Science review report 2011 - 2012

Molecular Strain Typing

Desmond M. Collins, Marian Price-Carter, Geoffrey W. de Lisle

AgResearch NCBID Wallaceville

EXECUTIVE SUMMARY

DNA subtyping of *Mycobacterium avium* subsp. *paratuberculosis* (*MAP*) is an important general tool because of its ability to answer crucial epidemiological questions, and its usefulness in infection, vaccination and pathogenicity studies. In previous years of this programme, we established a subtyping system for New Zealand isolates of *MAP* based on five variable number tandem repeats (VNTRs) and one short sequence repeat (SSR) and used this system to subtype isolates from dairy cattle, beef cattle, sheep and deer. We found four major subtypes that between them accounted for 89% of recent isolates. In the current year, we investigated a further 5 VNTRs and 5 SSRs for their ability to subdivide these four major subtypes. Eight of the assays gave no additional subdivision, VNTR4356 gave a very limited subdivision of the common subtype found in deer and SSR2 gave substantial subdivision of all four common subtypes. SSR2 subtyping is based on changes in the copy number of a single nucleotide which makes it inherently hypervariable. With our isolates it did however appear to be stable enough in most cases for it to be used for epidemiological purposes in situations where more subdivision of strains is required. This is in contrast to SSR1 which we had tested in previous years and had frequently found to give inconsistent results for some isolates. The two new assays for VNTR4356 and SSR2 together with the standard panel of assays were applied to 94 deer isolates of *MAP* obtained from an earlier Johne's Management Limited study and the subtyping results were compared with the acid fast organism (AFO) status of the samples (a measure of disease severity). While the two rarer subtypes found in deer (one of which was the common Type S subtype found in sheep) were both associated with low AFO status, there were too few samples to state definitively that these subtypes are less virulent for deer. In contrast, the two common subtypes that were found (the common deer subtype and the common dairy subtype), were a frequent cause of progressive disease in deer. While the common deer subtype was more often associated with high AFO levels, the difference from the common dairy subtype did not reach significance with the numbers tested. In the remaining two months of the current year, the assays for VNTR4356 and SSR2 are being applied to the dairy and mixed farm isolates that were typed in 2010-11, to better subdivide the isolates from those studies. DNA from the four common *MAP* subtypes found in New Zealand will also be submitted to the Sanger Institute in the UK for genome sequencing.

1. INTRODUCTION

The ability to distinguish between different strains of a pathogenic bacterial species provides a basis for answering important epidemiological questions about sources of infection and spread of disease and enables potential variation in pathogenicity of different strains to be more easily investigated. Since 1980, this process of distinguishing or typing strains has increasingly relied on the direct detection of DNA differences between strains. The history and current methods of typing *M. avium* subsp. *paratuberculosis* (*MAP*) are summarised in a recent review [1]. The first DNA typing system based on restriction endonuclease analysis [2] was quickly superseded by restriction fragment length polymorphism analysis based on the insertion sequence IS*900* [3]. IS*900* typing very clearly distinguished strains of *MAP* into three groups. Strains of one group called Type C (also called Type II) infect cattle and deer and are uncommon in sheep, while the other strains called Type S (also called Types I and III) infect primarily sheep and occasionally deer but in most countries are rarely isolated from cattle [4-7], although exceptions to this have been reported from Iceland [6] and Spain [8]. While IS*900* has been very helpful in distinguishing *MAP* into strains that preferentially infect sheep and cattle, the technique is cumbersome and relatively few differences have been identified within each group. Until six years ago, other DNA based approaches were also cumbersome [9] or relatively non-discriminating [7, 10]. This lack of discrimination has made it difficult to answer questions about sources of infection when one group of strains overwhelmingly predominates. This has been the situation for Type C strains in cattle in many countries including New Zealand. The recent availability of large amounts of DNA sequence from mycobacterial genome projects has led to the identification of many repetitive DNA sequences; often referred to as variable number tandem repeats (VNTRs) and short sequence repeats (SSRs). The recent application of VNTR typing to *MAP* [7, 11-13] indicated that some VNTRs [7, 11, 13], and two SSRs [13], appeared to have a better ability to subtype strains of *MAP* than IS*900* and other methods. Not all VNTRs and SSRs have yet been tested, and those used so far have only been applied to modest numbers of strains from a few regions.

