Genotyping methods and molecular epidemiology of *Mycobacterium avium* subsp. *paratuberculosis* (MAP)

Ahmad Fawzy¹, ², ³, *, Michael Zschöck², Christa Ewers³, Tobias Eisenberg², ³

¹ Faculty of Veterinary Medicine, Department of Medicine and Infectious Diseases, Cairo University, Egypt
² Hessian State Laboratory (LHL), Giessen, Germany
³ Institute of Hygiene and Animal Infectious Diseases, Justus-Liebig University, Giessen, Germany

**ABSTRACT**

*Mycobacterium avium* subsp. *paratuberculosis* (MAP) is the causative agent of Johne's disease (JD) which affects mainly ruminants and is characterized by chronic diarrhea and emaciation. Johne's disease is highly prevalent in many countries around the world and leads to high economic losses associated with decreased production. Genotyping of the involved pathogen could be used in the study of population genetics, pathogenesis and molecular epidemiology including disease surveillance and outbreak investigation. Principally, researchers have first assumed the presence of two different MAP strains that are associated with the animal host species (cattle and sheep). However, nowadays MAP characterization depends mainly upon genetic testing using genetic markers such as insertion elements, repetitive sequences and single nucleotide polymorphisms. This work aims to provide an overview of the advances in molecular biological tools used for MAP typing in the last two decades, discuss how these methods have been used to address interesting epidemiological questions, and explore the future prospects of MAP molecular epidemiology given the ever decreasing costs of the high throughput sequencing technology.

1. Introduction

*Mycobacterium (M.) avium* subsp. *paratuberculosis* (MAP) is a member of the *M. avium* complex (MAC). It is the causative agent of paratuberculosis or Johne's disease (JD), a chronic gastroenteritis primarily affecting domestic ruminants and causing high economic losses especially in the dairy industry worldwide [1]. The disease was also found to sporadically exist in wild ruminants [2]. MAP was also identified – but without showing clinical signs – in a wide range of hosts such as non-human primates [3], non-ruminant wildlife [4], dogs [5], feral cats [6], rabbits [7], parrots [8] and bears [9]. Over a century, a possible role for MAP in the pathogenesis of Crohn's disease (CD), a chronic debilitating gastroenteritis affecting humans has been debated. One of the first MAP isolates that was obtained from a Crohn's disease patient is the “Linda strain” [10]. The difficulty of constant and reliable MAP isolation from human CD patients was one of the main obstacles that hindered the proof or refutation of this hypothesis over such a long period of time [11,12].

Molecular epidemiology of infectious diseases is a branch of epidemiology that involves the use of molecular biological techniques in studying the determinants and distribution of disease occurrence in a certain population [13]. In this review, we will focus on the current understanding of MAP diversity, the advances achieved in the molecular biological approaches of MAP genotyping in the last two decades and the epidemiological applications of these approaches.

2. MAP types

Historically, MAP types were described based on epidemiologic and phenotypic characteristics. Based on the host from which MAP strains were first isolated, researchers proposed cattle (C) and sheep (S) type strains as the two main MAP types infecting cattle and sheep, respectively [14]. However, advances in the molecular biology have provided a more robust basis for classification. Using pulsed-field gel electrophoresis (PFGE), two MAP genetic groups were identified that showed only host preferences but not exclusivity and hence a new nomenclature for MAP types (type I and II) has been proposed [15]. Type I isolates were slow growers (more than 16 weeks to achieve visible growth) and strongly associated with sheep; while type II isolates were readily growing (4–16 weeks) strains commonly isolated from cattle, but with a...
broadens the host range. A third group called “intermediate” or “type III” was also characterized and was primarily thought to be different from type S and type C isolates [16]. However, a recent whole genome sequencing-based study revealed that both types (I and III) are subgroups of the sheep or type S [17].

Based on a SNP at bp 223 of the insertion element IS1311, a new MAP type termed “bison” or “type B” was described for isolates originating from bison (Bison bison) in Montana, USA [18]. All IS1311 copies in “type B” isolates had a thymidine transition compared to cytosine in type S isolates. In type C isolates, some copies possessed T and others C nucleotides (nt). Whole genome sequencing analysis later showed that the type B is a subgroup of type C and that some type C strains exhibited the IS1311 profile of type S raising concerns about the reliability of IS1311 analysis as a method of choice to discriminate between types C, S, and B [17]. Subsequent molecular analysis of Indian MAP type B isolates revealed variations from those isolated in the USA and consequently respective isolates were termed “Indian bison type” [19]. Later, the Indian bison type isolates were found to possess a TG deletion at bp positions 64 and 65 of the IS1311 locus 2 [20]. The current classification of MAP types based on a recent whole genome sequencing-based study [17] is illustrated in Fig. 1.

3. Genotyping methods

Many genotyping methods have been developed and applied for MAP characterization. In Table 1, the different methods described in this review are listed.

