Célia Cristina Fialho Leão
Graduated in Microbiology

Molecular tools in the diagnostic and epidemiology of infections caused by members of *Mycobacterium avium* Complex

Dissertation to obtain the PhD Degree in
Biology – Specialization in Microbiology

Supervisor: João José Inácio Silva, Senior Lecturer, U. Brighton, UK
Co-supervisor: Ilda Santos Sanches, Associate Professor, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa
Co-supervisor: Karen Stevenson, Principal Research Scientist, Moredun Research Institute, Edinburgh, UK

Jury:

President: Prof. Doctor António Manuel Dias de Sá Nunes dos Santos

Examiners: Prof. Doctor Helena Maria Vala Correia
Prof. Doctor Maria Isabel Nobre Franco Portugal Dias Jordão

Vowels: Prof. Doctor Ilda Maria Barros dos Santos Gomes Sanches
Prof. Doctor Isabel Maria dos Santos Leitão Couto
Doctor Ana Rosa Pombo Botelho
Prof. Doctor João José Inácio Silva

September, 2015

Célia Cristina Fialho Leão
Graduated in Microbiology

Molecular tools in the diagnostic and epidemiology of infections caused by members of *Mycobacterium avium* Complex

Dissertation to obtain the PhD Degree in
Biology - Specialization in Microbiology

Supervisor: João José Inácio Silva, Senior Lecturer, U. Brighton, UK
Co-supervisor: Ilda Santos Sanches, Associate Professor, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa
Co-supervisor: Karen Stevenson, Principal Research Scientist, Moredun Research Institute, Edinburgh, UK

Departamento de Ciências da Vida
Faculdade de Ciências e Tecnologia
Universidade Nova de Lisboa

September, 2015

Molecular tools in the diagnostic and epidemiology of infections caused by members of *Mycobacterium avium* Complex

© Célia Cristina Fialho Leão; FCT-UNL; UNL
ISBN 978-989-20-6008-8

“A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objectivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor”.

"Life isn't about waiting for the storm to pass...

It's about learning to dance in the rain."

Vivian Greene

To my mother, sister and nephew

To the memory of my father

Acknowledgements

In this section I want to express my genuine gratitude to all the people that contributed, directly or indirectly in some way, to the accomplishment of this work and that made part of my life during these five years.

Firstly, I want to thank Doctor João Inácio for being my supervisor and for providing me the opportunity to realize this work, for his support, guidance and friendship during my academic journey; to Professor Ilda Santos-Sanches and Doctor Karen Stevenson for being my co-supervisors and to Doctor Ana Botelho and Professor Isabel Couto for being my Thesis Advisory Committee members.

I am particularly indebted to Instituto Nacional de Investigação Agrária e Veterinária (INIAV, IP), Portugal and to Moredun Research Institute (MRI), United Kingdom, for accepting me as host institutions and for providing me conditions to perform the work described here.

I want to express my special appreciation to Doctor Ana Botelho for receiving me in the Bacteriology and Mycology Laboratory (INIAV, IP), and to Doctor Karen Stevenson, for accepting me in the pathogenic mycobacteria group (MRI), and for all the knowledge, support, guidance, dedication and friendship that they shared with me during these five years. Without that it wouldn't be possible to carry out the work presented in this dissertation.

I also want to acknowledge Carlos Pinto, from Serviço de Desenvolvimento Agrário de S. Miguel, Ricardo Bexiga from Faculdade de Medicina Veterinária da Universidade Técnica de Lisboa, Helena Vala from Escola Superior Agrária de Viseu and Fernando Esteves from Associação Nacinal de Criadores de Ovinos Serra da Estrela, for providing me the samples used in this study.

Secondly, I would like to thank to all the people that I met and with whom I worked.
From INIAV, IP, Portugal, especially to my dear friends and office mates Ana Canto and Zé Barahona, and to Ana Amaro, for the friendship, happiness, support and for all the moments that we shared during this years. To Lurdes Clemente, Ivone Correia, Maria Helena Ferronha, Mónica Cunha, Teresa Albuquerque, Jacinto Gomes and to all the technician personnel, for the contribution to my working progress. To my colleagues Pedro Costa, Ana Sofia Ferreira, Inês Guinote, and all the students who went through INIAV, IP during these five years.