In 2008-10, we selected eight VNTRs and two SSRs that were reported from overseas studies to be particularly useful for subtyping and applied them to archived New Zealand isolates from cattle, deer and sheep. A combination of the best 5 of these VNTRs gave good discrimination of 65 Type C isolates, resulting in a Simpson's index of diversity (h) of 0.69, and this was increased to 0.83 by including results from one of the short sequence repeats, SSR8. By also including the results from SSR1, the diversity index was improved to 0.92. Another finding to emerge from this first study of New Zealand Type C isolates was that one subtype that was common in deer was not found in any of the cattle isolates. The five VNTRs that showed some discrimination among the Type C isolates together with SSR1 and SSR8 were applied to 58 archived Type S isolates. Combining the results from all the VNTRs and the two SSRs resulted in a diversity index of 0.47, indicating that there was much less diversity among New Zealand Type S isolates than Type C isolates. The other major finding from the VNTR and SSR subtyping of Type C and Type S isolates was that in all cases the VNTR/SSR types of Type C isolates were different from those of Type S isolates, indicating agreement with previous IS*900* typing results.

In 2010-11, in collaboration with Livestock Improvement Corporation (LIC) and Massey University, we applied the VNTR/SSR subtyping system established during

the previous two years to 211 isolates obtained from dairy cattle and 154 isolates obtained from beef cattle, sheep and deer on mixed farms. We found that the subtypes of *MAP* infecting animals had changed significantly over the previous 15-20 years. In particular, infection of cattle with Type S isolates is much more prevalent than previously found and Type S isolates are now more common in New Zealand beef cattle than Type C isolates. To what extent these Type S isolates are causing progressive disease in New Zealand cattle is unknown. The frequency of the different subtypes found for the Type C isolates from dairy cattle was substantially different to that found for the Type C isolates from the Massey mixed farm project which included many deer isolates. Infection of individual cattle with more than one subtype was frequent in the LIC study, and this probably reflects a situation where levels of infection are high on some dairy farms. Studies on the stability of SSR1 were also carried out. While the use of SSR1 provided much additional subdivision of the common subtypes, it was ultimately found that it gave inconsistent results for some isolates and could not be relied upon for subtyping.

Three Milestones were approved for the 2011-12 year. In Milestone 3.2.8, we investigated the improvement of the subtyping system by establishing assays for another five VNTRs and another five SSRs and applying them to ten isolates each of the three common subtypes identified in previous years. We have also completed Milestone 3.2.9 that was based on our findings in previous years that had raised issues about the pathogenicity of different types and subtypes of *MAP* for different ruminant species. As a first step in addressing these issues, we subtyped 93 isolates from deer samples that had been supplied by Johne's Management Limited (JML) as part of another study and correlated these results with the number of *MAP* organisms found by histopathology in the same samples. In Milestone 3.2.10, which will not be completed till June 2012, we will apply two new assays identified from Milestone 3.2.8 to improve the subtyping of the LIC and Massey isolates from 2010-11.

2. METHODOLOGY

2.1. *MAP* isolates

For Milestone 3.2.8, we selected 30 *MAP* isolates that had been previously subtyped for the LIC and Massey studies; 10 isolates of the most common Type S subtype, 10 isolates of the most common Type C subtype found in dairy cattle and 10 isolates of the most common Type C subtype found in deer. These isolates were chosen based on differences in geographic origin to increase the chances of including unrelated isolates. For Milestone 3.2.9 we used 94 *MAP* isolates from deer that had been cultured at AgResearch Wallaceville from samples that had been supplied by JML as part of an earlier study. For Milestone 3.2.10 which has yet to be completed we will use the 350 isolates of *MAP* from LIC and Massey University that we subtyped with five VNTRs and SSR8 in 2010-11.

2.2. DNA extraction methods

Isolates of *MAP* were cultured in BACTEC medium containing egg yolk as described by Whittington [16]. DNA was extracted as described by de Lisle et al. [17]. In this method, 0.2 ml of a BACTEC culture was added to 0.5ml glass beads and 1 ml of proprietary ASL buffer (QIAamp DNA stool kit) and shaken strongly in a 2 ml microcentrifuge tube for 2 periods of 20 sec each in a Ribolyser (FastPrep Cell

disrupter; ThermoSavant, Holbrook, New York, USA) set on 6.5, with 1 min cooling on ice between each period. The suspension was heated in a water bath at 95ºC for 10 min, vortexed for 15 sec and then centrifuged at 20,000g for 1 min. For each sample, 1 ml supernatant was removed and placed into a fresh 2 ml tube, half an inhibitEX tablet was added and the mixture was vortexed until the tablet was completely dissolved. Suspensions were incubated at room temperature for 1 min and then centrifuged at 15,000 g for 3 min. The supernatant was transferred to a new tube and centrifuged for another 3 min and then 0.2 ml supernatant was treated with 15 µl of Proteinase K and 0.2 ml of AL buffer and incubated at 70ºC for 10 min. To precipitate the DNA, 0.2 ml of ethanol was added to the suspension and vortexed. The suspension was then added to a Qiagen column and centrifuged for 1 min at 15,000 g. The column was washed with 0.5 ml AW1 buffer, centrifuged, washed a second time with 0.5 ml of AW2 buffer and centrifuged. Residual washing buffer was removed by centrifugation for 3 min and then DNA was eluted with 0.1 ml of AE buffer by centrifugation at 15,000 g for 1 min.