3.1. Randomly amplified polymorphic DNA (RAPD)

RAPD analysis is a simple and cost effective PCR-based fingerprinting technique. In this analysis, the genomic DNA is directly amplified under low stringency PCR-conditions using a single short primer (10–22 bases) of arbitrary sequence. In the last decade of the 20th century, RAPD was commonly used for subtyping of different organisms such as Mycoplasma [28], M. tuberculosis [29] and M. avium [30]. In Germany, RAPD was used for the first time to characterize MAP isolates (n = 16) from cattle [21]. The isolates displayed heterogeneity; however the data was insufficient to conclude that RAPD could be applied in MAP genotyping. More recently, one primer -among 20 different primers available in a commercial kit- was found to be suitable for the identification and subtyping of MAP and M. avium using RAPD analysis [31]. Six different genotypes were identified for both MAP and M. avium isolates investigated. MAP isolates from cattle, goat and humans were genetically similar, while one sheep isolate had a distinct genetic profile.

3.2. Amplified fragment length polymorphism (AFLP)

AFLP is a rapid PCR-based fingerprinting technique. It involves DNA fragmentation with the help of restriction enzymes followed by ligation of adaptors that are complimentary to the restriction regions. The adapted restriction fragments are then amplified with PCR and visualized on polyacrylamide gels either by autoradiography or fluorescence based methods. A variety of restriction enzymes and adaptors/primers could be combined making AFLP a flexible tool that for different applications such as genetic characterization and genetic mapping. Moreover, AFLP is sensitive to polymorphism detection at the whole genome level.

In one study in the USA, AFLP was used to characterize 104 genetically diverse MAP isolates obtained from different hosts and geographic regions [22]. In contrast to earlier studies, AFLP fingerprints of the MAP isolates obtained from humans did not cluster with either the bovine or ovine isolates [22,32]. Among the studied MAP isolates, bovine isolates showed a low degree of genetic diversity regardless of geographic origin, while isolates obtained from humans and sheep displayed a higher degree of genetic heterogeneity. Contrarily, another American research group found a high degree of genetic diversity among MAP isolates characterized by AFLP using 96 primer set (eleven genotypes for 21 isolates; [33]). More recently, the same research group reported that AFLP could even discriminate between MAP isolates based on epigenetic variations [34]. Despite no sequence differences were detected, MAP isolates obtained from tissue samples possessed distinct AFLP fingerprints compared with isolates cultured from faecal samples. Interestingly, they were able to identify restriction sites that were not digested in the tissue-associated isolates and thus accounting for the apparent heterogeneity. The authors assumed that this could be due to the presence of a DNA sequence they identified upstream of the undigested sites for possible methyltransferase recognition.

AFLP suffers however from some limitations such as the need for a purified, high molecular weight DNA, poor reproducibility especially between different platforms and difficulty of standardization due to a subjective interpretation of the banding patterns [35].

3.3. Pulsed-field gel electrophoresis (PFGE)

PFGE is a fingerprinting-based genotyping technique that was first described by Schwartz et al. in 1984 for the separation of yeast chromosomes [36]. However, it is still the gold standard for characterizations of many pathogenic bacteria [37]. PFGE is a special form of gel electrophoresis where restricted DNA fragments are subjected to periodic reorientation of the electric field relative to the gel direction allowing a better separation of DNA fragments [38]. In Australia, PFGE was used for the first time for genotyping of MAP isolates obtained from different ruminant species [23]. PFGE was able to discriminate between the three main MAP types (I, II and III). As a first step, it has clustered pigmented and non-pigmented MAP isolates into two genetic groups.

| Table 1 | MAP genotyping methods described in this review. |
|---------|-----------------------------------------------|
| Genotyping method | Reference |
| Randomly amplified polymorphic DNA | [21] |
| Amplified fragment length polymorphism | [22] |
| Pulsed-field gel electrophoresis | [23] |
| IS900-restriction fragment length polymorphisms | [14] |
| Mycobacterial interspersed repetitive unit-variable number tandem repeat | [24] |
| Multi-locus short sequence repeats | [25] |
| Single nucleotide polymorphisms | [26] |
| Whole genome sequencing | [27] |

* The first study that developed or used the respective genotyping methods for analysis of MAP isolates.
PFGE was not commonly used in the literature as a MAP typing technique. Few studies have used PFGE for characterization of MAP isolates originating from different European countries [39-42].

3.4. IS900-restriction fragment length polymorphisms (IS900-RFLP)

RFLP is a fingerprinting-based typing approach that involves digestion of an organism's genome using one or more restriction endonuclease enzymes followed by electrophoresis of DNA fragments. These fragments are then transferred to a membrane ("Southern blot"), where they are hybridized with a labelled DNA probe targeting the MAP-specific insertion sequence (IS900). IS900-RFLP utilizes the random nature of IS900 insertion sites that differ from one MAP strain to another. Consequently, upon hybridization each MAP strain would exhibit different banding patterns [43]. Most studies have used one or more of the three enzymes BstEII, PvuII and PstI. However, combining the results of two or more IS900-RFLP reactions employing different restriction enzymes increases the overall discriminatory power/index (DI) of MAP genotyping [44]. DI is a numerical value ranging from 0 to 1 which measures the probability that a certain typing method would discriminate between two epidemiologically unrelated strains of a particular microbial population [45]. IS900-RFLP was also able to differentiate between the three main MAP types “type I or S”, “type II or C” and “type III or intermediate” [14].

