28.6%]) was the most prevalent in all years. The 22 isolates recovered from feedlot Q in 2007 were categorized into six STs; 12 (54.5%) were in ST24. Two STs (ST2 and ST21) were of particular interest because

TABLE 1 Numbers of MLST sequence types of *Mycoplasma bovis* isolates by production year and health status^a

Sequence type	No. of isolates											
	Production yr								Health status			Total
	2006	2007	2008	2014	2015	2016	2017	2018	Healthy	Diseased	Dead	
2		2				1				1	2	3
14							1		1			1
21	1	4					2		2	3	2	7
24		12			1				5	6	2	13
27								3			3	3
40		5							3	2		5
42	3	2		1	1				3	2	2	7
43								1	1			1
44							9				9	9
45							3	2	4		1	5
48		1							1			1
52		1				3					4	4
60				1		13	1	8			23	23
61						1					1	1
62		1							1			1
65						6	1		1		6	7
66							2		2			2
67							12				12	12
70		1			2					1	2	3
75			1								1	1
76				2							2	2
77						2					2	2
79						1					1	1
80						3					3	3
149								1			1	1
150								1	1			1
151	1								1			1
152							1		1			1
153								2	2			2
154							2		2			2
Total	5	29	1	4	4	30	34	18	31	15	79	125

^aA total of 125 *M. bovis* isolates were tested. PG45 (ST12) is excluded from this table.

they were separated in time and space. ST2 was isolated from feedlot Q in 2007 and from feedlot N in 2016, while ST21 was recovered from feedlot Q in 2007 and from feedlot N in 2017. These feedlots were separated by a distance of approximately 500 km. Five isolates were derived from cattle imported from the northern United States.

Isolates clustered into two clonal complexes (CC1, CC2) and as four singletons, including the PG45 reference strain (ST12, ST42, ST43, ST75) (Fig. 1). The ST52 and ST60 isolates formed the foci of the CC in the MLST minimum spanning tree (Fig. S1). Two isolates (MPLM0830, MPLM0608) from 2018 had allelic combinations that had not been described previously. Five STs persisted within the western Canadian cattle population for many years: ST2 (2007 to 2016), ST21 (2006 to 2017), ST24 (2007 to 2015), ST42 (2006 to 2015), and ST70 (2007 to 2015). Three of these STs (ST2, ST21, ST24) were grouped in CC2, whereas ST70 was grouped with CC1 and ST42 was a singleton.

The data set included 31 pairs of lung-joint samples obtained from individual animals; for 28 lung-joint pairs, both isolates were successfully typed using MLST. In 18 (64.3%) instances, the ST recovered from the lung matched the ST found in the joint of the same animal.

cgMLST and cgSNV analyses. A total of 102 genomes (101 isolates and PG45) were typed by the cgMLST (Fig. 2; Fig. S2) and cgSNV (Fig. 3; Fig. S3) methods. Figures S2 and S3 provide the MSTs of the isolates, with three clades (C1, C2, C3) branching from a single focus consisting of an isolate with MLST ST62. Isolates from these clades tended to cluster together in the neighbor joining (NJ) tree, but small subclades (Fig. 2 and 3) branched from the root and contained isolate-specific clades in the MSTs. Both

