

RESEARCH ARTICLE

Open Access

Limitations of variable number of tandem repeat typing identified through whole genome sequencing of *Mycobacterium avium* subsp. *paratuberculosis* on a national and herd level

Christina Ahlstrom¹, Herman W Barkema¹, Karen Stevenson², Ruth N Zadoks^{2,3}, Roman Biek³, Rowland Kao³, Hannah Trewby³, Deb Hauptstein⁴, David F Kelton⁵, Gilles Fecteau⁶, Olivia Labrecque⁷, Greg P Keefe⁸, Shawn L B McKenna⁸ and Jeroen De Buck^{1*}

Abstract

Background: *Mycobacterium avium* subsp. *paratuberculosis* (MAP), the causative bacterium of Johne's disease in dairy cattle, is widespread in the Canadian dairy industry and has significant economic and animal welfare implications. An understanding of the population dynamics of MAP can be used to identify introduction events, improve control efforts and target transmission pathways, although this requires an adequate understanding of MAP diversity and distribution between herds and across the country. Whole genome sequencing (WGS) offers a detailed assessment of the SNP-level diversity and genetic relationship of isolates, whereas several molecular typing techniques used to investigate the molecular epidemiology of MAP, such as variable number of tandem repeat (VNTR) typing, target relatively unstable repetitive elements in the genome that may be too unpredictable to draw accurate conclusions. The objective of this study was to evaluate the diversity of bovine MAP isolates in Canadian dairy herds using WGS and then determine if VNTR typing can distinguish truly related and unrelated isolates.

Results: Phylogenetic analysis based on 3,039 SNPs identified through WGS of 124 MAP isolates identified eight genetically distinct subtypes in dairy herds from seven Canadian provinces, with the dominant type including over 80% of MAP isolates. VNTR typing of 527 MAP isolates identified 12 types, including "bison type" isolates, from seven different herds. At a national level, MAP isolates differed from each other by 1–2 to 239–240 SNPs, regardless of whether they belonged to the same or different VNTR types. A herd-level analysis of MAP isolates demonstrated that VNTR typing may both over-estimate and under-estimate the relatedness of MAP isolates found within a single herd.

Conclusions: The presence of multiple MAP subtypes in Canada suggests multiple introductions into the country including what has now become one dominant type, an important finding for Johne's disease control. VNTR typing often failed to identify closely and distantly related isolates, limiting the applicability of using this typing scheme to study the molecular epidemiology of MAP at a national and herd-level.

Keywords: *Mycobacterium avium* subspecies *paratuberculosis*, Variable number tandem repeat typing, Whole genome sequencing

* Correspondence: jdebuck@ucalgary.ca

¹University of Calgary, Calgary, Alberta, Canada

Full list of author information is available at the end of the article

Background

Mycobacterium avium subsp. *paratuberculosis* (MAP) is the causative bacterium of Johne's disease in ruminants and has significant impacts on the health of cattle and the economics of dairy production systems in particular [1]. This organism is primarily transmitted via the fecal-oral route and is generally introduced into herds through the purchase of one or more infected animals. Despite control efforts in the Canadian dairy industry, up to 76% of herds are infected with MAP [2]. Additionally, a potential association with Crohn's disease in humans is increasing the pressure to control MAP on farms and in the environment [3,4].

Limited tools exist to assess within-species diversity of pathogens with low genetic variation [5], such as MAP. Despite this challenge, a number of molecular typing methods have been developed that target repeat regions in the genome and have been moderately successful in observing apparent diversity [6,7]. Variable number tandem repeat (VNTR) and mycobacterial interspersed repetitive unit (MIRU) typing are based on repetitive element differences and have been used in molecular epidemiological studies in a number of organisms and settings [8-12]. Most recently, Oakey et al. [13] used VNTR typing to investigate MAP incursions in a low-prevalence region of Australia and concluded that two separate VNTR types were circulating in this region. This genetic information was used to evaluate epidemiological links and inform policy decisions and control strategies.

To date, whole genome sequencing (WGS) of MAP has predominantly focused on *de novo* assembly of a limited number of relevant genomes [14-16]. While this has provided immense resources to understand the genomic structure and evolution of major strain types (types I, II, and III), WGS of many epidemiologically linked isolates can determine the true degree of diversity and quantify relatedness. Single nucleotide polymorphisms (SNPs) identified through WGS are evolutionarily stable and can be reliably used to identify true evolutionary relationships [17]. In a MAP-endemic environment, this level of detail is invaluable in understanding the molecular epidemiology and transmission dynamics.

Many reports have described the diversity of MAP isolates using VNTR typing with or without additional molecular targets [12,18-20]. The validity of using VNTR typing to help clarify transmission and the true diversity of MAP, however, has yet to be determined. Whilst the instability of repetitive elements contributes to their discriminatory ability, there is also the risk of convergent evolution, which may lead to incurred epidemiological inferences from VNTR data [21]. Therefore, the aim of this study was to assess the diversity of MAP isolates (type II) in Canadian dairy herds using WGS to assess

the ability of VNTR typing to identify unrelated and related isolates at the national and herd level.