IS900-RFLP was the first and the most commonly used MAP genotyping approach for almost two decades [41,44,46-48]. Nevertheless, it has drawbacks limiting its use especially after the advancement of sequencing technologies in the last years. It is laborious, time consuming, requires a substantial amount of cultured material for genomic DNA extraction [49] and the inter-laboratory interpretation of results might be challenging.

3.5. Tandem repeats-based genotyping

Tandem repeats (micro- or minisatellites) are DNA sequences that are repeated in a head to tail manner. They are common in prokaryotic genomes and the repeat copy number may vary even between closely related strains [50]. Based on the repeat size, tandem repeats are classified into three main categories. Repeats ranging from 1 to 9 nt are called microsatellites, short sequence repeats (SSR), or short tandem repeats. Larger repeats (10–100 nt) are usually known as minisatellites or variable number tandem repeats (VNTR), while repeats of more than 100 nt are termed macrosatellites [51]. Two mechanisms were hypothesized as the drivers for the addition or deletion of repeat copies, namely strand slippage caused by DNA polymerase errors on one side and recombination events on the other side [50]. Tandem repeats were first thought to be junk DNA of no function; however, an accumulating body of evidence has shown that they affect the expression of genes associated with bacterial adaptation to the surrounding environment [52]. Epidemiologists have exploited the mutable nature of such repeats to characterize bacterial pathogens based on the polymorphism that results from variations in repeat copy numbers [53]. The characterization is founded on PCR-based amplification of tandem repeat loci followed by calculating the repeat copy numbers at each locus either by size polymorphism mainly in gel or capillary electrophoresis (for VNTR) or by sequencing the PCR products (for SSR) [54,55]. The repeat copy numbers at different tandem repeat loci are then combined together in a numerical code defining the genotype of each isolate. As a PCR-based approach, tandem repeats exhibit some advantages compared with the other typing methods. They have made the characterization faster and less laborious. Furthermore, they enabled direct typing from clinical specimens or from non-viable cultures. Other advantages are reproducibility, standardization, and comparability of results between laboratories [25].

3.5.1. Mycobacterial interspersed repetitive unit-variable number tandem repeat

Mycobacterial interspersed repetitive units (MIRU) are a specific group of minisatellites found only in mycobacteria. They were first described in the genome of M. tuberculosis [56]. Compared with other minisatellites (VNTR), MIRU have a unique structure of two DNA sequence motifs; A (24 bp) 5′-TGACGAGAGGGCGGAGTGAG-3′ and B (29 bp) 5′-GGCCGCGTGGTCAGTCGAGGTAGTT-3′ [57]. Since both types of repeats (MIRU and VNTR) have already been used for MAP characterization in the majority of studies, the nomenclature MIRU-VNTR was established in the field.

MIRU loci (MIRU 1-MIRU 4) were first described in MAC strains by nucleotide basic local alignment search tool (BLASTN) analysis of the previously identified MIRU sequences in M. tuberculosis [24]. These loci were able to discriminate between MAP and other MAC subspecies and between pigmented ovine and all other MAP isolates as well. One year later [58], VNTR loci were investigated in the publicly available complete genome of M. avium strain 104 (GenBank accession number NC_008595.1) using the tandem repeat finder software [59]. By investigating 49 MAP isolates, eight VNTR candidates were tested and five of them were found to be polymorphic dividing the isolates into six genotypes. One of these five VNTR loci (VNTR 1658) was identical to the previously described MIRU 3 locus [24]. In 2007, a set of eight MIRU-VNTR loci (VNTR 3, 7, 10, 25, 32, 47 and MIRU X3 and 292) was discovered in the MAP genome [60]. MIRU X3 and 292 are the same loci as the previously described MIRU 3 and 2, respectively [24]. These eight loci have been used to set up an online database (available at http://mac-inmv.tours.inra.fr/index.php) to facilitate the documentation and comparison of MIRU-VNTR-based MAP genotypes detected in different laboratories all over the world. This has encouraged many researchers to use this MIRU-VNTR panel to carry out epidemiological investigations for MAP isolates from different countries [42,44,47,61-66]. In 2010, a new VNTR locus (VNTR 259) was described which was able to discriminate between MAP type II and type III isolates [67].

MIRU-VNTR loci of MAP were recently found to be a subject of homoplasy and hence compromising their accuracy as genetic markers for phylogenetic analysis [68]. Homoplasy is the independent development of the same alleles in isolates belonging to different genetic lineages of an organism based on convergent evolution [17].

3.5.2. Multi-locus short sequence repeats (MLSSR)

In a first screening of SSR in the MAP genome, an American research group identified 78 perfect mono-, di- and tri-nucleotide repeats by in silico analysis of the first complete genome of MAP strain K10 [25]. Perfect repeats are characterized by an integer number of identical copies of the original repeat sequence, whereas imperfect repeats may possess a sequence of n integer copies plus a fragment of the original repeat sequence. In a preliminary analysis using a subset of diverse MAP isolates (n = 6), eleven of these repeats were found to be polymorphic (SSR 1-SSR 11) [25]. In a sequencing-based approach, these polymorphic loci were subsequently utilized for the typing of 27 genetically diverse MAP isolates. The analysis exhibited a high DI of 0.96 by identifying 20 different SSR types (two major phylogenetic groups) which was sufficient to support epidemiological investigation. Shortly afterwards, American research groups started to use this genotyping approach to characterize MAP isolates originating from different host species and geographical regions in the USA [69-73].