2.3. Selection of additional VNTRs and SSRs to test for Milestone 3.2.8

Shotgun genome sequences of ten Australian strains of *Mycobacterium avium* subsp. *paratuberculosis* (*MAP*) recently became available for access on public DNA databases. These are not completely assembled genomes but contain a collection of large sequenced DNA segments that between them incorporate nearly all of the DNA sequences of those strains. These sorts of partially assembled sequences can be used for many bioinformatic purposes. Bioinformatic analysis of these sequences by us revealed five VNTRs that we had not yet used and which showed some variation in copy number between those strains. Three of these VNTRs, at genome positions 279, 703 and 4356, have previously been investigated by others [11, 14, and 15]; the other two VNTRs were at genome positions 480 and 788. Five SSRs that we had not previously investigated but which had been shown by Amonsin et al [13] to give reasonable discrimination of *MAP* strains were also tested; these SSRs were denoted as 2, 5, 6, 9 and 11. Apart from SSR2 which has a 1-nucleotide repeat, the SSRs we selected had 2-nucleotide or 3-nucleotide repeats.

2.4. PCR amplification conditions for VNTR assays

For PCR analysis of the five VNTR loci chosen for Milestone 3.2.8, 20 ng of *MAP* DNA was subjected to PCR using the forward (F) and reverse (R) primers in Table 1.

Table 1. Primers for VNTR analysis.

VNTR	Primer name	Sequence
279-F	DMC1768	GGGTGTGGAGCTACGACTTC
279-R	DMC1769	GAGCTGCTTGACCAGGTGAT
480-F	DMC1774	GTGGTGGTGCGGCGGTC
480-R	DMC1775	GTCGACGACGGACTGCAGC
703-F	DMC1770	CGGGACATCACAAATACAGAAGAA
703-R	DMC1771	CGCCGCCCGCCGAAAAAG
788-F	DMC1776	CCCGTATTACCTGCAATACACC
788-R	DMC1777	GCATCTATCCGCCACCTTCC
4356-F	DMC1772	GCGTGTCGCCGCCCTGT
4356-R	DMC1773	AACCTGGCCCGCGTCGC

PCR reaction conditions were optimized for each VNTR locus. The 25 µl PCR reactions contained 1 X standard reaction buffer, 1-2 mM $MgCl_2$, 0.2mM dNTPs, 0.8 µM primer mix, 1.25 U Amplitaq Gold (Roche) 1mg/ml bovine serum albumin, 0 - 12% DMSO and 2 µl *MAP* DNA. PCR reaction conditions were denaturation at 94ºC for 10 minutes, then 35 cycles of denaturation at 94 ºC for 45 sec, annealing at 58ºC for 1 minute, extension at 72ºC for 1 minute and a final extension at 72ºC for 7 min. The variable conditions of the PCR reactions optimized for each VNTR are summarized in Table 2.

Table 2. PCR reaction mixtures, annealing temperatures and % agarose gels used for analysis of PCR product size.

VNTR	Repeat length (bp)	[Mg] mM	% DMSO	% agarose gel *
279	54	2	10	2.5
480	56	1	16	2.5
703	53	2	10	2.5
788	30	1	10	2.5
4356	8	1	10	4

*** Using standards of known repeat length**

2.5. PCR amplification conditions for SSR assays

The primers for the SSR loci 2, 5, 6, 9, and 11 were the same as those used by Amonsin et al. [13]. The 50 µl reactions contained 1 X standard PCR buffer, 2 mM $MgCL_2$, 250 µM dNTPs, 0.8 µM primers, 5% DMSO and 0.5 U of TaqGold [13]. PCR conditions were denaturation at 94ºC for 15 minutes, then 35 cycles of denaturation at 94ºC for 45 sec, annealing at 60ºC for 1 min, extension at 72ºC for 2 min and a final extension at 72ºC for 7 min. The PCR products were then outsourced for purification and sequencing using the same primers as for PCR.