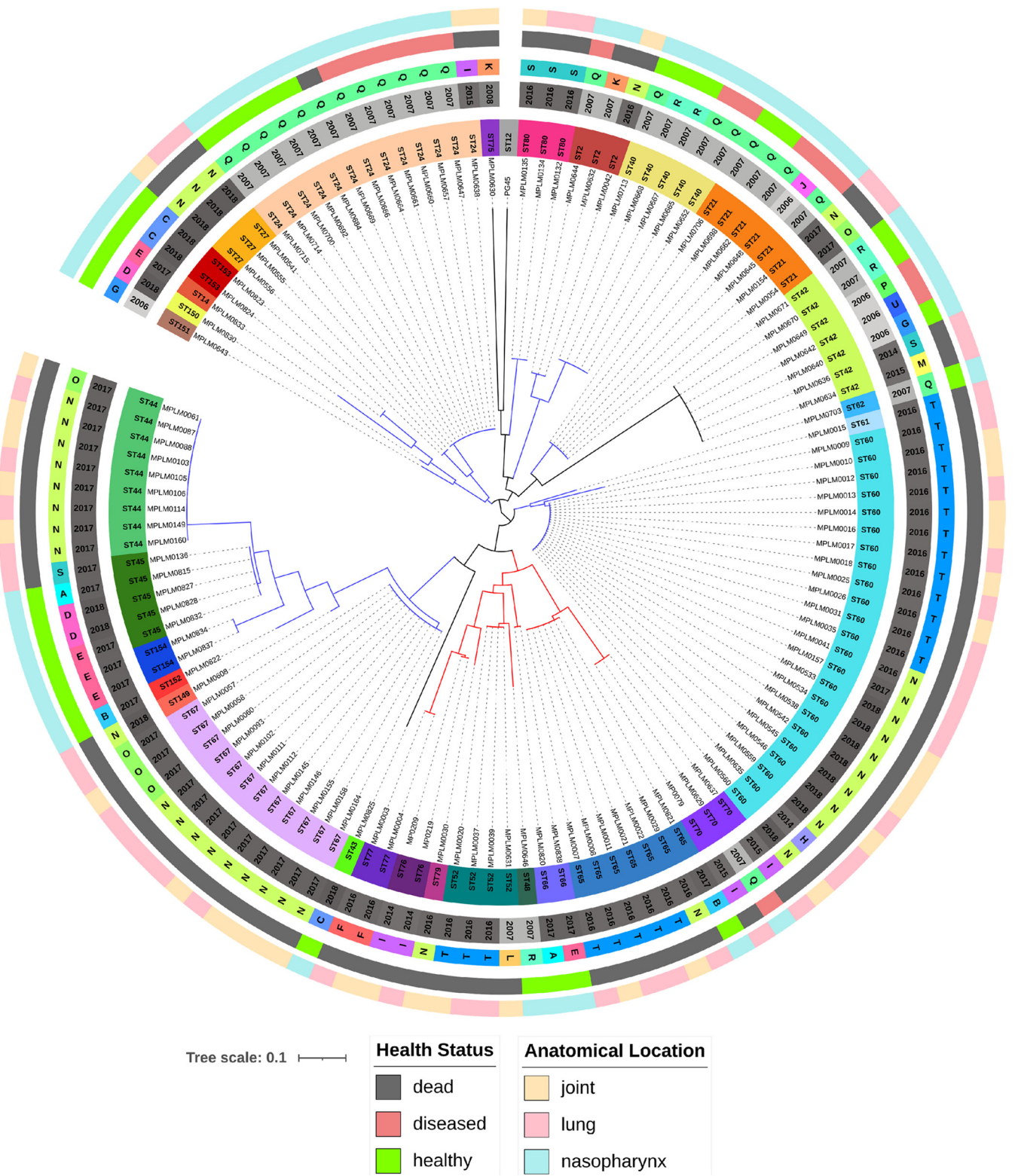


FIG 1 MLST neighbor joining tree of 126 *Mycoplasma bovis* isolates typed by MLST. The rings (starting from the innermost) contain information on isolate identifiers, MLST sequence type (ST), production year, feedlot (A to U), anatomical sampling location, and animal health status at the time of sampling. Two clonal complexes are depicted with red (CC1) and blue (CC2) branch lines. Four singletons (ST12, ST42, ST43, ST75) are depicted as black branch lines.

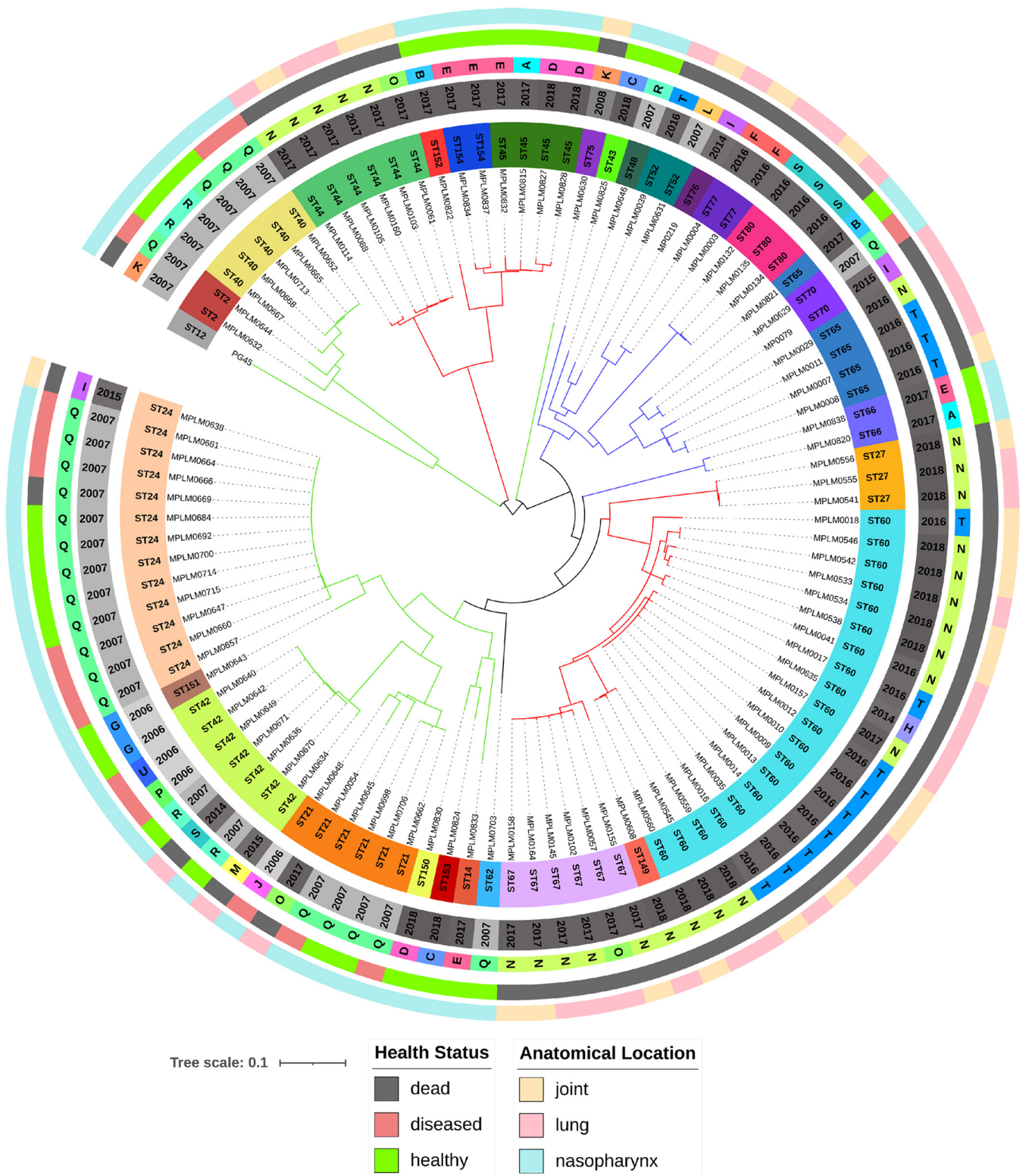


FIG 2 cgMLST neighbor joining tree of 102 *Mycoplasma bovis* genomes based on alleles at 296 core genome loci. The rings (starting from the innermost) contain information on isolate identifiers, MLST sequence type (ST), production year, feedlot (A to U), anatomical sampling location, and animal health status at the time of sampling. The branches are colored in accordance with the three clades (C1, red; C2, blue; C3, green) identified in Fig. S2, highlighting the differences between the tree construction methods.