Furthermore, the stability of SSR loci as genetic markers was investigated. In one study, SSR markers remained unchanged (stable) after ten in vitro serial passages of three genetically distinct MAP isolates [73]. Contrarily, SSR locus 2 exhibited a repeat copy number variation both after in vivo passage and in a natural infection context within single cattle herds and individual animals as well [74], hence it was recommended to refrain from using this locus in epidemiological investigations due to its instability. Based on technical limitations of
sequencing a long stretch of a mononucleotide repeat, it was proposed that more than ten guanine bases at SSR locus 1 should be interpreted as ≥11 G [74], which limits the discriminatory capacity of this locus. Different attempts were carried out to overcome this problem by determination of the copy number at this locus either by Matrix-Assisted Laser Desorption Ionization Time of Flight Mass Spectrometry (MALDI-TOF MS; [75]) or by capillary gel electrophoresis-based fragment analysis [76]. This in turn extended the detection limit of copy numbers at SSR locus 1 to 15 and 21 nt, respectively.

3.6. Combination of different genotyping markers

Bacteria are described to be monomorphic, when they exhibit very low genomic diversity. Some important pathogens such as M. tuberculosis, Bacillus anthracis, Yersinia pestis and MAP are monomorphic [77]. With respect to MAP genotyping, different authors therefore used combinations of two or more genotyping techniques to increase the overall typing DI in order to obtain epidemiologically meaningful data (Table 2).

3.7. Single nucleotide polymorphisms (SNP)

Single nucleotide polymorphisms (SNPs) are point mutations of single nt along the genome. They are stable genetic markers that can be used to trace back the phylogenetic history of an organism. The availability of the first complete MAP genome [27] has directed the scientists’ efforts to search for SNPs that could be used for MAP typing and phylogenetic analyses.

Investigation of single or multiple gene sequences such as 65 kDa heat shock protein (hsp65), superoxide dismutase (sodA), chromosomal replication initiator protein (dnaA), DNA polymerase III subunit beta (dnaN), DNA replication and repair protein (recF), DNA topoisomerase IV subunit B (gyrB), enoyl-(acyl carrier protein) reductase (inhA), polypeptide synthase (pks8), DNA gyrase subunit A (gyrA) and PPE family protein (MAP1506) enabled scientists to find SNPs differentiating between the major MAP types I, II and III [26,85–87]. Partial sequencing of IS900, a robust and merely MAP-specific insertion element, revealed SNP characteristics for the MAP types I, II, and III adding a typing feature to the well-established diagnostic value of this element [88]. The emergence of new technologies allowed researchers to adopt novel approaches such as microarray technology and high resolution melting (HRM) for detection of the previously described MAP-specific SNPs [67,89]. Recently, the decrease in the costs of high throughput sequencing (HTS) has made the whole genome sequencing (WGS) of bacteria more accessible to microbiologists. Accordingly, MAP isolates have been sequenced and informative SNPs were used for developing assays that can be used to address various epidemiological questions. For example, seven SNPs were derived from WGS data of human MAP isolates and used to investigate the occurrence of identical lineages of human MAP isolates in other hosts in Australia [90,91]. Similarly, five SNPs were extracted from WGS data of a national Canadian collection of cattle MAP isolates [92,93] and were used to detect the four most common phylogenetic groups in Canada. The preliminary climax was reached, when a comprehensive SNP-based MAP typing assay was developed based on an international collection of MAP isolates for the first time [94]. This was feasible by extracting only 14 SNPs that have been used to classify MAP isolates into 14 different phylogenetic groups.

3.8. Whole genome sequencing (WGS)

Using a shotgun sequencing approach, the first closed and annotated genome sequence of the MAP cattle strain K10 was published in 2005 [27], which has later been revisited and improved [95]. This has facilitated subsequent MAP genetic studies for diagnostic, comparative genomic and epidemiological purposes. It also gave insight into the molecular basis of MAP pathogenicity and virulence characteristics. Since then, researchers have worked on reinforcing the scientific knowledge by publishing complete genomes of other MAP isolates from different geographic areas. Currently, 44 genomes are publicly available in the genome database (https://www.ncbi.nlm.nih.gov/genome/ genomes) of the National Center for Biotechnology Information (NCBI) for MAP isolates obtained from different countries such as India, Egypt, Germany, and the USA [96–98].

The recent application of HTS enabled researchers to investigate the MAP sequence variations at the whole genome level. Consequently, comparative genomic studies were carried out to gain insight into genomic variations between different MAP types obtained from several host species including sheep [99], camel [100], bison [101], human [90,102] and other domestic and wild animals [103]. The exponential decrease in the costs of HTS has enabled the WGS of a large number of MAP isolates. This helped elucidating phylogeography of MAP isolates at a national level in Canada [68,92] and – in the meantime – also on a worldwide scale [17].

4. Epidemiological applications of MAP genotyping

In this section, examples of the application of MAP genotyping in addressing interesting epidemiological aspects of JD will be discussed.