3. RESULTS

3.1. Milestone 3.2.8 - Improve the subtyping of *MAP* isolates

The results from applying five new VNTR and five new SSR assays to 10 isolates each of three common *MAP* subtypes are given in Table 3. One of the newly applied VNTRs (4356) showed a different copy number for one of the ten isolates of the most common Type C subtype found in deer. No other differences were found within the groups of ten isolates. VNTR4356 was applied in Milestones 3.2.9 and will also be used in Milestone 3.2.10. The inability of the other VNTRs to detect any further differences within the groups of ten isolates emphasizes how genetically similar the isolates are in each group. While no other differences were found within the groups of ten isolates, three of the VNTRs gave a different result for the Type S group compared to the two Type C groups. This further illustrates the substantial genetic differences between Type S and Type C isolates.

Assays for four of the new SSRs tested showed no differences within each of the three groups of ten isolates but two of them (SSR9 and SSR11) did distinguish the Type S group from the two Type C groups. SSR5 had the same number of repeats in all of the analysed samples, but sequencing of the PCR products revealed that Type S strains were distinguished from Type C strains by a point mutation in the repeated sequence.

SSR6 showed differences between the common Type C subtype in dairy cattle and the common Type S subtype and common Type C subtype in deer, but this distinction is already accomplished by VNTR292 and also by SSR8, which distinguishes all three types. The 1-nucleotide repeat (SSR2) showed four or five difference among each of the groups of ten isolates. This high variability is very reminiscent of the results obtained in previous years with SSR1 which was also a 1-nucleotide repeat. However SSR1 could not be usefully added to the subtyping panel because re-assay of it in multiple clones from selected samples had shown that it was not sufficiently stable to use with New Zealand samples. With this previous problem in mind, we selected five samples from 2010-11 that had given a mixed result for one of the VNTR assays (indicating that two different VNTR types were present in the sample) and assayed them twice for SSR2. In all five cases, the SSR2 assay gave consistent results. We then re-assayed all 30 isolates for SSR2 and found that all but one isolate gave consistent results.

Table 3. Results from applying five new VNTR and five new SSR assays to 10 isolates each of three common *MAP* subtypes

Subtype	Isolate	New VNTR					New SSR				
		279	703	788	4356	480	5	6	9	11	2
Common sheep subtype of Type S	1	1	1	3	7	2	5t*	5	5	4	9
	2	1	1	3	7	2	5t*	5	5	4	9
	3	1	1	3	7	2	5t*	5	5	4	10
	4	1	1	3	7	2	5t*	5	5	4	10
	5	1	1	3	7	2	5t*	5	5	4	10
	6	1	1	3	7	2	5t*	5	5	4	10
	7	1	1	3	7	2	5t*	5	5	4	10
	8	1	1	3	7	2	5t*	5	5	4	11
	9	1	1	3	7	2	5t*	5	5	4	11
	10	1	1	3	7	2	5t*	5	5	4	12
Common deer subtype of Type C	1	2	2	3	9	2	5	5	6	5	9
	2	2	2	3	9	2	5	5	6	5	10
	3	2	2	3	10	2	5	5	6	5	10
	4	2	2	3	9	2	5	5	6	5	10
	5	2	2	3	9	2	5	5	6	5	10
	6	2	2	3	9	2	5	5	6	5	10
	7	2	2	3	9	2	5	5	6	5	11
	8	2	2	3	9	2	5	5	6	5	11
	9	2	2	3	9	2	5	5	6	5	11
	10	2	2	3	9	2	5	5	6	5	13
Common dairy subtype of Type C	1	2	2	3	9	2	5	6	6	5	9
	2	2	2	3	9	2	5	6	6	5	10
	3	2	2	3	9	2	5	6	6	5	10
	4	2	2	3	9	2	5	6	6	5	10
	5	2	2	3	9	2	5	6	6	5	10
	6	2	2	3	9	2	5	6	6	5	10
	7	2	2	3	9	2	5	6	6	5	11
	8	2	2	3	9	2	5	6	6	5	11
	9	2	2	3	9	2	5	6	6	5	11
	10	2	2	3	9	2	5	6	6	5	13

5t* point mutation G to T in the repeated sequence

3.2. Milestone 3.2.9 – Subtyping for JML deer project

The 94 deer isolates of *MAP* from JML samples were subtyped using VNTR4356 and SSR2 together with our standard assay panel of five VNTRs and SSR8. The results are shown in Tables 4 and 5 together with the numbers of acid fast organisms (AFOs) found in microscopic sections of each sample. The AFO results were supplied by

JML. Since all the samples whose isolates are included in the tables were positive on culture only for *MAP*, the number of AFOs gives a measure of the number of *MAP* organisms within that sample. As disease progresses, AFOs will increase, so AFOs can be used as a proxy for the degree of pathogenesis. The results of subtyping without including SSR2 are given in Table 4 and the results with the inclusion of SSR2 are given in Table 5. The three subtypes we have previously referred to as the common deer subtype, the common dairy subtype and the common sheep subtype are indicated in Tables 4 and 5.