Methods

MAP isolates and DNA preparation

A collection of Canadian MAP isolates was derived from previously obtained samples through provincial Johne's disease control programs and research projects. Positive culture broth from environmental manure samples was obtained from Johne's disease control programs in Alberta and Saskatchewan [2], the Atlantic region (Prince Edward Island and New Brunswick) [22] and Québec (Table 1), while individual cow fecal samples were obtained from British Columbia and Québec. Six environmental manure samples were collected from participating Alberta and all Saskatchewan dairy herds. Additionally, individual cow fecal samples from cows with high milk ELISA titers in Ontario were obtained and cultured as described previously [23]. Samples from Québec were cultured using the BACTEC MGIT 960 ParaTB culture system (Becton, Dickinson and company, Franklin Lakes, NJ, USA), whereas the remaining samples were cultured using the TREK ESP Culture System reagents (TREK Diagnostics, Cleveland, OH, USA). Approximately 10 µl of culture broth was plated onto Middlebrook 7H11 agar supplemented with 2 mg/L mycobactin J to isolate individual MAP colonies. Plates were incubated at 37°C for 4–6 weeks. A single colony was subsequently sub-streaked onto 7H11 agar and incubated for 4–8 weeks prior to DNA extraction. DNA was extracted using a modified protocol of the Qiagen DNeasy blood and tissue kit (Qiagen, Mississauga, ON, Canada), as described previously [24].

VNTR typing

VNTR typing targeting eight loci was performed on 527 DNA samples extracted from MAP isolates using a previously established protocol [7] and VNTR types (INRA

Table 1 Province of origin, number of herds, VNTR types and number of *Mycobacterium avium* subsp. *paratuberculosis* isolates analyzed

Province/Region	# Herds	# VNTR typed	VNTR types (INMV #)	# Whole genome sequenced
Alberta	98	300	1, 13, 17, 2, 3, 6, 68, 77, NEW2	58
Atlantic regions	10	62	17, 2, 77	11
British Columbia	8	14	2, 3	5
Ontario	17	52	17, 2, 3, 6	16
Quebec	24	45	1, 2, 3, 68	14
Saskatchewan	25	54	1, 2, 3, 6, 7, 117, NEW	20
Total	182	527		124

Nouzilly MIRU-VNTR [INMV] profile number) were assigned according to the MAC-INMV database [25].

Whole genome sequencing

MAP isolates were selected for WGS to 1) represent the approximate proportion of isolates per VNTR type in Canada (while minimizing duplicate types from the same herd), 2) include all Canadian provinces currently represented in the MAP collection, and 3) represent three different scenarios of VNTR type diversity at the herd-level in Alberta. The three Alberta herd-level isolate sample sets included a herd with (A) four isolates of the same VNTR type, (B) three isolates with three different VNTR types, and (C) four isolates with two different VNTR types. As these isolates were cultured from environmental manure samples, true herd-level diversity was not assessed.

MAP DNA was prepared for sequencing using the NexteraXT sample preparation kit (Illumina, San Diego, CA, USA). Samples were multiplexed to achieve paired end reads with an average coverage of 50X and sequenced using V2 (250 bp reads) or V3 (300 bp reads) chemistry using the Illumina MiSeq sequencing platform (Illumina, San Diego, CA, USA).

Data analysis

Raw reads were trimmed using ConDeTri [26] and mapped to a revised version of the K10 reference genome (NCBI Sequence Read Archive study SRR060191) [27] using BWA [28]. Variant sites were identified using SAMtools [29] and were subsequently filtered based on depth of coverage (at least 2 high quality SNPs on both the forward and reverse strand), mapping quality (>40), and heterozygosity (<5%). SAM files were visualized in Tablet [30] to investigate specific polymorphisms of interest, such as the 2 base-pair deletion in IS1311 “bison type” isolates [31]. An alignment of concatenated SNPs was used to create a maximum likelihood phylogenetic tree using PhyML [32], using the nucleotide substitution model, as determined by jModelTest [33]. Clade support was evaluated based on the analysis of 100 bootstrap pseudo-replicates. To determine how specific VNTR loci contribute to the phylogenetic congruence, a circularized phylogenetic tree was annotated according to the six loci that were polymorphic in the WGS dataset (292, X3, 25, 47, 7, and 10) [34].

Additionally, the absolute minimum, maximum, and average pairwise genetic distance of each isolate to every other isolate within the same VNTR type and between VNTR types was calculated for the three most common types by averaging across the PhyML generated distance matrix. To identify how frequently closely related isolates differ in VNTR type, the number of pairwise distances smaller than a range of cut-off values (<5, <10,

<15, <20, <25, <30, <35, <40, <45, <50, <75, <100, <200) was calculated within and between types.

Results

VNTR typing

MAP isolates evaluated in this study came from a total of 201 Canadian dairy herds from British Columbia, Alberta, Saskatchewan, Ontario, Québec, and the Atlantic region. VNTR typing was performed on 527 MAP isolates and resulted in 12 different types (Table 2). A total of 76, 8, and 7% of MAP isolates belonged to INMV type 2, 3, and 1 respectively, with 9 types representing the remaining 9% of isolates. Two loci (3 and 32) were not found to be polymorphic, while loci 292, X3, 25, 47, 7, and 10 contained 3, 3, 3, 2, 2, and 2 alleles, respectively. MAP isolates belonging to INMV68, which was previously identified as “bison type” isolates in Québec [18], were identified in seven herds (two from Québec and five from Alberta) (Table 1).