From MRI, UK, I want to acknowledge Val, Craig, Patricia, Emma, Sam and their families for everything they did for me during the time that I was in Edinburgh and for all the guidance, support and friendship. I also want to thank to Robert Goldstone, David GE Smith, Joyce, and to everyone I met in Edinburgh.

From Neiker-Tecnalia, Spain, I want to express my gratitude to Professor Ramon Juste, Joseba Garrido, Iker Sevilla and all the Neiker's team, for receiving me and for all the support and knowledge they shared with me.

Thirdly, I'm very grateful to my dear friends Sónia, Marlene, Felicia, Catarina, Emanuel, Paula, Fernando and Marcos for always believing in me and in my capability to go further, for all the encouraging words and strength that you gave me and for making part of my life.

Thank you Nuno for all the moments we shared during these years, for listening and for giving me all the support I needed to conclude this stage of my life.

I want to thank to all members of my family that were always there for me and contributed in some way to this journey; especially to my mother Zulmira, to my sister Isabel, to my nephew Afonso and to my cousins for being always by my side, for believing in me and for everything else.

In memory of my father Joaquim; I know that, wherever he is, he will be very proud of me.

Thank you mom and dad for making me the women I am today.

And finally, thank you all!!

Abstract

Mycobacterium avium Complex (MAC) comprises microorganisms that affect a wide range of animals including humans. The most relevant are *Mycobacterium avium* subspecies *hominissuis* (*Mah)* with a high impact on public health affecting mainly immunocompromised individuals and *Mycobacterium avium* subspecies *paratuberculosis* (*Map)* causing paratuberculosis in animals with a high economic impact worldwide.

In this work, we characterized 28 human and 67 porcine *Mah* isolates and evaluated the relationship among them by Multiple-Locus Variable number tandem repeat Analysis (MLVA). We concluded that *Mah* population presented a high genetic diversity and no correlations were inferred based on geographical origin, host or biological sample.

For the first time in Portugal *Map* strains, from asymptomatic bovine faecal samples were isolated highlighting the need of more reliable and rapid diagnostic methods for *Map* direct detection. Therefore, we developed an IS*900* nested real time PCR with high sensitivity and specificity associated with optimized DNA extraction methodologies for faecal and milk samples. We detected 83% of 155 faecal samples from goats, cattle and sheep, and 26% of 98 milk samples from cattle, positive for *Map* IS*900* nested real time PCR.

A novel SNPs (single nucleotide polymorphisms) assay to *Map* characterization based on a Whole Genome Sequencing analysis was developed to elucidate the genetic relationship between strains. Based on sequential detection of 14 SNPs and on a decision tree we were able to differentiate 14 phylogenetic groups with a higher discriminatory power compared to other typing methods.

A pigmented *Map* strain was isolated and characterized evidencing for the first time to our knowledge the existence of pigmented Type C strains.

With this work, we intended to improve the *ante mortem* direct molecular detection of *Map*, to conscientiously aware for the existence of *Map* animal infections widespread in Portugal and to contribute to the improvement of *Map* and *Mah* epidemiological studies.

Keywords

Mah; *Map*; epidemiology; *ante mortem* molecular diagnostic; SNPs analysis; Type C pigmented strains

Resumo

O Complexo *Mycobacterium avium* (MAC) é composto por microrganismos que causam infecções em animais e humanos. Os membros mais relevantes são *Mycobacterium avium* subespécie *hominissuis* (*Mah)*, que causa infeções em doentes imunocomprometidos com elevado impacto na saúde pública humana, e *Mycobacterium avium* subespécie *paratuberculosis* (*Map)*, agente etiológico da paratuberculose em animais, uma doença que acarreta prejuízos económicos consideráveis.

Neste trabalho foi avaliada a relação genotipica entre 28 isolados de *Mah* de origem humana e 67 de origem suína por análise de múltiplas regiões do genoma repetidas em tandem (MLVA) e concluiu-se que os isolados apresentavam elevada variabilidade genética, sem qualquer relação entre genótipos, origem geográfica, hospedeiro ou amostra biológica.