4.1. Within-herd and within-animal strain diversity

Using a wide spectrum of typing techniques, multiple MAP strains were identified in the same herd and even in the same animal [32,39,40,47,62,71–74,76,104–108]. Different explanations were proposed in the literature for the observed multiple strain superinfections. Since animals’ movements are well known as an important risk factor for inter-herd MAP transmission [109], multiple introduction events associated with purchasing subclinically infected animals from different

| Markers’ combination | No. of MAP isolates | No. of genotypes | DI value | Reference |
|----------------------|---------------------|-----------------|----------|-----------|
| MIRU-VNTR + SSR      | 34                  | 18              | Not calculated | [78]      |
|                      | 38                  | 22              | 0.94     | [79]      |
|                      | 35                  | 9               | Not calculated | [80]      |
|                      | 91                  | 25              | 0.93     | [81]      |
| MIRU-VNTR + IS900-RFLP| 71                  | 24              | 0.997    | [44]      |
| SSR + PFGE           | 268                 | 37 (PFGE)       | 0.817    | [40]      |
|                      |                     | 19 (SSR)        |          |           |
| IS900-RFLP + PFGE + MIRU-VNTR | 164 | 44              | 0.941    | [42]      |
| MIRU-VNTR + SSR + IS900-RFLP | 53  | 17              | 0.95     | [47]      |
| MIRU-VNTR + SSR + IS13H1-PCR-REA** | 21  | 11              | Not calculated | [82]      |
|                      |                     | 200             | 0.88     | [83]      |

Did Calculated according to Hunter and Gaston [84]. Values are those reported by the respective studies.

** Restriction endonuclease analysis.
sources is the most popular and likely the most plausible explanation [44,72]. However, one study has assumed that polyclonality could be a basic aspect of the organism’s ecology and has proposed that clinical cases of paratuberculosis might be provoked by multiple MAP-strain infections [62]. A third explanation was that MAP strains exhibiting minor genetic diversity could exist due to within-herd or within-animal evolution over time (microevolution) given the chronic nature of the organism [107]. The extent of microevolution in MAP infections and its possible phenotypic consequences need to be further investigated. Moreover, the potential consequences of multiple strain infections on the clinical outcome, immune reaction, and vaccination efficacy should be addressed in upcoming studies.

4.2. Informing paratuberculosis management policies

One of the most important epidemiological applications of genotyping is tracing back sources of infections and hence providing invaluable information for decision makers concerning disease management policies [110]. In this regard, a high resolution genotyping that could classify strains in the greatest number of groups is necessary for an efficient tracing back process [111].

MAP isolates obtained from cattle showing clinical signs of JD in Queensland, a subtropical Australian area known to have a very low prevalence of JD, were subjected to genotyping using a combination of MIRU-VNTR and MLSSR [112]. The genotypes were compared with national and international strains to identify the potential source(s) of infection and to evaluate the applied control measures. The results indicated that the two identified MAP strains (Q2012 and Q2013) resulted from two different incursion events and were not introduced from other Australian regions, where the disease is more prevalent. Although no definitive clarification of the sources of infections was achieved, the findings of this study provided evidence that restriction of animal movements between high- and low-prevalence regions as a control measure applied has not been violated in Australia.

4.3. Revealing the inter-species MAP transmission

MAP genotypic characterization has helped elucidating MAP transmission events between different host species. Cross-species transmission between cattle and different wildlife animal species has been reported. Using a combination of three genotyping approaches (SSR, MIRU-VNTR, and IS900-RFLP), a German study investigated the transmission of MAP between free-living red deer and farmed cattle sharing the same habitat [47]. Four of the identified 17 genotypes were shared between the two animal species, where one genotype was found to dominate in both of them suggesting a cross-species transmission between cattle and deer [47]. In Australia, a small proportion (1.7%) of macropods co-grazing with sheep was found to be infected with MAP [113]. Genotypic characterization of the isolates using PCR and restriction endonuclease analysis has indicated that a cross-species transmission endonuclease analysis has indicated that a cross-species transmission of the MAP type S has occurred. However, macropods seems not to be a wildlife reservoir of infection, since they were found to shed the organism at low levels [113]. On the other side, wild rabbits are believed to represent a reservoir of MAP infection [114] and a genotyping approach combining IS900-RFLP, PFGE, and MIRU-VNTR has revealed that wild rabbits and cattle within the same region often share the same genotypes [42]. Moreover, MAP isolates obtained from CD patients showed a great homogeneity (based on WGS) with bovine MAP isolates from the same geographic location in Australia indicating a possible zoonotic transmission, presumably through consumption of milk and dairy products [90].

5. Conclusions

The advances in the molecular biological techniques have revolutionized our understanding of the MAP diversity and the epidemiology of JD has witnessed a shift from depending on classical phenotypic traits such as host preference, time to visible growth and colony pigmentation for MAP characterization towards using genotyping approaches. The latter have so far successfully been used for interesting epidemiological applications. This included informing paratuberculosis management policies by identifying the sources of infections, detection of multiple MAP strains both at the intra-herd and intra-animal levels and revealing the inter-species MAP transmission. In the near future, HTS technology is expected to be more accessible to more laboratories throughout the world. Consequently, WGS will transform our ability to characterize MAP isolates at an unprecedented high resolution which would be applied to extend our understanding of MAP molecular epidemiology at intra-animal, intra-herd, regional, national and international levels.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing interests

The authors have no competing interests to declare.