Whole genome sequencing

124 isolates were successfully sequenced, representing 11 different VNTR types, with an average coverage of 33X. 3,039 high quality variant sites were identified and used to create a maximum likelihood phylogenetic tree (Figure 1). Whole genome sequencing revealed more than 650 SNPs unique to the “bison type” isolate, and visualization of the reads in Tablet confirmed a ‘TG’ deletion unique to “bison type” strains in the IS1311 sequences. Eight divergent subtypes were identified that contained at least 45 unique SNPs relative to the most recent common ancestor of all isolates, excluding the “bison type” outgroup. The dominant subtype (subtype VIII) included 86% of isolates analyzed.

Table 2 The number of MAP isolates, herds, and whole genomes sequenced (WGS) within each VNTR type

VNTR type	Locus								# isolates	# herds	# WGS
	292	X3	25	47	3	7	10	32			
INMV 2	3	2	3	3	2	2	2	8	398	139	83
INMV 3	3	2	3	3	2	2	1	8	41	25	16
INMV 1	4	2	3	3	2	2	2	8	36	16	13
INMV 17	3	1	3	3	2	2	2	8	18	3	3
INMV 68	2	2	5	3	2	2	2	8	17	6	1
INMV 13	2	2	3	3	2	2	2	8	6	2	1
INMV 6	3	2	3	3	2	1	2	8	4	4	2
INMV 77	3	2	2	3	2	2	2	8	3	2	2
INMV 7	3	2	3	3	2	1	1	8	1	1	1
INMV117	3	2	3	2	2	2	2	8	1	1	1
NEW1	3	3	3	3	2	2	1	8	1	1	1
NEW2	4	2	3	2	2	2	2	8	1	1	0

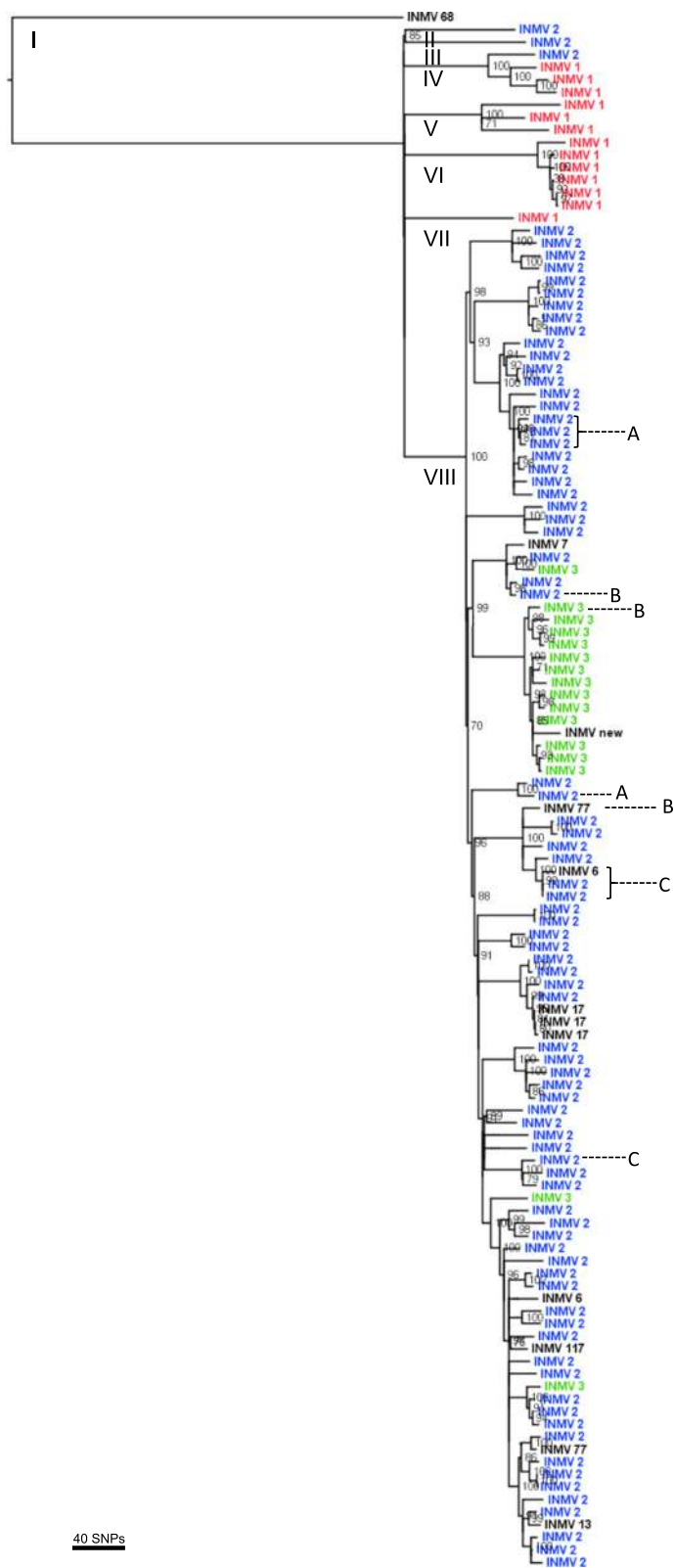


Figure 1 (See legend on next page.)

(See figure on previous page.)

Figure 1 Maximum likelihood phylogenetic tree created from 3,039 concatenated variant sites using PhyML and the TPM1uf nucleotide substitution model, rooted to the “bison type” isolate (INMV 68). The tips are labeled as the VNTR type and the three most prevalent types are color-coded (INMV 2 = blue, INMV 3 = green, INMV 1 = red). Dotted lines indicate samples belonging to Herds A, B, and C, the eight divergent subtypes are labeled I-VIII, and bootstrap values of node support ≥ 70 are displayed.