Table 4. Correlation of AFOs to *MAP* subtype based on 6 VNTRs and SSR8

Subtype description	***MAP* subtype**[1]	**Total**	**AFO status**[2] **(fraction of total isolates)**					
			a	**b**	**c**	**d**	**e**	**f**
Uncommon	3322249	4	4 (1)					
Common dairy	3322259	14	7 (0.5)	2 (0.14)	2 (0.14)	1 (0.07)	2 (0.14)	
Common sheep	4311137	3	3 (1)					
Common deer	4322249	70	41 (0.59)	1 (0.01)	6 (0.09)	1 (0.01)	18 (0.26)	3 (0.04)
	Mixed	3	2 (0.67)				1 (0.33)	

[1] The seven digit *MAP* subtype gives the allele frequencies of VNTRs and SSR-8 in the following order: VNTR292, VNTR25, VNTRX3, VNTR7, VNTR3, SSR8, VNTR4356.

[2] AFO status: a, no AFOs seen; b, one AFO seen in section; c, one AFO seen in a number of macrophages; d, 2-10 AFOs seen in macrophages; e, 10-60 AFOs seen in macrophages; f, > 60 AFOs seen in macrophages.

The four subtypes found in Table 4 and the frequencies of each subtype were similar to that seen in our earlier study of deer samples last year. Only three samples had mixed subtypes and in one of these there appeared to be both a Type S and a Type C strain. VNTR4356 which had been added to the panel because it distinguished one of ten isolates of the common deer subtype in Milestone 3.2.8 gave disappointing results; it did not distinguish between any of the 70 isolates of the common deer subtype in this study. It did distinguish the common sheep subtype which is a Type S strain from the other isolates which are Type C strains but this is a distinction which is already achieved by several other VNTRs on the panel.

Two of the subtypes with a total of 4 (uncommon) and 3 (common sheep) isolates respectively all had the lowest AFO status of "a". It was unsurprising to find that the common sheep isolate was associated with low numbers of *MAP*, as earlier reports have indicated that it is of low virulence in deer. The subtype 3322249 may also be of lower virulence for deer but the number of isolates is too few to be certain. It is clear from inspecting the AFO status for the 14 isolates with the common dairy subtype and the 70 isolates for the common deer subtype that these subtypes are frequently associated with a high AFO status of "d-f" in deer. While the high AFO values were a little more frequent for the common deer subtype indicating it may be the more virulent strain in deer, the difference in AFO levels between the common deer and common dairy subtypes was not statistically significant (Fisher's exact test). However, even these two subtypes that were sometimes correlated with high levels of

AFOs were more often correlated with a very low AFO status of "a". Two possible reasons for this are that either a majority of deer that become infected with these subtypes have some resistance to becoming clinically affected, or the animals have been culled in the early stages of disease before overt disease has occurred. While JML may have further information on some of these samples that relates to these different possibilities, it is probable that a prospective study would be needed to investigate which possibility is more likely in each case. Another plausible explanation is that some of the common deer subtype are more virulent but are not distinguished from the others by the VNTR and SSR assays.

The inclusion of SSR2 in the subtyping panel substantially increased the number of subtypes from four to 17 (Table 5) which makes the expanded panel of tests of more use for epidemiological studies. However, results with SSR2 should be treated with some caution as, like SSR1, the variation in allele number is due to changes in a single nucleotide, a situation that leads to hyper variation. Although there were many more subtypes when SSR2 was included, none of these showed a marked correlation with a higher AFO status (Table 5). Differences in virulence may be better identified with a combination of VNTR/SSR assays and assays for changes anywhere in the genome of single nucleotides called single nucleotide polymorphisms (SNPs).

Table 5. Correlation of AFOs to *MAP* subtype based on 6 VNTRs and 2 SSRs

Subtype description	***MAP* subtype** [1]	**Total**	**AFO status**[2]					
			a	**b**	**c**	**d**	**e**	**f**
Uncommon	3322249-10	2	2					
	3322249-11	1	1					
	3322249-13	1	1					
Common dairy	3322259-10	6	3	1		1	1	
	3322259-11	1	1					
	3322259-12	5	3	1	1			
	3322259-13	1					1	
	3322259-14	1			1			
Common sheep	4311137-10	1	1					
	4311137-11	1	1					
	4311137-12	1	1					
Common deer	4322249-9	18	12				5	1
	4322249-10	22	12		3		6	1
	4322249-11	17	10		2		4	1
	4322249-12	8	4			1	3	
	4322249-13	5	3	1	1			

[1] The seven digit *MAP* subtype gives the allele frequencies of VNTRs and SSR-8 in the following order: VNTR292, VNTR25, VNTRX3, VNTR7, VNTR3, SSR8, VNTR4356; the SSR2 results are given after the hyphen.