References

[1] Losinger WC. Economic impact of reduced milk production associated with Johne’s disease on dairy operations in the USA. J Dairy Res 2005;72(4):425–32.
[2] Machackova M, Svastova L, Parnova I, Liska V, Smolik J, et al. Paratuberculosis in farmed and free-living wild ruminants in the Czech Republic (1999-2001). Vet Microbiol 2004;101(1):225–34.
[3] Fechner K, Matz-Bensing K, Lamp K, Karpf EJ, Czerny CP, Schäfer J. Detection of Mycobacterium avium subsp. paratuberculosis in non-human primates. J Med Primatol 2017.
[4] Floros M, Leontides I, Kostoulas P, Billinis C, Sofia M, Kyriazakis I, et al. Isolation of Mycobacterium avium subsp. paratuberculosis from non-ruminant wildlife living in the sheds and on the pastures of Greek sheep and goats. Epidemiol Infect 2008;136(5):644–52.
[5] Miller MA, Davey SC, Van Helden LS, Kettnner F, Welton SM, Last R. Paratuberculosis in a domestic dog in South Africa. J S Afr Vet Assoc 2017;88.
[6] Palmer MV, Stoffregen WC, Carpenter JG, Stabel JR. Isolation of Mycobacterium avium subsp. paratuberculosis (MAP) from feral cats on a dairy farm with MAP-infected cattle. J Wildl Dis 2005;41(3):629–35.
[7] Raizman EA, Wells SJ, Jordan PA, DeFrandeke GD, Bey RR. Mycobacterium avium subsp. paratuberculosis from free-ranging deer and rabbits surrounding Minnesota dairy herds. Can J Vet Res 2005;69(1):32–8.
[8] Münster F, Fechner K, Volkel I, von Buchholz A, Caemyn CP. Distribution of Mycobacterium avium subsp. paratuberculosis in the central European Carpathians. J Wildl Dis 2006;42(3):691–5.
[9] Chiodini RJ, Van Kruiningen HJ, Thayer WR, Merkal RS, Coutu JA. Possible role of mycobacteria in inflammatory bowel disease. 1. An unclassified Mycobacterium species isolated from patients with Crohn’s disease. Dig Dis Sci 1984;29(12):1073–9.
[10] Sechi LA, Dow CT. Mycobacterium avium subsp. Paratuberculosis zoonosis – the hundred year war – beyond Crohn’s disease. Front Immunol 2015;6:96.
[11] Fawzy A, Prince A, Hassan AA, Fayed A, Zischik M, Naga M, et al. Epidemiological studies on Johne’s disease in ruminants and Crohn’s disease in humans in Egypt. Int J Vet Sci Med 2013;1(2):79–86.
[12] Foxman B, Riley L. Molecular epidemiology: Focus on infection. Am J Epidemiol 2001;153(12):1135–41.
[13] Collins DM, Gabric DM, De Lisle GW. Identification of two groups of Mycobacterium paratuberculosis strains by restriction endonuclease analysis and DNA hybridization. J Clin Microbiol 1996;28(2):1591–6.
[14] Stevenson K, Hughes VM, Juan LD, Inglis NF, Wright F, Sharp JM. Molecular characterization of pigmented and nonpigmented isolates of Mycobacterium avium subsp. paratuberculosis. J Clin Microbiol 2002;40(5):1798–804.
[15] Juan LD, Mateos A, Domínguez I, Sharp JM, Stevenson K. Genetic diversity of Mycobacterium avium subsp. paratuberculosis isolates from goats detected by pulsed-field gel electrophoresis. Vet Microbiol 2005;106(3-4):249–57.
[16] Bryant JM, Thibault VC, Smith DG, McGuckie J, Heron I, Sevilla I, et al. Phylogenetic exploration of the relationships between strains of Mycobacterium avium subsp. paratuberculosis. BMC Genomics 2016;17:79.
[17] Whittington RJ, Marsh IB, Whittlock RH. Typing of 1311 polymorphisms confirms that bovis (Bison bison) with paratuberculosis in Montana are infected with a
strain of Mycobacterium avium subsp. paratuberculosis distinct from that occurring in cattle and other domesticated livestock. Mol Cell Probes 2001;15(3):139–45.

[19] Sevilla I, Singh SV, Garrido JM, Aduriz G, Rodriguez S, Geijo MV, et al. Molecular typing of Mycobacterium avium subsp. paratuberculosis strains from different hosts and regions. J Clin Microbiol 2002;40(3):1126–31.

[20] Sohal JS, Singh SK, Singh PK, Singh AV, Kumar N. A new marker s1311 I2 PCR-REA for identification of Indian bison type Mycobacterium avium subsp. paratuberculosis. Indian J Biotechnol 2013;12:204–7.

[21] Schleif F, Gerlach GF. Differentiation of Mycobacterium paratuberculosis isolates by rDNA-spacer analysis and random amplified polymorphic DNA patterns. Vet Microbiol 1997;57(2–3):151–8.

[22] Motiwalla AS, Strother M, Amosin 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(4):1694–702.