Analysis of VNTR typing and WGS

In Figure 1, tips of the phylogenetic tree were labeled according to VNTR type and the three most common types were color-coded. INMV 2 isolates in this dataset originated from all six provinces, whereas INMV 3 and INMV 1 isolates were identified from five and three provinces, respectively. Of the 11 VNTR types identified, nine were located within the dominant subtype (VIII).

The frequency with which isolates differing by a range of SNPs belong to different VNTR types is depicted in Figure 2. In this dataset, MAP isolates differing by <10 SNPs belonged to different VNTR types 10% of the time. This value increased as the number of SNP differences increased.

The average number of pairwise SNP differences between isolates within the three most common VNTR types and outside of each type is presented in Figure 3. The majority of isolates within a type had lower average SNP differences when compared to the differences between VNTR types (Figure 3). However, in each of the three major VNTR types, several isolates with within-type differences clustered with the between-type differences, with average SNP differences that were considerably higher than would be expected for isolates of the same VNTR type. INMV 3 isolates were on average more closely related, with less than 50 SNP differences in the largest cluster, compared to INMV 2 isolates, which were approximately 100 SNPs different. INMV 1, on the other hand, had on average over 150 SNP differences between isolates within that type, with one isolate exceeding a difference of 200 SNPs. The absolute maximum number

of SNP differences between any two isolates within each type was 240, 215, and 116 for INMV 1, 2, and 3 respectively, while the minimum number of SNP differences was one for INMV 1 and INMV 2 and four for INMV 3 (Table 3). The absolute maximum and minimum number of SNP differences between two different VNTR types for INMV 1, 2, and 3 was also calculated. In the case of INMV 1, the maximum SNP difference within the type was larger than between types (Table 3). The dominant subtype included five different VNTR types that differed from INMV 2 by 15 SNPs or less, including INMV 17 that differed from an INMV 2 isolate by only two SNPs (Figure 1).

Three herds with multiple MAP isolates per herd were evaluated using VNTR typing and WGS. Four isolates of the same VNTR type were analyzed in Herd A, three isolates of different VNTR types were analyzed in Herd B, and four isolates representing two VNTR types were analyzed in Herd C. Herd B VNTR typing results correctly differentiated distantly related isolates; however, isolates from Herd A and Herd C had inconsistent typing results (Figure 1) with either unrelated isolates belonging to the same VNTR type (Herds A and C) or different VNTR types being closely related (Herd C).

Additionally, there was evidence of convergent evolution in four of six VNTR types in which more than one isolate per type was sequenced (INMV 2, 3, 6, and 77) (Figure 1). Locus-specific changes in repeat length are presented in Figure 4. Locus 10 had the most instances of convergent evolution (five times), whereas loci 47 and X3 appeared to be relatively stable.

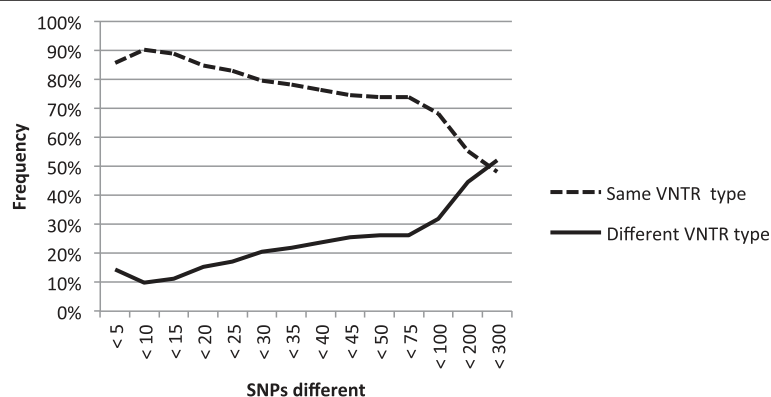
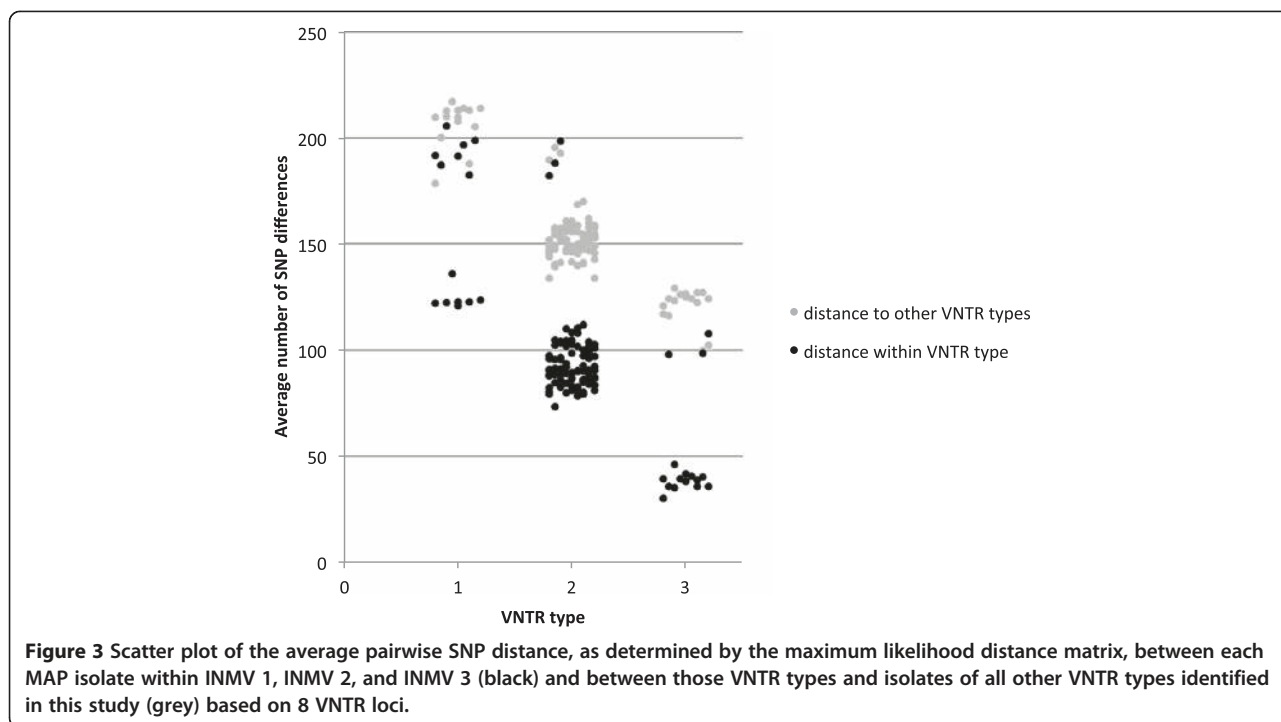