Isolámos pela primeira vez em Portugal estirpes de *Map* de bovinos assintomáticos e evidenciámos a necessidade de técnicas de diagnóstico mais rápidas e robustas para a detecção directa em amostras biológicas. Deste modo, desenvolvemos uma metodologia para amplificação de um fragmento da IS*900* por PCR "nested" em tempo real com elevada sensibilidade e especificidade, acoplado a métodos de extracção de DNA optimizados para amostras de fezes e leite. Com esta técnica detectámos DNA de *Map* em 83% de 155 amostras de fezes de caprinos, bovinos e ovinos e em 26% de 98 amostras de leite de bovinos.

Desenvolvemos também uma técnica para caracterização de isolados de *Map* tendo por base a detecção de SNPs ("single nucleotide polymorphisms") no genoma completo. Pela análise sequencial de 14 SNPs foi possível caracterizar 14 grupos filogenéticos, com um poder discriminatório superior comparando com outros métodos de tipificação.

No decurso deste estudo foi isolada uma estirpe pigmentada de *Map* caracterizada como Tipo C/II.

Com este trabalho pretendemos contribuir para o melhoramento do diagnóstico molecular *ante mortem*, alertar para a existência de *Map* nas explorações Portuguesas e contribuir com novos marcadores de tipificação molecular para estudos epidemiológicos.

Palavras-chave

Mah; *Map*; epidemiologia; diagnóstico molecular *ante mortem*; análise de SNPs; estirpes pigmentadas de *Map* Tipo C

Contents

Acknowledgements .... **VII**

Abstract .... **IX**

Resumo .... **XI**

Contents .... **XIII**

Figures .... **XIX**

Tables .... **XXI**

List of abbreviations and Units .... **XXIII**

Publications and communications .... **XXIX**

***Chapter I* - General Introduction** 1

1.1. *Mycobacterium avium* Complex (MAC) and related Mycobacteria 3

1.2. Historical perspective 4

1.3. *Mycobacterium avium* Complex (MAC) major characteristics and classification 5

1.4. Post-genomic era of *Mycobacterium avium* subspecies 7

1.4.1. Insertion Sequences (IS) 9

1.4.2. Large Sequence Polymorphisms (LSPs) 10

1.4.3. Single Nucleotide Polymorphisms (SNPs) 10

1.5. *Mycobacterium avium* subspecies typing tools 11

1.5.1. Restriction Fragment Length Polymorphism (RFLP) 12

1.5.2. Pulsed-field gel electrophoresis (PFGE) 12

1.5.3. Variable number tandem repeats (VNTR) 13

1.6. Diagnostic of *Mycobacterium avium* (*M. avium*) infections 14

1.6.1. Bacteriological diagnostic 14

1.6.2. Immunological and Serological diagnostic 15

1.6.3. Molecular diagnostic 16

1.7. *Mycobacterium avium* subspecies *hominissuis* (*Mah*) 18

1.8. *Mycobacterium avium* subspecies *paratuberculosis (Map)* 19

1.9. Objectives and outline 24

1.10. References 27

Chapter II **- Relatedness of *Mycobacterium avium* subspecies *hominissuis* clinical isolates of human and porcine origins assessed by Multiple-Locus Variable number tandem repeat Analysis (MLVA)** 35

2.1. Abstract 37

2.2. Introduction 38

2.3. Material and Methods 39

2.3.1. Reference strain 39

2.3.2. Clinical isolates 39

2.3.3. DNA extraction 39

2.3.4. Confirmation of MAH isolates 40

2.3.5. Selection of VNTR *loci* 40

2.3.6. Data analysis 46

2.4. Results 47

2.4.1. VNTR allelic diversity 47

2.4.2. Genetic diversity of *Mah* isolates 47

2.4.3. Multi-dimensional scaling analysis 50

2.5. Discussion 51

Acknowledgements 54

Author's contribution 54

2.6. References 55

Chapter III **- Search for *Mycobacterium avium* subsp. *paratuberculosis* in Portuguese asymptomatic cattle** 59