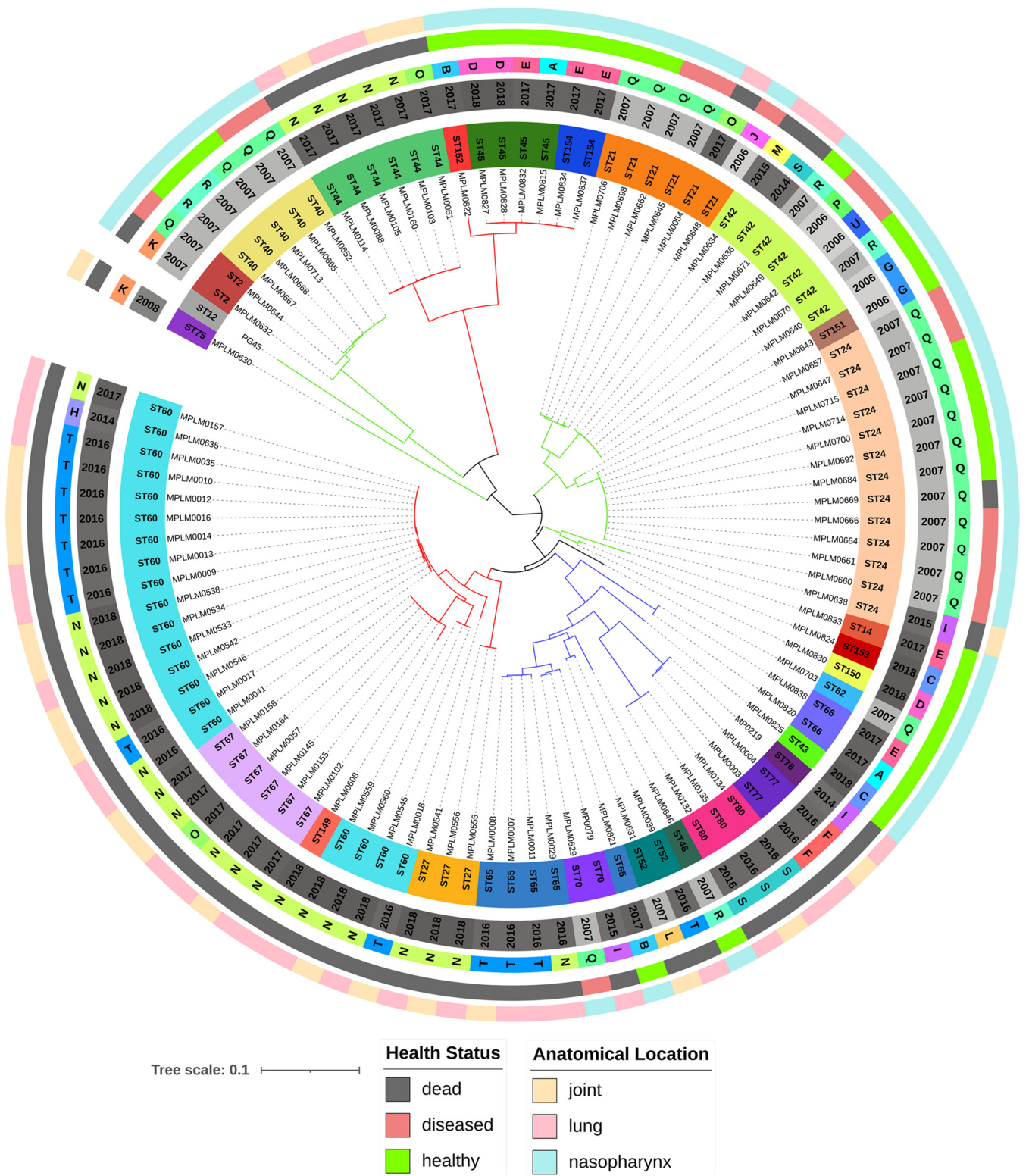


FIG 3 cgSNV neighbor joining tree of 102 *Mycoplasma bovis* isolates based on 283 core genome loci (3,925 SNVs). The rings (starting from the innermost) contain information on isolate identifiers, MLST sequence type (ST), production year, feedlot (A to U), anatomical sampling location, and animal health status at the time of sampling. The branches are colored to match the clades (C1, red; C2, blue; C3, green) identified in Fig. S3, highlighting the differences between the tree construction methods.

cg analyses grouped the five STs discussed above (ST2, ST21, ST24, ST42, and ST70) into two clades: ST2, ST21, ST24, and ST42 were grouped in C3, and ST70 was allocated to C2 (Fig. S2 and S3). Although minor differences in overall clade membership were observed between the MLST method and the two cg methods, the results were fairly consistent. Simpson's diversity index was 0.987 for the cgMLST and 0.984 for the cgSNV analysis, indicating strong separation of the isolates into individual STs, with most isolates assigned a unique ST.

wgSNV analysis. All 130 genomes were typed by the wgSNV method, and no two isolates shared an identical SNV matrix; hence, each isolate was unique (Fig. 4; Fig. S4). As a result, Simpson's diversity index was 1.000. Two primary clades (C1, C2) branched into two subclades of approximately equal size. A single isolate (ST62) was positioned evenly between the two clades. The wgSNV method was able to resolve isolates with the same MLST ST assignment as the cgMLST and cgSNV methods. Like the cgMLST and cgSNV methods, the wgSNV method grouped ST2, ST21, ST24, and ST42 in C1, while the ST70 isolates grouped with a separate lineage (C2) (Fig. S4).