[2] Same as footnote for Table 4.

4. DISCUSSION

Investigation of 10 additional VNTRs and SSRs for their ability to improve the subtyping of New Zealand isolates of *MAP* did provide a substantially increased

number of subtypes. However, nearly all of this improvement relied on use of the hypervariable SSR2 assay. SSR2 was found to give consistent results in 29 of 30 isolates that were retested, so while it is not as highly reliable as the VNTRs and SSR8, its ability to greatly expand the number of subtypes will be a useful attribute in some detailed epidemiological studies. Previously, we had assessed the usefulness of the hypervariable SSR1 sequence and had concluded that the results from its assay were too inconsistent to be used with New Zealand isolates of *MAP*. While SSR2 appears to give somewhat more consistent results than SSR1 for New Zealand isolates of *MAP*, conclusions made from its use need to be tempered with the knowledge that for a small percentage of isolates it may give reliable results. There is a tradeoff when using SSR2; much better discrimination of isolates into subtypes but a slightly less than 100% certainty on the result for any individual isolate. One strategy to increase the reliability of SSR2 analysis would be to repeat the assay on each isolate, although this would also increase the time and cost of subtyping isolates.

Application of the expanded panel of assays to the JML deer isolates indicated that different subtypes may have a different pathogenicity for deer but did not provide conclusive evidence. Even with the addition of the SSR2 assay, which further subdivided the common dairy and common deer subtypes, none of these subtypes was associated significantly more often with high AFO status. Future subtyping of other isolates from animals of known disease status might yield more conclusive results on the pathogenicity of different subtypes for different animal hosts, but it is more likely that a better understanding of the pathogenicity of different subtypes will only be obtained when studies specifically designed to test this are carried out.

5. CONCLUSIONS

1. Many VNTRs that provide moderate discrimination of *MAP* strains in other countries provide little or no discrimination of New Zealand *MAP* strains. This suggests that New Zealand has a less genetically diverse range of *MAP* subtypes than many other countries.
2. A sufficiently large number of VNTRs and SSRs have now been investigated for us to conclude that further investigation of this method of subtyping is unlikely to significantly improve the VNTR/SSR system that we are now using for New Zealand isolates of *MAP*.
3. Use of SSR2 significantly improved the discrimination of the subtyping. However, SSR2 is a hypervariable sequence that does not give consistent results for all isolates. The slight inconsistency that is introduced by including SSR2 in the subtyping panel would by justified in most epidemiological studies by the greatly improved discrimination that it would provide.
4. A significantly better subtyping system than that now being used is only likely to be achieved by developing a subtyping method based on the use of genetic strain differences such as SNPs that are not in hypervariable regions. SNPs are revealed by genome sequencing of *MAP* strains followed by comparison of the genome sequences to each other.

5. While the rarer *MAP* subtypes found in deer were all associated with low AFO status there were too few samples to state definitively that these subtypes are less virulent for deer.
6. Addition of SSR2 further subdivided the common dairy and deer subtypes but none of these subtypes were associated significantly more often with high AFO status.
7. Both the common deer subtype and the common dairy subtype can cause progressive disease in deer but many deer with these subtypes had little evidence of disease progression. This may have been due to resistance to paratuberculosis in some deer or to some deer being culled before their disease had become advanced.

6. RECOMMENDATIONS FOR FUTURE DIRECTION

Projects A – C itemized below have been proposed for year 5. These involve *MAP* genome sequencing as well as providing a subtyping service for intervention studies. There are other *MAP* DNA typing projects that could be carried out to answer a range of important questions. Many of these projects can be carried out on samples already held at AgResearch Wallaceville that have come from other investigations.