[23] Turenne CY, Semert M, Cousins DV, Collins DM, Behr MA. Sequencing of hsp65 distinguishes among subsets of the Mycobacterium avium complex. J Clin Microbiol 2006;44(2):433–40.

[24] Pillai SR, Jayarao BM, Gummo JD, Hue EC, Tiwari D, Stabel JR, et al. Molecular epidemiology of Mycobacterium paratuberculosis in cattle and other domesticated livestock. Mol Cell Probes 2001;15(3):139–45.

[25] Fritsch I, Luyven G, Köhler H, Lutz W, Möbius P. Suspicions of Mycobacterium avium subsp. paratuberculosis transmission between cattle and wild-living red deer (Cervus elaphus) by multilocus genotyping. Appl Environ Microbiol 2012;78(4):1132–9.

[26] Sun Z, Li W, Xu S, Huang H. The discovery, function and development of the variable number tandem repeat in different Mycobacterium species. Crit Rev Microbiol 2016;42(5):738–58.

[27] Zhou X, Aarten P, Michaels CW. The role of variable DNA tandem repeats in bacterial adaptation. FEMS Microbiol Rev 2014;38(1):119–41.

[28] Jansen A, Gemeyel R, Versterpen KJ. Unstable microsatellite repeats facilitate rapid evolution of coding and regulatory sequences. Genome Dyn 2012;7:108–25.

[29] Chou CS. Multilocus variable-number tandem repeat analysis as a molecular tool for subtyping and phylogenetic analysis of bacterial pathogens. Expert Rev Mol Diagn 2010;10:5–7.

[30] El-Sayed A, Natur S, Abdel N-EMI, Salem M, Hassaan A, Zschick M. Genotyping of Mycobacterium avium field isolates based on repetitive elements. Int J Vet Med Sci 2011;13(1):36–42.

[31] Salem M, Natur S, El-Sayed AA, Hassan A, Baljer G, Zschick M. Molecular characterization of Mycobacterium avium subsp. paratuberculosis field isolates recovered 2005 from dairy cattle in Germany. J Vet Med Sci 2008;63(5):1406–18.

[32] Supply P, Magdalena J, Himpens S, Loot C. Identification of novel intergenic repetitive units in a mycobacterial two-component system operon. Mol Microbiol 1997;26:991–1003.

[33] Castellanos E, Romero B, Rodríguez S, Juan LD, Bezos J, Mateos A, et al. Molecular characterization of Argentinian Mycobacterium avium subsp. paratuberculosis strains. BMC Microbiol 2012;12:264.

[34] Gieffert A, Correa Muñoz M, Alvarado Pinedo MF, Vaca R, Morrella C, Fiorentino MA, et al. Molecular typing of Argentinian Mycobacterium avium subsp. paratuberculosis isolates by multiple-locus variable-number-tandem repeat analysis. Braz J Microbiol 2015;46(2):557–64.

[35] Rónai Z, Czisvicsik A, Gyurancsik M, Kreizinger Z, Dán Á, Jánosi S. Molecular analysis and MIRU-VNTR typing of Mycobacterium avium subsp. paratuberculosis strains from various sources. J Clin Microbiol 2015;118(2):275–83.

[36] Imperiale BR, Moyano RD, Di Giulio AB, Romero MA, Alvarado Pinedo MF, Santangelo MP, et al. Genetic diversity of Mycobacterium avium complex strains isolated in Argentina by MIRU-VNTR. Epidemiol Infect 2017;145(7):1382–91.

[37] Kruijl M, Lessiau ON, Corrigan D, Ramirez E, Coffey O, A'Mizlany J. Low genetic diversity of bovine Mycobacterium avium subsp. paratuberculosis isolates detected by MIRU-VNTR typing. Vet Microbiol 2017;203:280–5.

[38] Castellanos E, Romero B, Rodríguez S, Juan LD, Bezos J, Mateos A, et al. Molecular characterization of Mycobacterium avium subsp. paratuberculosis and Mycobacterium avium subsp. paratuberculosis II and III isolates by a combination of MIRU-VNTR loci. Vet Microbiol 2015;184:217–22.

[39] Abitron C, Barkema HW, Stevenson K, Zadoks RN, Biek R, Kao R, et al. Limitations of variable number tandem repeat typing identified through whole genome sequencing of Mycobacterium avium subsp. paratuberculosis on a national and herd level. BMC Genomics 2015;16:161.

[40] Ghadiali AH, Strother M, Naser S, Malinge EJ, Sreevatsan S. Mycobacterium avium subsp. paratuberculosis isolated from Crohn's disease patients and animal species exhibit similar polymorphic locus patterns. J Clin Microbiol 2004;42(11):3545–8.

[41] Motiwala AS, Amosin A, Strother M, Manning EJ, Kapur V, Sreevatsan S. Molecular epidemiology of Mycobacterium avium subsp. paratuberculosis recovered from wild animal species. J Clin Microbiol 2004;42(6):7137–20.

[42] Motiwala AS, Strother M, Theus NE, Stich RW, Byram B, Shulaw WP, et al. Rapid detection and typing of strains of Mycobacterium avium subsp. paratuberculosis from
[83] Sohal JS, Arsenault J, Labrecque O, Fairbrother J-H, Roy J-P, Fecteau G, et al. SNP genotyping of Mycobacterium avium subsp. paratuberculosis in Canadian dairy cattle. PLoS One 2016;11(2):e0149017.