DISCUSSION

This is the first study to compare four different phylogenetic typing methods using a relatively large and diverse set of *M. bovis* isolates derived from western Canadian feedlot cattle. Overall, more genomes were typed by the MLST scheme ($n = 126$) than by the cgMLST or cgSNV method ($n = 102$), which required the assembly of hundreds of loci constituting a large portion of the genome. The wgSNV method successfully typed all 130 genomes, highlighting the robustness of the SPAdes and kSNP software, even with inputs of variable sequencing depth. All four methods had a high degree of discriminatory power, as judged by Simpson's diversity index. MLST had the lowest discriminatory power, while cgMLST had a modestly higher index than the cgSNV method. Despite some differences in the phylogenetic outputs of the four methods, the MLST method generated a pattern of clonal complexes (CC) comparable to the clades generated by the cgMLST, cgSNV, and wgSNV methods. This is noteworthy, since the MLST scheme relied on seven housekeeping genes, representing <1% of the *M. bovis* PG45 reference genome. In contrast, the cgMLST scheme was derived from 506 loci, covering approximately 55% of the genome. The cgSNV method analyzed 3,925 SNVs from 283 loci, while the wgSNV matrix utilized 14,383 SNVs, generating the highest genotypic resolution.

Since core genome and whole-genome methods are based on a larger representation of the genome, theoretically they should have greater potential to resolve relationships than the seven-locus MLST scheme. This makes the cgMLST, cgSNV, and wgSNV methods ideally suited for epidemiological investigations where small differences between STs may be consequential. Parker et al. applied wgSNV analysis to 75 Australian *M. bovis* isolates and found a very high level of homogeneity among the isolates; the maximum number of single nucleotide polymorphisms (SNPs) between any two isolates was 50 (47). This level of resolution led the researchers to conclude that a single strain of *M. bovis* was circulating within Australia's cattle population. This is quite unique, since a number of other country-level studies have found multiple clusters of genetically distant *M. bovis* STs within the cattle population (30–32, 48). Furthermore, these studies provided some insight into the movement of specific STs from country to country and over time. These higher-resolution methods have also been applied at the farm level, to examine transmission between cows and calves within the same dairy farm (29) and the introduction of *M. bovis* into dairy herds via contaminated semen (28).

The higher concordance and resolution of the cgMLST, cgSNV, and wgSNV methods were evident when the comparison focused on five STs (ST2, ST21, ST24, ST42, ST70) representing 33 isolates obtained between 2006 and 2017. The MLST scheme assigned three STs (ST2, ST21, ST24) to clonal complex 2 (CC2), ST70 to CC1, and ST42 as a singleton, whereas both the cgMLST and cgSNV analyses assigned four STs (ST2, ST21, ST24, ST42) to clade 3 (C3) and ST70 to C2. The wgSNV analysis yielded results similar

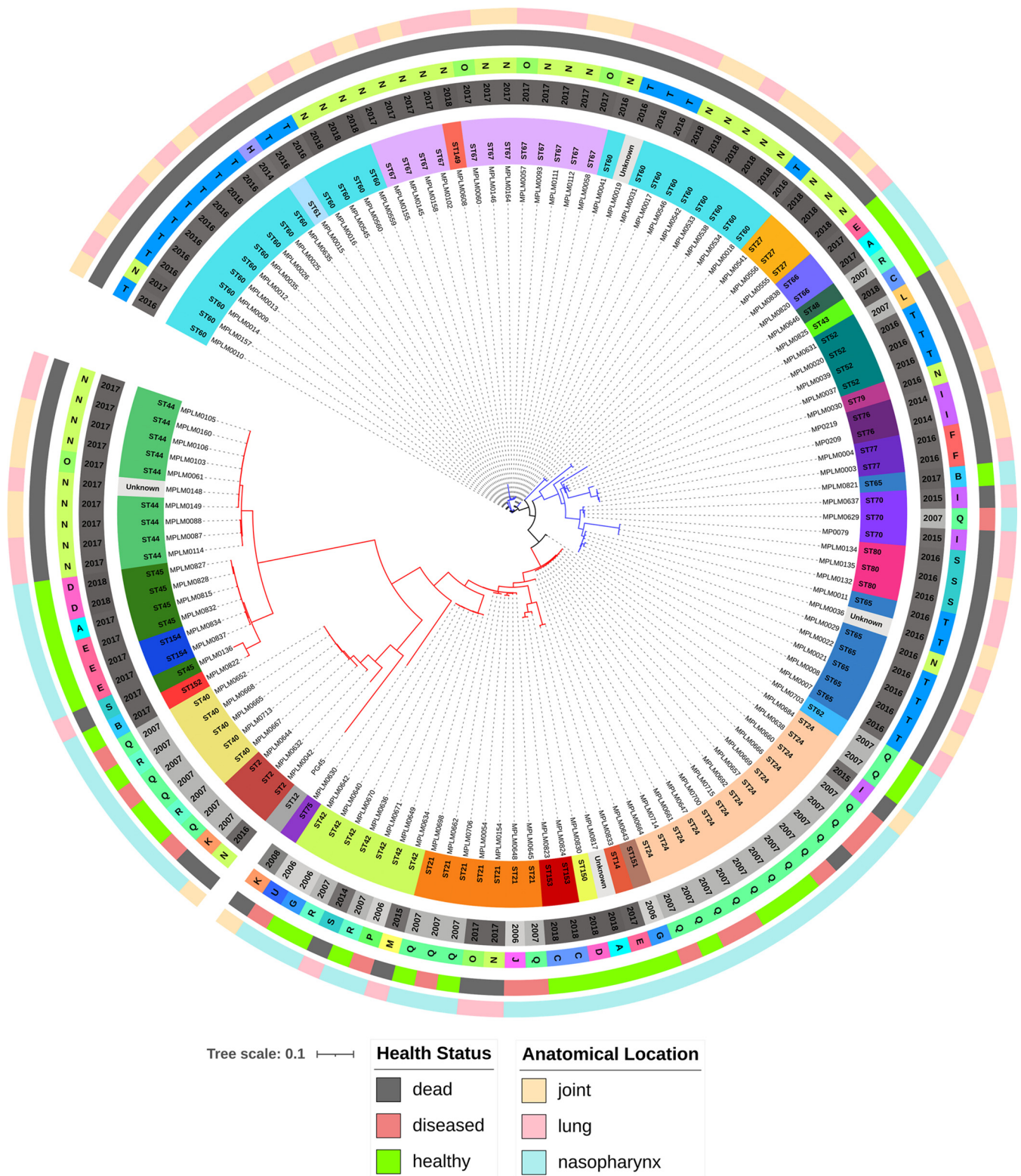


FIG 4 wgSNV maximum-likelihood tree of 130 *Mycoplasma bovis* genomes based on 14,383 SNVs in the core and accessory genomes. The rings (starting from the innermost) contain information on isolate identifiers, MLST sequence type (ST), production year, feedlot (A to U), anatomical sampling location, and animal health status at the time of sampling. Clades 1 and 2 are indicated by blue and red, respectively. Isolate MPLM0703 (ST62) falls evenly between these two clades and, as a result, was not included in either. These clades match the clades identified in Fig. S4.