A. Dr Karen Stevenson (Moredun Institute, Edinburgh) is responsible for a major European Initiative to genome sequence more that 100 isolates of *MAP*. This work is being carried out at the Sanger Institute in Cambridge, England and is nearing completion, but the panel of isolates they are genome sequencing contains no Australian or New Zealand isolates. Dr Stevenson has asked the Sanger Institute if they would be prepared to genome sequence an additional four isolates from New Zealand as part of this study and the Sanger Institute has indicated they are happy to do this. The JRDC Board has given permission for us to submit one isolate of each of the predominant four *MAP* subtypes in New Zealand to the Sanger Institute for sequencing. The strains are currently being cultured and should be able to be harvested for DNA and the DNA submitted to the Sanger Institute in May 2012. Since Sanger has already sequenced over 100 genomes of *MAP*, inclusion of our four common subtypes in the project will give an excellent indication of how similar our strains are to those found in other countries. In addition, when Sanger provide us with the results, we will compare the New Zealand genomes of *MAP* to each other and use bioinformatic methods to determine sequence differences between the strains and the possible phenotypic significance of some of those differences.

B. Genome sequence three additional strains of each of the four common subtypes of *MAP* that are found in New Zealand so that a total of four strains of each common subtype are available for analysis. Compare the genomes both within and between the groups of four and select SNPs or other genetic differences that are not in likely hypervariable regions and that could be used to improve the current subtyping system for New Zealand isolates of *MAP*. Since genome sequencing has the highest discriminatory power in characterising isolates it can provide information not available from other typing methods and can be used to improve current typing methods for mycobacteria or to develop new ones (18). In one recent study, genomes of 36 *Mycobacterium tuberculosis* isolates of the same VNTR subtype were sequenced and over 200 SNPs were identified (19).

High resolution characterisation of these isolates based on these SNPs was able to greatly enhance the epidemiological investigation, revealing that the transmission of disease was substantially different to that indicated by VNTR typing alone.

C. Investigate the longitudinal changes of VNTR/SSR subtypes of *MAP* in any new intervention studies carried out by the JRDC. These studies should be designed to enable this subtyping to take place so that better conclusions can be drawn from the results with particular regard to the infectivity and pathogenicity of individual subtypes and the identification of the major sources of infection within and between herds and flocks.

D. Develop PCR assays based on potentially useful SNPs identified in B. above and apply the assays to the samples from the JML deer study and to groups of isolates of the four most common subtypes from the Massey and LIC studies.

E Investigate the role of wildlife as a host for *MAP*. As part of a study initiated by Landcare with support from Massey University, we have approximately 200 *MAP* isolates from a number of different wildlife species associated with three heavily infected deer farms. We also have *MAP* isolates from deer samples acquired from these deer farms in the years before during and after the wildlife sampling. The question that needs to be addressed is whether the subtypes in wildlife are the same as those present in domesticated animals.

F. Investigate the longitudinal changes of VNTR/SSR types of *MAP* in farmed deer. Most of these will be Type C isolates. The Wallaceville laboratory has in their culture collections large numbers of isolates of *MAP* from deer dating back twenty years. This covers much of the period when Johne's disease developed into an endemic disease in farmed deer. An investigation of the strain types over this period would reveal which types have predominated, in which localities the types come from (North versus South Islands) and how they compare with the types present in cattle.

G. Investigate the *MAP* subtypes occurring in "outbreaks" of clinical disease in young deer. In some infected deer herds, multiple cases of clinical disease have been observed in young deer, approximately 1 year old. Subtyping would determine whether or not these "outbreaks" are due to a restricted set of Type C subtypes that are more pathogenic in deer than non-outbreak associated Type C subtypes. An alternative hypothesis is that all Type C subtypes present in deer have the potential to cause "outbreaks" in young animals provided a given set of environmental/host factors are present.

H. Investigate selected cultures that have given three PCR products for a single VNTR sequence in order to determine if all three products have arisen because of the presence of three different isolates with different allele copy numbers in the culture or whether there are only two different isolates in the culture and the third PCR product is a PCR artifact.