[91] Wynne JW, Beller C, Boyd V, Francis B, Gwozdz J, Carajias M, et al. SNP genotyping of Mycobacterium avium subsp. paratuberculosis k10 genome: Improved annotation and revised genome sequence. J Bacteriol 2010;192(23):6319–20.

[93] Amin AS, Huu CY, Darwish SF, Ghosh P, AbdEl-Fatah EM, Belhous TR, et al. Ecology and genonic features of infection with Mycobacterium avium subsp. paratuberculosis in Turkey. Vet Res 2016;47(1):83.

[94] Leao C, Goldstone RJ, Bryant J, McLuckie J, Inacio J, Smith DGE, et al. Single nucleotide polymorphism-based assay for genotyping Mycobacterium avium subsp. paratuberculosis. J Clin Microbiol 2011;49(10):3297–303.

[95] Liao C, Goldstone RJ, Bryant J, Mcluckie J, Inacio J, Smith DG, et al. single nucleotide polymorphism-based assay for genotyping Mycobacterium avium subsp. paratuberculosis. J Clin Microbiol 2011;49(10):3297–303.

[96] 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(23):6319–20.

[97] Amin AS, Huu CY, Darwish SF, Ghosh P, AbdEl-Fatah EM, Belhous TR, et al. Ecology and genonic features of infection with Mycobacterium avium subsp. paratuberculosis in Turkey. Vet Res 2016;47(1):83.

[98] Möbius P, Höhler M, Felder M, Nordsiek G, Groth M, Köhler H, et al. Genomic analysis of the emerging infection with Mycobacterium avium subsp. paratuberculosis in the Arabian camels (Camelus dromedarius). PLoS One 2012;7(2):e31947.

[99] Sohal JS, Sheoran N, Narayanasamy K, Brahmacari V, Singh S, Subodh S. Genetic analysis of local isolate of Mycobacterium avium subsp. paratuberculosis. Vet Microbiol 2009;134(3–4):375–82.

[100] Timms VJ, Hassan KA, Mitchell HM, Neilan BA. Comparative genomics between human and animal associated subspecies of the Mycobacterium avium complex: a basis for pathogenicity. BMC Genomics 2015;16:695.

[101] Huu CY, Wu CW, Talaat AM. Genome-wide sequence variation among Mycobacterium avium subsp. paratuberculosis isolates: a better understanding of Johne's disease transmission dynamics. Front Microbiol 2011;2:236.

[102] Pavlik I, Bejekova L, Pavlas M, Rossylova Z, Koskova S. Characterization by restriction endonuclease analysis and DNA hybridization using IS900 of bovine, ovine, caprine and human dependent strains of Mycobacterium paratuberculosis isolated in various localities. Vet Microbiol 1995;45:311–8.

[103] Cousins DV, Williams SN, Hope A, Eamens JG. DNA fingerprinting of Australian isolates of Mycobacterium avium subsp. paratuberculosis using IS900 RFLP. Aust Vet J 2000;78(3):184–90.

[104] Möbius P, Luyven G, Heitzel H, Kohler H. High genetic diversity among Mycobacterium avium subsp. paratuberculosis strains from German cattle herds shown by combination of IS900 restriction fragment length polymorphism analysis and mycobacterial interspersed repetitive unit-variable-number tandem-repeat typing. J Clin Microbiol 2008;46(3):972–91.

[105] van Hulzen KJE, Heuven HCM, Nielen M, Hoebb J, Santema WJ, Koops AP. Different Mycobacterium avium subsp. paratuberculosis MIRU-VNTR patterns co-exist within cattle herds. Vet Microbiol 2011;148(2–3):419–24.

[106] Davidson AW, Ablstrom C, De Buck J, Whitney HG, Tahan K. Examination of Mycobacterium avium subsp. paratuberculosis mixed genotype infections in dairy animals using a whole genome sequencing approach. PeerJ 2016;4:e2973.

[107] Bannantine JP, Li L, Mwangi M, Cote R, Raygoza Garay JA, Kapur V. Complete genome sequence of Mycobacterium avium subsp. paratuberculosis, isolated from human breast milk. Genome Announc 2014;2(1):1.

[108] Hsu CY, Wu CW, Talaat AM. Genome-wide sequence variation among Mycobacterium avium subsp. paratuberculosis isolates: a better understanding of Johne's disease transmission dynamics. Front Microbiol 2011;2:236.

[109] Bannantine JP, Li L, Mwangi M, Cote R, Raygoza Garay JA, Kapur V. Complete genome sequence of Mycobacterium avium subsp. paratuberculosis, isolated from human breast milk. Genome Announc 2014;2(1):1.

[110] Bannantine JP, Wu CW, Hsu C, Zhou S, Schwartz DC, Bayles DO, et al. Genome sequence of ovine isolates of Mycobacterium avium subsp. paratuberculosis offers insights into host association. BMC Genomics 2012;13:89.

[111] Marquetoux N, Heuer C, Wilson P, Ridler A, Stevenson M. Merging DNA typing and network analysis to assess the transmission of paratuberculosis between farms. Prev Vet Med 2015;123(1):6–13.