to those of the cgMLST and cgSNV methods. These findings underscore that all four methods had similar assignments for four of the five STs. However, because ST42 was a triple locus variant, it became an outlier or singleton. Since the cgMLST, cgSNV, and wgSNV methods utilize hundreds to thousands of discrete data points to compare the genetic relatedness of isolates, they grouped ST42 alongside other STs into a clade.

The MLST method identified 30 different STs, underscoring the genetic diversity of *M. bovis* in western Canada. This is most likely related to the underlying structure of the Canadian cattle industry and the way in which feedlot cattle are procured. At the time of the last agricultural census (2016), Canada had approximately 54,000 beef cattle farms, 38,700 (72%) of which were located in western Canada (<https://www150.statcan.gc.ca/n1/en/type/data?text=40221>). Furthermore, western Canadian feedlots also import cattle from the United States. Since most feeder cattle are sold at auction, extensive commingling of cattle from multiple owners occurs during procurement. Once sold, the cattle are assembled, transported, and then processed and further commingled at the feedlot. This also occurs in the autumn months, when inclement weather conditions arise. Conceivably, commingling, transport, and changing environmental conditions all contribute to stress, which facilitates increased shedding from carrier animals (14). Thus, given the broad catchment from which cattle are sourced and mixed, it is understandable that feedlots had multiple STs circulating during the same time period. Interestingly, ST2 was isolated from two feedlots in 2007 (feedlots Q and K) and then, 9 years later, from feedlot N. Similarly, ST21 was isolated from feedlot J in 2006, from feedlot Q in 2007, and then from feedlots N and O in 2017. Not only were these isolates separated in time, but feedlots N and Q were located approximately 500 km apart. This separation by time and space suggests that some STs may be more dominant and widespread than others.

While Canada's feedlot sector is concentrated in Alberta, there is bilateral trade in cattle between Canada and the United States, which is noteworthy because five isolates from American cattle were evenly distributed among the Canadian isolates within the NJ trees and MSTs (identified in Table S1 in the supplemental material). These results support the findings of a recent study in which wgSNV analysis found a high degree of genetic diversity among the American isolates, with Canadian isolates clustering within the same clade as the American isolates (32).

MLST identified 30 STs dispersed over 12 production years. Among these, two strains (ST21, ST52) had been reported in bovine isolates outside North America: ST21 was reported in Europe and Asia, whereas ST52 was reported in Europe, Asia, and Oceania (PubMLST isolate database [<https://pubmlst.org/organisms/mycoplasma-bovis/>], accessed 12 March 2021). This worldwide distribution underscores the international trade in cattle and the need for biosecurity measures to mitigate the transmission of *M. bovis* and other potentially production-limiting pathogens. It is also noteworthy that ST21 has also been isolated from bison (PubMLST isolate database, accessed 12 March 2021).

WGS allows typing by multiple methods to be done *in silico*, and Ridom Seqsphere+ makes high-resolution typing methods accessible to those without the knowledge required to construct a customized whole-genomic-analysis pipeline. Established typing methods, such as MLST, will invariably continue to support *in silico* efforts to conduct comparative studies using historical and contemporary data. WGS of isolates, particularly when analyzed with Ridom Seqsphere+, provides the opportunity to merge established methods (MLST) with more-robust core genome (i.e., cgMLST and cgSNV) approaches. Additionally, given that the cost of WGS is comparable to that of sequencing seven PCR amplicons, particularly for a small, 1-Mbp genome such as that of *M. bovis*, WGS is likely to become the standard for phylogenetic studies. Furthermore, the ability to generate high-quality *M. bovis* assemblies from long reads in a cost-effective manner will enable greater use of cgMLST, cgSNV, and wgSNV phylogenetic typing methods (49).

A caveat for using the cg methods is the need for greater sequencing depth and fewer sequencing artifacts in order to generate more-complete, higher-quality assemblies. This was evident from the fact that more isolates were typed by MLST than by

the cgMLST or cgSNV method, while the wgSNV software pipeline was able to generate a phylogenetic tree for all isolates. The wgSNV approach was better able to process poorer-quality sequencing data because it analyzes the entire genome to a greater degree than the other methods. However, it is not without its weaknesses. The assembly software may have erroneously assembled the small subsets of the genome, resulting in false SNPs contributing to the uniqueness of the genotypes. However, misassemblies occur only infrequently and in sections of an assembly with poor coverage, making this occurrence in our data set unlikely. Care must also be taken in choosing analysis software and associated input parameters. This is exemplified in the NJ trees and MSTs, which were similar, but with differences in the positions of ST2, ST40, ST44, ST45, and ST75. This highlights the need for high-quality sequence data as the analysis moves from MLST to more-complex methods, such as the cgSNV or wgSNV method.

The design of this study was appropriate for comparing the four genotyping methods. However, the results were equivocal with regard to whether the lack of association between genotypes (STs) and phenotypes (year, health status, anatomical location) was real or was related to the limited number of isolates collected. Obtaining a complete set of DNP swabs, lung samples, and joint samples from each animal would have helped in determining whether STs exhibited a tropism for specific tissues. On this point, it was salient that the data set included paired lung-joint samples from 28 animals, 18 (64.3%) of which had the same ST in the lung and joint, suggesting the absence of single-tissue tropism. However, in 10 cases, the genotypes of lung and joint isolates within the same animal differed, underscoring the need for polyvalent vaccines. Others have also suggested that a polyvalent vaccine maybe required to cover the broad functional diversity of isolates (32).