REFERENCES

1. Collins DM. (2010) Strain characterization of *Mycobacterium avium* subsp. *paratuberculosis*. In: Paratuberculosis: organism, disease, control; MA Behr and DM Collins eds., CABI, Wallingford, UK.
2. Collins DM, de Lisle GW. (1986) Restriction endonuclease analysis of strains of *Mycobacterium paratuberculosis* isolated from cattle. Am J Vet Res 47: 2226-2229.
3. Collins DM, Gabric DM, de Lisle GW. (1990) Identification of two groups of *Mycobacterium paratuberculosis* strains by restriction endonuclease analysis and DNA hybridization. J Clin Microbiol 28: 1591-1596.
4. de Lisle GW, Collins DM, Huchzermeyer HFAK. (1992) Characterization of ovine strains of *Mycobacterium paratuberculosis* by restriction endonuclease analysis and DNA hybridization. Onderstepoort J Vet Res 59: 163-165.
5. de Lisle, G.W., G.F. Yates, D.M. Collins. (1993) Paratuberculosis in farmed deer; case reports and DNA characterization of isolates of *Mycobacterium paratuberculosis*. J Vet Diagn Invest 5: 567-571.
6. Whittington RJ, Taragel CA, Ottaway S, Marsh I, Seaman J, Fridriksdottir V. (2001) Molecular epidemiological confirmation and circumstances of occurrence of sheep (S) strains of *Mycobacterium avium* subsp. *paratuberculosis* in cases of paratuberculosis in cattle in Australia and sheep and cattle in Iceland. Vet Microbiol 79: 311-322.
7. Motiwala AS, Li L, Kapur V, Sreevatsan S. (2006) Current understanding of the genetic diversity of *Mycobacterium avium* subsp. *paratuberculosis*. Microbes Infect 8: 1406-1418.
8. de Juan L, Alvarez J, Aranaz A, Rodriguez A, Romero B, Bezos J, Mateos A, Dominguez L. (2006) Molecular epidemiology of Types I/III strains of *Mycobacterium avium* subspecies *paratuberculosis* isolated from goats and cattle. Vet Microbiol 115:102-110.
9. Stevenson K, Hughes VM, de Juan L, Inglis NF, Wright F, Sharp JM. (2002) Molecular characterization of pigmented and nonpigmented isolates of *Mycobacterium avium* subsp. *paratuberculosis*. J Clin Microbiol 40: 1798-1804.
10. Sevilla I, Singh SV, Garrido JM, Aduriz G, Rodriguez S, Geijo MV, Whittington RJ, Saunders V, Whitlock RH, Juste RA. (2005) Molecular typing of *Mycobacterium avium* subspecies *paratuberculosis* strains from different hosts and regions. Rev Sci Tech 24: 1061-1066.
11. Overduin P, Schouls L, Roholl P, van der Zanden A, Mahmmod N, Herrewegh A, van Soolingen D. (2004) Use of multilocus variable-number tandem-repeat analysis for typing *Mycobacterium avium* subsp. *paratuberculosis*. J Clin Microbiol 42: 5022-5028.
12. Bull TJ, Sidi-Boumedine K, McMinn EJ, Stevenson K, Pickup R, Hermon-Taylor J. (2003) Mycobacterial interspersed repetitive units (MIRU) differentiate *Mycobacterium avium* subspecies *paratuberculosis* from other species of the *Mycobacterium avium* complex. Mol Cell Probes 17: 157-164.
13. Amonsin A, Li LL, Zhang Q, Bannantine JP, Motiwala AS, Sreevatsan S, Kapur V. (2004) Multilocus short sequence repeat sequencing approach for

differentiating among *Mycobacterium avium* subsp. *paratuberculosis* strains. J Clin Microbiol 42: 1694-1702.

14 Castellanos E, Aranaz A, Gould KA, Linedale R, Stevenson K, Alvarez J, Dominguez L, de Juan L, Hinds J, Bull TJ. (2009b) Discovery of stable and variable differences in the *Mycobacterium avium* subsp. *paratuberculosis* Type I, II, and III genomes by pan-genome microarray analysis. Appl Env Microbiol 75: 676-686.

15. Thibault VC, Grayon M, Boschiroli ML, Willery E, Allix-Béguec C, Stevenson K, Biet F, Supply P. (2008) Combined multilocus short sequence repeat and mycobacterial interspersed repetitive unit- variable-number tandem repeat typing of *Mycobacterium avium* subsp. *paratuberculosis* isolates. J Clin Microbiol 46: 4091-4094.

16. Whittington R. (2010) Cultivation of *Mycobacterium avium* subsp. *paratuberculosis*. In: Paratuberculosis: organism, disease, control; MA Behr and DM Collins eds., CABI, Wallingford, UK.

17. de Lisle GW, Cannon MC, Yates GF, Collins DM. (2006) Subtyping by PCR of *Mycobacterium avium* subsp. *paratuberculosis*, an increasingly important pathogen from farmed deer in New Zealand. N Z Vet J 54: 195-197.

18. Schürch AC, van Soolingen D. 2012. DNA fingerprinting of *Mycobacterium tuberculosis*: From phage typing to whole-genome sequencing. Infect Genet Evol. 12: 602-609.

19. Gardy JL, Johnston JC, Ho Sui SJ, Cook VJ, Shah L, Brodkin E, Rempel S, Moore R, Zhao Y, Holt R, Varhol R, Birol I, Lem M, Sharma MK, Elwood K, Jones SJ, Brinkman FS, Brunham RC, Tang P. 2011. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. N Engl J Med. 364:730-739.