3.1. Abstract 61

3.2. Introduction 62

3.3. Materials and Methods 63

3.3.1. Faecal samples 63

3.3.2. Culture and identification of isolates 63

3.3.3. DNA extraction and IS*900*-targeted PCR amplification 63

3.3.4. Characterization of *Map* isolates by MLVA 64

3.4. Results 65

3.5. Discussion 66

Acknowledgements 68

Author's contribution 68

3.6. References 69

***Chapter IV* - Effectiveness of nested IS*900*-targeted real time PCR to detect *Mycobacterium avium* subsp. *paratuberculosis* in faeces and milk** 71

4.1. Abstract 73

4.2. Introduction 74

4.3. Materials and Methods 75

4.3.1. Bacterial strains 75

4.3.2. Faecal samples 76

4.3.3. Milk samples 77

4.3.4. Spiked faecal samples 78

4.3.5. Spiked milk samples 78

4.3.6. Bacteriological culture of faeces 78

4.3.7. Bacteriological culture of milk 79

4.3.8. DNA extraction from faeces 79

4.3.9. DNA extraction from milk 79

4.3.10. *TaqMan* probes and primers 80

4.3.11. Identification of *Map* isolates using *F57*-targeted PCR 82

4.3.12. Detection of *Map* in samples using IS*900*-targeted real time PCR 82

4.4. Results 83

4.4.1. Analytical specificity and sensitivity of probes and primers 83

4.4.2. Culture and identification of *Map* from faecal and milk samples 84

4.4.3. Detection of *Map* by IS*900* real time PCR 85

4.5. Discussion 87

Acknowledgements 90

Author's contribution 90

4.6. References 91

Chapter V* - Novel SNP-based assay for genotyping *Mycobacterium avium* subsp. *paratuberculosis 95

5.1. Abstract 97

5.2. Introduction 98

5.3. Materials and methods 99

5.3.1. Selection of genome sequences and *Map* strains 99

5.3.2. Preparation of genomic DNA 101

5.3.3. SNP analysis and phylogenomics 101

5.3.4. SNP selection and primer design 102

5.3.5. Discriminatory power of SNP-based genotyping assay 105

5.4. Results 106

5.4.1. SNP validation and genotyping 109

5.4.2. Discriminatory power of SNP-based genotyping assay 111

5.5. Discussion 111

Acknowledgements 114

Author's contribution........ 114

5.6. References........ 115

Chapter VI - **Molecular characterization of a rare pigmented *Mycobacterium avium* subsp. *paratuberculosis* Type C strain**........ 117