Figure 2 The frequency in which any two MAP isolates share the same VNTR type (solid line) or different VNTR types (dotted line) at a range of pairwise SNP differences (<5 to <300 SNPs) (i.e. two isolates that have fewer than 300 pairwise SNP differences belong to different VNTR types 52% of the time).



Discussion

In this study, the genetic relationship of 124 Canadian bovine MAP isolates was assessed using WGS and VNTR typing. Main findings were that 1) at least eight genetically distinct MAP subtypes exist in Canada, with over 80% of isolates belonging to a single dominant subtype, and 2) VNTR typing may both overestimate and underestimate relatedness. Additionally, evidence of convergent evolution was observed multiple times and individual VNTR loci contributed differently to the levels of VNTR-level homoplasy observed in the SNP-based phylogeny.

The phylogenetic analysis of MAP isolates based on over 3,000 SNPs identified through WGS revealed eight divergent subtypes, one of which contains over 80% of isolates. The presence of a dominant subtype in Canada could be attributed to a variety of factors, such as increased virulence, increased culturability, or a founder effect (e.g. the first MAP introduction into Canada). The branch lengths separating the major subtypes exceed 100 SNPs, suggesting there were likely at least eight separate introduction events into Canada, as the accumulation of this number of SNPs

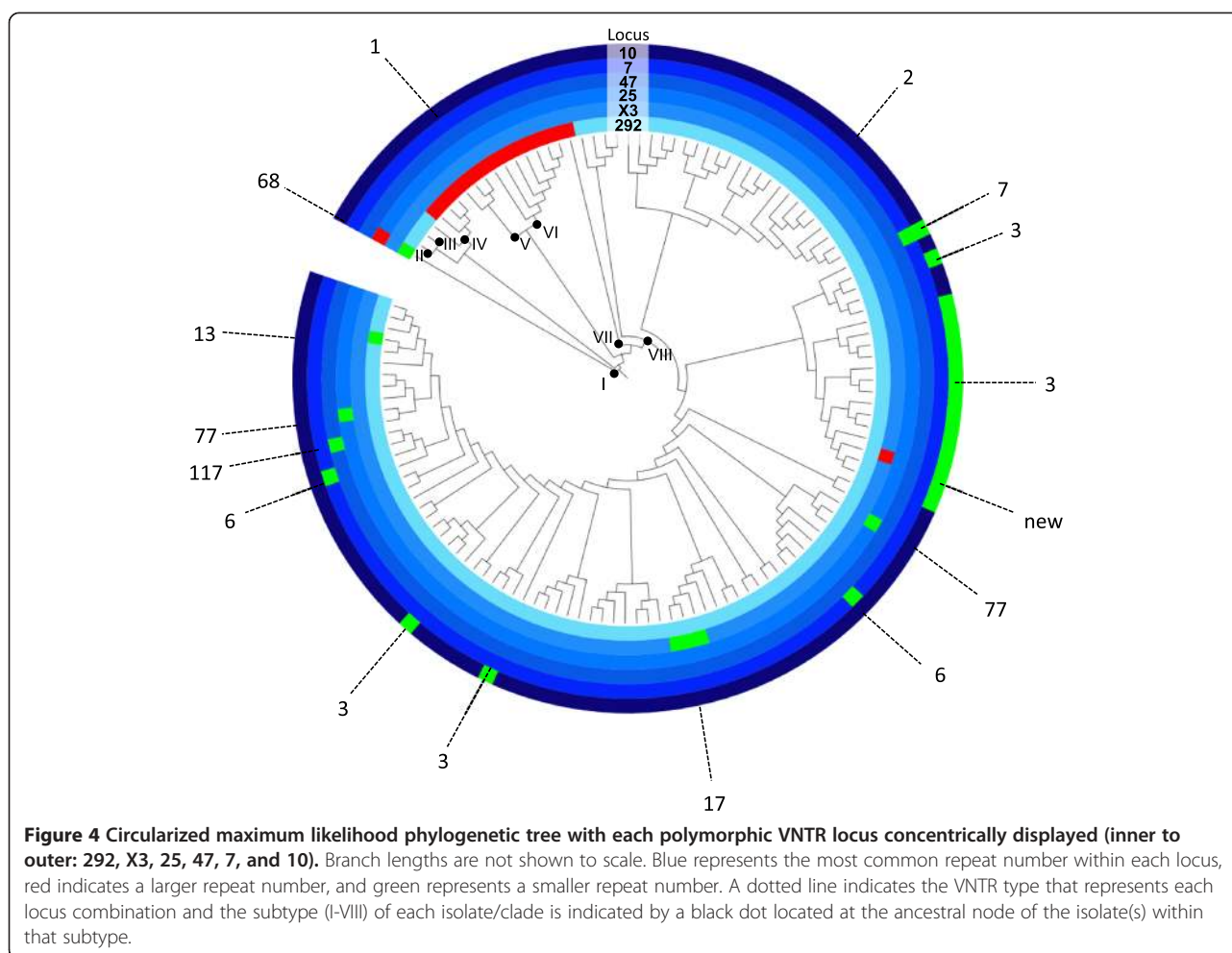
is unlikely within the timeframe of Holstein-Friesian dairy farming in Canada (1881 to the present) [35].