The lack of association between genotype and phenotype is certainly not unique; rather, it is the emerging consensus. Multiple studies using high-resolution typing methods have been unable to show linkages between clusters and anatomical sample locations (47, 48) or health status (31, 47). However, one of the issues is that these studies have not been specifically designed to investigate these associations. This is problematic, since the lack of association may be related not only to an inadequate number of samples but also to an unbalanced study design. Many of the studies have biases in sampling related to year, anatomical location, geographical location, and health status. These confounding factors may result in type I and II errors. This is particularly true when one is investigating the association between genotype and health status in feedlot cattle. Animals deemed healthy on arrival and at the time of sampling may develop BRD within days. Conversely, BRD is a polymicrobial disease, and hence, clinical disease does not equate to mycoplasmosis. This conundrum is best addressed by comparing isolates from healthy animals with those obtained at the time of postmortem examination from tissues (lungs and joints) exhibiting pathology consistent with mycoplasmosis.

Each genotyping method has strengths and weaknesses depending on the research question. MLST is best suited for use as an initial screening method to detect the presence of genetically distinct strains and is amenable to both PCR and *in silico* methods. Furthermore, the *M. bovis* PubMLST database is curated and is accessible through a publicly available website. While cgMLST, cgSNV, and wgSNV analyses allow for the typing of strains, they also provide a higher level of genetic resolution, which may be used to discern clinically relevant differences, such as tissue tropisms, antimicrobial resistance, or virulence.

Conclusion. The wgSNV method successfully typed all 129 field isolates, whereas the cg methods typed 101 isolates, an outcome that may be rectified with greater sequencing depth. Overall, assignments of clade membership by MLST and the higher-resolution methods were similar on visual assessment of the NJ trees and MSTs; the cgMLST and cgSNV methods had the highest degree of concordance. The wgSNV approach provided an incrementally higher level of genomic resolution and detail, which may have utility in some epidemiological investigations and in addressing

research questions relating to gene function and characterization. Although the wgSNV method was very powerful and robust, it is less user-friendly, requiring specialized training in bioinformatics. Conversely, the cg analyses were performed using Ridom SeqSphere+, which is graphical interface software that nonbioinformaticians with moderate background knowledge can use. Thus, it provides greater resolution than MLST and requires less specialized training, but it is relatively expensive if it is being used only on a limited number of isolates. None of the methods could show a clear association between genotype and phenotype, which may reflect the limitations of these methods or could be related to a relatively small sample size and inadequate study design.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 0.9 MB.

ACKNOWLEDGMENTS

We thank Karen Gesy (University of Saskatchewan), Taylor Gibson (University of Saskatchewan), and Sara Andrés-Lasheras (Lethbridge Research and Development Centre) for assistance with isolation and culturing, Champika Fernando (University of Saskatchewan) and Janet Hill for assistance with WGS, and Tony Ruzzini for support throughout this study. We also extend our gratitude to the veterinary practices that provided the samples (Feedlot Health Management Services, Veterinary Agri-Health Services, and Alberta Beef Health Solutions) and to the feedlot producers who allowed their cattle to be sampled.

This study was supported by the Saskatchewan Ministry of Agriculture (Agricultural Development Fund: 20160253) and the Canadian Cattlemen's Association (Beef Cattle Research Council: ANH.30.17).