In the absence of observational data, a suboptimal typing technique will falsely classify related isolates as unrelated, masking the true relationship, or overestimate relatedness, incorrectly linking unrelated isolates. Within the three most common VNTR types, some isolates were highly unrelated based on the number of SNP differences, with as many as 240 differences within a type (INMV 1). Given the slow mutation rate of MAP, estimated to be slower than the 0.3 SNPs/genome/year in *Mycobacterium tuberculosis* ([36]; Bryant et al., unpublished), these isolates are highly unrelated. Falsely classifying isolates as related is characteristic of a typing technique with a low discriminatory ability, which can be resolved by including additional discriminatory genetic markers.

In contrast, this dataset provides examples where evidence of transmission based on VNTR typing would be lost due to a change in repeat number at a single locus (i.e. INMV 2 and 17; INMV 2 and INMV 6). This has similarly been demonstrated in the evaluation of VNTR typing of *M. bovis*, where single locus variants were present in the

Table 3 Absolute minimum and maximum pairwise SNP difference within and between the three major VNTR types

	INMV 1		INMV 2		INMV 3	
	Within	Between	Within	Between	Within	Between
Absolute minimum	1	70	1	2	4	15
Absolute maximum	240	239	215	236	116	230



same herd, suggesting the presence of clonal variants [37]. While it has been demonstrated that genetic data based on repetitive elements are not appropriate for deep phylogenetic inference [5], the observation that a VNTR type can be identical by convergent evolution indicates that classifying isolates based on VNTR type may lead to erroneous epidemiological conclusions regarding transmission events occurring within a timeframe of a few years. Given that the most severe cases of convergent evolution occur in the dominant subtype (VIII), the use of a typing scheme that first defines major lineages based on SNPs followed by VNTR and/or additional typing methods, as has been proposed for other organisms [5,9], will not substantially clarify transmission dynamics at this limited spatial scale.

The analysis of three herd-level datasets illustrates the value of using WGS data for the study of epidemiologically linked isolates and that VNTR typing data may lead to an incorrect assessment of diversity and relatedness of strains. Herd A harbored at least two different strains that appear to be identical based on VNTR typing. Herd C also had at least two different strains circulating; however,

despite the presence of two different VNTR types, presence of INMV 6 is likely due to within-herd evolution of the closely related INMV 2 isolates, whereas the more divergent INMV 2 isolate was probably the result of a separate introduction event.

Locus-specific differences in repeat length were also evaluated. Convergence was observed in four of six loci and overall a higher frequency of loss, rather than gain, of repeat numbers relative to the most common allele at each locus was found, which has been observed previously in *M. tuberculosis* [38,39]. Prediction models for VNTR locus evolution in other organisms have been developed [21,40,41], but they have not been based on comparison with WGS data. It has been reported previously [9] that different genetic markers should be selected for different spatial and temporal scales, so that the molecular clock speed of the marker is matched to the scale of the investigation. However, results from the current study suggest that the current VNTR typing scheme includes loci that are too unstable, and therefore unreliable, to be used for molecular epidemiological analysis of MAP at both broad and limited spatial scales.

Conclusions

The usefulness of molecular typing strongly depends on its ability to differentiate epidemiologically meaningful differences between isolates. Molecular epidemiological studies of MAP have not yet provided sufficient results to truly impact our understanding of the transmission dynamics and genotype-specific virulence characteristics. Correctly classifying strains as “same” or “different” is important for transmission studies and in identifying phenotypic associations. Based on results from this study, caution should be used when using VNTR typing as a tool to assess the diversity and relatedness of MAP isolates at both a national and herd-level. Whole genome sequencing, on the other hand, provides unparalleled detail regarding the genetic profiles of MAP. The dominance of a single clade of MAP in Canada shows how important it is to have such detailed information when attempting to use molecular data to study MAP transmission dynamics.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

HWB, DH, DFK, GF, OL, GPK, SLBM, and JDB designed sample collection. CAA performed labwork. CAA, HWB, and JDB designed the study, performed data analysis, and drafted the manuscript. HT conceived the bioinformatic pipeline. KS, RNZ, RB, and RK provided expertise in WGS and VNTR analyses. All authors read and approved the final manuscript.