REFERENCES

- Nicholas RAJ, Ayling RD. 2003. *Mycoplasma bovis*: disease, diagnosis, and control. *Res Vet Sci* 74:105–112. [https://doi.org/10.1016/S0034-5288\(02\)00155-8](https://doi.org/10.1016/S0034-5288(02)00155-8).
- Maunsell FP, Woolums AR, Francoz D, Rosenbusch RF, Step DL, Wilson DJ, Janzen ED. 2011. *Mycoplasma bovis* infections in cattle. *J Vet Intern Med* 25:772–783. <https://doi.org/10.1111/j.1939-1676.2011.0750.x>.
- Calcutt MJ, Lysnyansky I, Sachse K, Fox LK, Nicholas RAJ, Ayling RD. 2018. Gap analysis of *Mycoplasma bovis* disease, diagnosis and control: an aid to identify future development requirements. *Transbound Emerg Dis* 65:91–109. <https://doi.org/10.1111/tbed.12860>.
- Levison LJ, Miller-Cushon EK, Tucker AL, Bergeron R, Leslie KE, Barkema HW, DeVries TJ. 2016. Incidence rate of pathogen-specific clinical mastitis on conventional and organic Canadian dairy farms. *J Dairy Sci* 99:1341–1350. <https://doi.org/10.3168/jds.2015-9809>.
- Bauman CA, Barkema HW, Dubuc J, Keefe GP, Kelton DF. 2018. Canadian National Dairy Study: herd-level milk quality. *J Dairy Sci* 101:2679–2691. <https://doi.org/10.3168/jds.2017-13336>.
- Fox LK. 2012. *Mycoplasma* mastitis: causes, transmission, and control. *Vet Clin North Am Food Anim Pract* 28:225–237. <https://doi.org/10.1016/j.cvfa.2012.03.007>.
- Gagea MI, Bateman KG, Shanahan RA, van Dreumel T, McEwen BJ, Carman S, Archambault M, Caswell JL. 2006. Naturally occurring *Mycoplasma bovis*-associated pneumonia and polyarthritis in feedlot beef calves. *J Vet Diagn Invest* 18:29–40. <https://doi.org/10.1177/104063870601800105>.
- Martin SW, Bateman KG, Shewen PE, Rosendal S, Bohac JG, Thorburn M. 1990. A group level analysis of the associations between antibodies to seven putative pathogens and respiratory disease and weight gain in Ontario feedlot calves. *Can J Vet Res* 54:337–342.
- Maunsell FP, Donovan GA. 2009. *Mycoplasma bovis* infections in young calves. *Vet Clin North Am Food Anim Pract* 25:139–177. <https://doi.org/10.1016/j.cvfa.2008.10.011>.
- Grissett GP, White BJ, Larson RL. 2015. Structured literature review of responses of cattle to viral and bacterial pathogens causing bovine respiratory disease complex. *J Vet Intern Med* 29:770–780. <https://doi.org/10.1111/jvim.12597>.
- Radaelli E, Luini M, Loria GR, Nicholas RAJ, Scanziani E. 2008. Bacteriological, serological, pathological and immunohistochemical studies of *Mycoplasma bovis* respiratory infection in veal calves and adult cattle at slaughter. *Res Vet Sci* 85:282–290. <https://doi.org/10.1016/j.rvsc.2007.11.012>.
- Maunsell FP, Donovan GA, Risco C, Brown MB. 2009. Field evaluation of a *Mycoplasma bovis* bacterin in young dairy calves. *Vaccine* 27:2781–2788. <https://doi.org/10.1016/j.vaccine.2009.02.100>.
- Punyapornwithaya V, Fox LK, Hancock DD, Gay GM, Alldredge JR. 2010. Association between an outbreak strain causing *Mycoplasma bovis* mastitis and its asymptomatic carriage in the herd: a case study from Idaho, USA. *Prev Vet Med* 93:66–70. <https://doi.org/10.1016/j.prevetmed.2009.08.008>.
- Alabdullah HA. 2017. *Mycoplasma bovis* associated diseases: modeling stress and its effect on bovine immune function. PhD thesis. Washington State University, Pullman, WA.
- Caswell JL, Bateman KG, Cai HY, Castillo-Alcala F. 2010. *Mycoplasma bovis* in respiratory disease of feedlot cattle. *Vet Clin North Am Food Anim Pract* 26:365–379. <https://doi.org/10.1016/j.cvfa.2010.03.003>.
- Arcangioli M-A, Aslan H, Tardy F, Poumarat F, Grand DL. 2012. The use of pulsed-field gel electrophoresis to investigate the epidemiology of *Mycoplasma bovis* in French calf feedlots. *Vet J* 192:96–100. <https://doi.org/10.1016/j.tvjl.2011.05.004>.
- Biddle MK, Fox LK, Evans MA, Gay CC. 2005. Pulsed-field gel electrophoresis patterns of *Mycoplasma* isolates from various body sites in dairy cattle with *Mycoplasma* mastitis. *J Am Vet Med Assoc* 227:455–459. <https://doi.org/10.2460/javma.2005.227.455>.
- Spergser J, Macher K, Kargl M, Lysnyansky I, Rosengarten R. 2013. Emergence, re-emergence, spread and host species crossing of *Mycoplasma bovis* in the Austrian Alps caused by a single endemic strain. *Vet Microbiol* 164:299–306. <https://doi.org/10.1016/j.vetmic.2013.02.007>.
- Castillo-Alcala F, Bateman KG, Cai HY, Schott CR, Parker L, Clark ME, McRaid P, McDowall RM, Foster RA, Archambault M, Caswell JL. 2012. Prevalence and genotype of *Mycoplasma bovis* in beef cattle after arrival