Acknowledgements

This research was supported by the Alberta Livestock and Meat Agency and Dairy Farmers of Canada. The authors would like to thank Brian Radke of the British Columbia Ministry of Agriculture for providing MAP isolates from British Columbia and Uliana Kanevets, Clara Lee, and Marguerite Speelman for their help with MAP culture and VNTR typing.

Author details

¹University of Calgary, Calgary, Alberta, Canada. ²Moredun Research Institute, Penicuik, Scotland. ³University of Glasgow, Glasgow, Scotland. ⁴SaskMilk, Regina, Saskatchewan, Canada. ⁵University of Guelph, Guelph, Ontario, Canada. ⁶Université de Montréal, Montréal, Québec, Canada. ⁷Laboratoire d'épidémiologie animale du Québec, Saint-Hyacinthe, Québec, Canada. ⁸University of Prince Edward Island, Charlottetown, Prince Edward Island, Canada.

Received: 29 October 2014 Accepted: 24 February 2015

Published online: 08 March 2015

References

- McKenna SLB, Keefe GP, Tiwari A, VanLeeuwen J, Barkema HW. Johne's disease in Canada part II: disease impacts, risk factors, and control programs for dairy producers. *Can Vet J.* 2006;47:1089–99.
- Wolf R, Barkema HW, De Buck J, Slomp M, Flaig J, Hauptstein D, et al. High herd-level prevalence of *Mycobacterium avium* subspecies *paratuberculosis* in Western Canadian dairy farms, based on environmental sampling. *J Dairy Sci.* 2014;97:6250–9.
- Barkema HW, Hendrick S, De Buck JM, Ghosh S, Kaplan GG, Rioux KP. Crohn's Disease in Humans and Johne's Disease in Cattle - Linked Diseases? In: Krause D, Hendrick S, editors. *Zoonotic Pathogens in the Food Chain*. Wallingford: CABI; 2010. p. 197–213.
- Behr MA. The path to Crohn's disease: is mucosal pathology a secondary event? *Inflamm Bowel Dis.* 2010;16:896–902.
- Comas I, Homolka S, Niemann S, Gagneux S. Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies. *PLoS One.* 2009;4:e7815.
- Amonsin A, Li LL, Zhang Q, Bannantine JP, Motiwala AS, Sreevatsan S, et al. Multilocus short sequence repeat sequencing approach for differentiating among *Mycobacterium avium* subsp. *paratuberculosis* strains. *J Clin Microbiol.* 2004;42:1694–702.
- Thibault VC, Grayon M, Boschirolu ML, Hubbans C, Overduin P, Stevenson K, et al. New variable-number tandem-repeat markers for typing *Mycobacterium avium* subsp. *paratuberculosis* and *M. avium* strains: comparison with IS900 and IS1245 restriction fragment length polymorphism typing. *J Clin Microbiol.* 2007;45:2404–10.
- Sola C, Filliol I, Legrand E, Lesjean S, Locht C, Supply P, et al. Genotyping of the *Mycobacterium tuberculosis* complex using MIRUs: association with VNTR and spoligotyping for molecular epidemiology and evolutionary genetics. *Infect Genet Evol.* 2003;3:125–33.
- Keim P, Van Ert MN, Pearson T, Vogler AJ, Huynh LY, Wagner DM. Anthrax molecular epidemiology and forensics: using the appropriate marker for different evolutionary scales. *Infect Genet Evol.* 2004;4:205–13.
- Srisungnam S, Rudeeaneksin J, Lukebua A, Wattanapokayakit S, Pasadorn S, Mahotam K, et al. Molecular epidemiology of leprosy based on VNTR typing in Thailand. *Lepr Rev.* 2009;80:280–9.
- Skuce RA, Mallon TR, McCormick CM, McBride SH, Clarke G, Thompson A, et al. *Mycobacterium bovis* genotypes in Northern Ireland: herd-level surveillance (2003 to 2008). *Vet Rec.* 2010;167:684–9.
- Stevenson K, Alvarez J, Bakker D, Biet F, De Juan L, Denham S, et al. Occurrence of *Mycobacterium avium* subspecies *paratuberculosis* across host species and European countries with evidence for transmission between wildlife and domestic ruminants. *BMC Microbiol.* 2009;9:212.
- Oakey J, Gavey L, Singh SV, Platell J, Waltisbuhl D. Variable-number tandem repeats genotyping used to aid and inform management strategies for a bovine Johne's disease incursion in tropical and subtropical Australia. *J Vet Diagn Invest.* 2014;26:651–7.
- Bannantine JP, Wu C, Hsu C, Zhou S, Schwartz DC, Bayles DO, et al. Genome sequencing of ovine isolates of *Mycobacterium avium* subspecies *paratuberculosis* offers insights into host association. *BMC Genomics.* 2012;13:89.
- Ghosh P, Hsu C, Alyamani EJ, Shehata MM, Al-Dubaib MA, Al-Naeem A, et al. Genome-wide analysis of the emerging infection with *Mycobacterium avium* subspecies *paratuberculosis* in the Arabian camels (*Camelus dromedarius*). *PLoS One.* 2012;7:e31947.
- Wynne JW, Bull TJ, Seemann T, Bulach DM, Wagner J, Kirkwood CD, et al. Exploring the zoonotic potential of *Mycobacterium avium* subspecies *paratuberculosis* through comparative genomics. *PLoS One.* 2011;6:e22171.
- Pearson T, Okinaka RT, Foster JT, Keim P. Phylogenetic understanding of clonal populations in an era of whole genome sequencing. *Infect Genet Evol.* 2009;9:1010–9.
- Sohal JS, Arsenault J, Labrecque O, Fairbrother J-H, Roy J-P, Fecteau G, et al. Genetic structure of *Mycobacterium avium* subsp. *paratuberculosis* population in cattle herds in Quebec as revealed by using a combination of multilocus genomic analyses. *J Clin Microbiol.* 2014;52:2764–75.
- Thibault VC, Grayon M, Boschirolu ML, Willery E, Allix-Béguec C, Stevenson K, et al. Combined multilocus short-sequence-repeat and mycobacterial interspersed repetitive unit-variable-number tandem-repeat typing of *Mycobacterium avium* subsp. *paratuberculosis* isolates. *J Clin Microbiol.* 2008;46:4091–4.
- Fritsch I, Luyven G, Köhler H, Lutz W, Möbius P. Suspicion of *Mycobacterium avium* subsp. *paratuberculosis* transmission between cattle and wild-living red deer (*Cervus elaphus*) by multitarget genotyping. *Appl Environ Microbiol.* 2012;78:1132–9.
- Reyes JF, Chan CHS, Tanaka MM. Impact of homoplasy on variable numbers of tandem repeats and spoligotypes in *Mycobacterium tuberculosis*. *Infect Genet Evol.* 2012;12:811–8.
- Lavers CJ, McKenna SLB, Dohoo IR, Barkema HW, Keefe GP. Evaluation of environmental fecal culture for *Mycobacterium avium* subspecies *paratuberculosis* detection in dairy herds and association with apparent within-herd prevalence. *Can Vet J.* 2013;54:1053–60.
- Mortier RA, Barkema HW, Orsel K, Wolf R, De Buck J. Shedding patterns of dairy calves experimentally infected with *Mycobacterium avium* subspecies *paratuberculosis*. *Vet Res.* 2014;45:71.