- at a feedlot. *Am J Vet Res* 73:1932–1943. <https://doi.org/10.2460/ajvr.73.12.1932>.
20. Pinho L, Thompson G, Rosenbusch R, Carvalheira J. 2012. Genotyping of *Mycoplasma bovis* isolates using multiple-locus variable-number tandem-repeat analysis. *J Microbiol Methods* 88:377–385. <https://doi.org/10.1016/j.mimet.2012.01.003>.
 21. Sulyok KM, Kreizinger Z, Fekete L, Jánosi S, Schweitzer N, Turcsányi I, Makrai L, Erdélyi K, Gyuranecz M. 2014. Phylogeny of *Mycoplasma bovis* isolates from Hungary based on multi locus sequence typing and multiple-locus variable-number tandem repeat analysis. *BMC Vet Res* 10:108. <https://doi.org/10.1186/1746-6148-10-108>.
 22. Becker CAM, Thibault FM, Arcangioli M-A, Tardy F. 2015. Loss of diversity within *Mycoplasma bovis* isolates collected in France from bovines with respiratory diseases over the last 35 years. *Infect Genet Evol* 33:118–126. <https://doi.org/10.1016/j.meegid.2015.04.019>.
 23. Jolley KA, Bray JE, Maiden MCJ. 2018. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res* 3:124. <https://doi.org/10.12688/wellcomeopenres.14826.1>.
 24. Lysnyansky I, Freed M, Rosales RS, Mikula I, Khateb N, Gerchman I, van Straten M, Levisohn S. 2016. An overview of *Mycoplasma bovis* mastitis in Israel (2004–2014). *Vet J* 207:180–183. <https://doi.org/10.1016/j.tvjl.2015.10.057>.
 25. Register KB, Lysnyansky I, Jelinski MD, Boatwright WD, Waldner M, Bayles DO, Pilo P, Alt DP. 2020. Comparison of two multilocus sequence typing schemes for *Mycoplasma bovis* and revision of the PubMLST reference method. *J Clin Microbiol* 58:e00283-20. <https://doi.org/10.1128/JCM.00283-20>.
 26. Rosales RS, Churchward CP, Schnee C, Sachse K, Lysnyansky I, Catania S, lob L, Ayling RD, Nicholas RAJ. 2015. Global multilocus sequence typing analysis of *Mycoplasma bovis* isolates reveals two main population clusters. *J Clin Microbiol* 53:789–794. <https://doi.org/10.1128/JCM.01910-14>.
 27. Ghanem M, Wang L, Zhang Y, Edwards S, Lu A, Ley D, El-Gazzar M. 2017. Core genome multilocus sequence typing: a standardized approach for molecular typing of *Mycoplasma gallisepticum*. *J Clin Microbiol* 56:e01145-17. <https://doi.org/10.1128/JCM.01145-17>.
 28. Haapala V, Pohjanvirta T, Vähänikkilä N, Halkilahti J, Simonen H, Pelkonen S, Soveri T, Simojoki H, Autio T. 2018. Semen as a source of *Mycoplasma bovis* mastitis in dairy herds. *Vet Microbiol* 216:60–66. <https://doi.org/10.1016/j.vetmic.2018.02.005>.
 29. Timonen AAE, Autio T, Pohjanvirta T, Häkkinen L, Katholm J, Petersen A, Mötus K, Kalmus P. 2020. Dynamics of the within-herd prevalence of *Mycoplasma bovis* intramammary infection in endemically infected dairy herds. *Vet Microbiol* 242:108608. <https://doi.org/10.1016/j.vetmic.2020.108608>.
 30. Tardy F, Aspan A, Autio T, Ridley A, Tricot A, Colin A, Pohjanvirta T, Smid B, Harders F, Lindegaard M, Tølbøll Lauritsen K, Lyhs U, Wisselink H, Lenz Strube M. 2020. *Mycoplasma bovis* in Nordic European countries: emergence and dominance of a new clone. *Pathogens* 9:875. <https://doi.org/10.3390/pathogens9110875>.
 31. Yair Y, Borovok I, Mikula I, Falk R, Fox LK, Gophna U, Lysnyansky I. 2020. Genomics-based epidemiology of bovine *Mycoplasma bovis* strains in Israel. *BMC Genomics* 21:70. <https://doi.org/10.1186/s12864-020-6460-0>.
 32. Kumar R, Register K, Christopher-Hennings J, Moroni P, Gioia G, Garcia-Fernandez N, Nelson J, Jelinski MD, Lysnyansky I, Bayles D, Alt D, Scaria J. 2020. Population genomic analysis of *Mycoplasma bovis* elucidates geographical variations and genes associated with host-types. *Microorganisms* 8:1561. <https://doi.org/10.3390/microorganisms8101561>.
 33. Jelinski M, Kinneer A, Gesy K, Andrés-Lasheras S, Zaheer R, Weese S, McAllister TA. 2020. Antimicrobial sensitivity testing of *Mycoplasma bovis* isolates derived from western Canadian feedlot cattle. *Microorganisms* 8:124. <https://doi.org/10.3390/microorganisms8010124>.
 34. Hayflick L. 1965. Tissue cultures and mycoplasmas. *Tex Repl Biol Med* 23:285–303.
 35. Register KB, Boatwright WD, Gesy KM, Thacker TC, Jelinski MD. 2018. Mistaken identity of an open reading frame proposed for PCR-based identification of *Mycoplasma bovis* and the effect of polymorphisms and insertions on assay performance. *J Vet Diagn Invest* 30:637–641. <https://doi.org/10.1177/1040638718764799>.
 36. Miles K, McAuliffe L, Ayling RD, Nicholas RAJ. 2004. Rapid detection of *Mycoplasma dispar* and *M. bovirhinis* using allele specific polymerase chain reaction protocols. *FEMS Microbiol Lett* 241:103–107. <https://doi.org/10.1016/j.femsle.2004.10.010>.
 37. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
 38. Jünemann S, Sedlazeck FJ, Prior K, Albersmeier A, John U, Kalinowski J, Mellmann A, Goesmann A, Von Haeseler A, Stoye J, Harmsen D. 2013. Updating benchtop sequencing performance comparison. *Nat Biotechnol* 31:294–296. <https://doi.org/10.1038/nbt.2522>.
 39. Souvorov A, Agarwala R, Lipman DJ. 2018. SKESA: strategic k-mer extension for scrupulous assemblies. *Genome Biol* 19:153. <https://doi.org/10.1186/s13059-018-1540-z>.
 40. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
 41. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prijbelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
 42. Gardner SN, Slezak T, Hall BG. 2015. kSNP3: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics* 31:2877–2878. <https://doi.org/10.1093/bioinformatics/btv271>.
 43. Hunter PR, Gaston MA. 1988. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J Clin Microbiol* 26:2465–2466. <https://doi.org/10.1128/JCM.26.11.2465-2466.1988>.
 44. Letunic I, Bork P. 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23:127–128. <https://doi.org/10.1093/bioinformatics/btl529>.
 45. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol Biol Evol* 35:1547–1549. <https://doi.org/10.1093/molbev/msy096>.
 46. Zhou Z, Alikhan NF, Sergeant MJ, Luhmann N, Vaz C, Francisco AP, Carriço JA, Achtman M. 2018. GrapeTree: visualization of core genomic relationships among 100,000 bacterial pathogens. *Genome Res* 28:1395–1404. <https://doi.org/10.1101/gr.232397.117>.
 47. Parker AM, Shukla A, House JK, Hazelton MS, Bosward KL, Kokotovic B, Sheehy PA. 2016. Genetic characterization of Australian *Mycoplasma bovis* isolates through whole genome sequencing analysis. *Vet Microbiol* 196:118–125. <https://doi.org/10.1016/j.vetmic.2016.10.010>.
 48. Bokma J, Vereecke N, De Bleecker K, Callens J, Ribbens S, Nauwynck H, Haesebrouck F, Theuns S, Boyen F, Pardon B. 2020. Phylogenomic analysis of *Mycoplasma bovis* from Belgian veal, dairy and beef herds. *Vet Res* 51:121. <https://doi.org/10.1186/s13567-020-00848-z>.
 49. Vereecke N, Bokma J, Haesebrouck F, Nauwynck H, Boyen F, Pardon B, Theuns S. 2020. High quality genome assemblies of *Mycoplasma bovis* using a taxon-specific Bonito basecaller for MinION and Flongle long-read nanopore sequencing. *BMC Bioinformatics* 21:517. <https://doi.org/10.1186/s12859-020-03856-0>.