24. Ahlstrom C, Barkema HW, De Buck J. Improved short-sequence-repeat genotyping of *Mycobacterium avium* subsp. *paratuberculosis* by using matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Appl Environ Microbiol.* 2014;80:534–9.
25. Cochard T, Branger M, Thibault V, Supply P, Biet F. A web tool for analysis the genotyping of *Mycobacterium avium* subsp. *paratuberculosis* strains and other MAC members. In: Proceedings of the Twelfth International Colloquium on Paratuberculosis. Parma, Italy: International Association for Paratuberculosis; 2014. p. 233. 22–26 June 2014.
26. Smeds L, Künstner A. ConDeTri—a content dependent read trimmer for Illumina data. *PLoS One.* 2011;6:e26314.
27. Wynne JW, Seemann T, Bulach DM, Coutts SA, Talaat AM, Michalski WP. Resequencing the *Mycobacterium avium* subsp. *paratuberculosis* K10 genome: improved annotation and revised genome sequence. *J Bacteriol.* 2010;192:6319–20.
28. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25:1754–60.
29. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25:2078–9.
30. Milne I, Stephen G, Bayer M, Cock PJA, Pritchard L, Cardle L, et al. Using Tablet for visual exploration of second-generation sequencing data. *Brief Bioinform.* 2013;14:193–202.
31. Sohal JS, Singh SV, Singh PK, Singh AV. On the evolution of "Indian Bison type" strains of *Mycobacterium avium* subspecies *paratuberculosis*. *Microbiol Res.* 2010;165:163–71.
32. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol.* 2003;52:696–704.
33. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods.* 2012;9:772.
34. Letunic I, Bork P. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res.* 2011;39(Web Server issue):W475–8.
35. Reaman GE. History of the Holstein-Friesian breed in Canada. Toronto: Collins; 1946.
36. Bryant JM, Schürch AC, Van Deutekom H, Harris SR, De Beer JL, De Jager V, et al. Inferring patient to patient transmission of *Mycobacterium tuberculosis* from whole genome sequencing data. *BMC Infect Dis.* 2013;13:110.
37. Boniotti MB, Gorla M, Loda D, Garrone A, Benedetto A, Mondo A, et al. Molecular typing of *Mycobacterium bovis* strains isolated in Italy from 2000 to 2006 and evaluation of variable-number tandem repeats for geographically optimized genotyping. *J Clin Microbiol.* 2009;47:636–44.
38. Grant A, Arnold C, Thorne N, Gharbia S, Underwood A. Mathematical modelling of *Mycobacterium tuberculosis* VNTR loci estimates a very slow mutation rate for the repeats. *J Mol Evol.* 2008;66:565–74.
39. Arnold C, Thorne N, Underwood A, Baster K, Gharbia S. Evolution of short sequence repeats in *Mycobacterium tuberculosis*. *FEMS Microbiol Lett.* 2006;256:340–6.
40. Aandahl RZ, Reyes JF, Sisson SA, Tanaka MM. A model-based Bayesian estimation of the rate of evolution of VNTR loci in *Mycobacterium tuberculosis*. *PLoS Comput Biol.* 2012;8:e1002573.
41. Estoup A, Jarne P, Cornuet J-M. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Mol Ecol.* 2002;11:1591–604.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