6.1. Abstract........ 119

6.2. Introduction........ 120

6.3. Material and Methods........ 121

6.3.1. Isolation and identification of pigmented colonies........ 121

6.3.2. DNA extraction........ 121

6.3.3. Identification and confirmation of pigmented isolates........ 122

6.3.4. Molecular characterization........ 122

6.4. Results........ 123

6.4.1. Isolation, sub-culture and confirmation of identity of pigmented isolate........ 123

6.4.2. Molecular characterization........ 124

6.5. Discussion........ 128

Acknowledgements........ 130

Author's contribution........ 130

6.6. References........ 131

Chapter VII - **Final Conclusions and Future perspectives**........ 133

7.1. Conclusions........ 135

7.2. Future perspectives........ 144

7.3. References........ 146

Figures

Figure 1.1.	Illustration of the transmission routes of MAC members between environment, wildlife, livestock animals and humans…………………………………………...	5
Figure 1.2.	Schematic representation of the structure of mycobacteria cell wall composed by three main layers..………………………………….......................................	6
Figure 1.3.	Representation of the phylogenetic tree of a subcluster of slow growers mycobacteria, including MAC members, generated by 16S rRNA gene sequences analysis…………………………………………………………………..	8
Figure 1.4.A.	Illustrative representation of the phylogenetic evolution of *M. avium* subspecies based on LSPs and genetic variations and recombination………………………...	11
Figure 1.4.B.	Illustrative representation of the phylogenetic evolution of *M. avium* subspecies based on SNPs analysis…………………………………………………………	11
Figure 2.1.	Dendrogram constructed from the analysis of VNTR profiles including 28, 67 and 2 *Mycobacterium avium* subsp. *hominissuis* strains of human (H), pig (P) and wild boar (WB) origins, respectively………………………………………...	49
Figure 2.2.	Relatedness of *Mycobacterium avium* subsp. *hominissuis* populations of different sources (humans and pigs) and regions (Portugal and Japan) plotted in a multi-dimensional scaling graph based on the genetic distances of the respective VNTR profiles………………………..	51
Figure 4.1.	Schematic representation of complementary *Map* IS*900*-targeted primers and probe used in this study…………………………………………………………	81
Figure 5.1.	Whole genome SNP-based phylogenetic tree of 133 *Map* isolates included in this study………………………………………………………………………….	106
Figure 5.2.	Whole genome SNP-based phylogenetic tree of 115 Type C *Map* isolates…........	107
Figure 5.3.	Work flow with a schematic representation of the decision tree with the sequential numbered steps, SNPs positions, expected bases and SNP profiles obtained based on the SNPs analysis…………………………………………...	108
Figure 6.1.A.	Image of the yellow pigmented mixed with non-pigmented *Map* colonies obtained from faecal culture after 8-10 weeks of incubation at 37ºC on HEYM with mycobactin J………………………………………………………………	123

Figures (cont.)

Figure 6.1.B. Microscopical observation of the pigmented colonies stained by auramin-rhodamin procedure showing fluorescence characteristic of acid-fast bacilli……. 123

Figure 6.2. Image of the yellow pigmented *Map* strains suspension obtained from sub-culture in Middlebrook 7H9 supplemented with 10% OADC and mycobactin J after 2 months of incubation at 37ºC…………………………………………... 124

Figure 6.3. F*57* real time PCR result showing amplification for the F*57* specific *Map* market……………………………………………………………………….. 124

Figure 6.4. Image of the Blast analysis presenting 100% homology of $LSP^{A}20$ (with 8050 bp), found to be present in the pigmented strain genome data…………………… 127

Figure 6.5. Image of the Blast analysis showing no similarity between $LSP^{A}18$ (with 16424 bp) with the pigmented strain genome data (indicating to be absent from the genome)………………………………………………………………………… 127

Tables

Table 1.1.	Characterization of each *M. avium* subspecies based on the presence or absence of the principal Insertion Sequences…………………………………………………	9
Table 1.2.	Commonly used molecular techniques to identify and genotype *M. avium* subsp. *paratuberculosis*………………………………………………………..................	21
Table 1.3.	Summary of the principal characteristics of *M. avium* subsp. *paratuberculosis* strain Types..	23
Table 2.1.	MATR/MIRU allelic profile for a subset of 21 Portuguese *Mycobacterium avium* subsp. *hominissuis* isolates of porcine origins…………………………………….	42
Table 2.2.	MATR/MIRU allelic distribution and diversity for a subset of 21 Portuguese *Mycobacterium avium* subsp. *hominissuis* isolates of porcine origins……….….....	44
Table 2.3.	VNTR *loci* analyzed in this study, respective amplicon sizes and number of tandem repeats……………………………………………………………………….	45
Table 3.1.	Sequences of the primers used for the MLVA characterization of the isolates…...	65
Table 3.2.	Results obtained for the culture-based and molecular detection of *Map* in bovine faecal samples……………………………………………………………………..	66
Table 4.1.	Bacterial strains used for the determination of the analytical specificity of the real time PCR assay and respective results………………………………………..	76
Table 4.2.	Faecal and milk samples used in this study……………………………………….	77
Table 4.3.	Primers and probes used in this study……………………………………………..	81
Table 4.4.	Results for the *Map* detection in faecal and milk samples by culture and real time PCR assays………………………………………………………………………....	85
Table 5.1.	Additional *Map* field isolates used for validation of SNP assay………………….	100
Table 5.2.	PCR primers and restriction endonucleases used in the SNP assay for this study……………………………………………………………………………….	103
Table 5.3.	SNP profiles of Type C *Map* isolates in phylogenetic sub-group A used in this study……………………………………………………………………………….	110
Table 6.1.	Obtained results for the first steps of the decision tree of the SNP assay based on amplification and restriction endonuclease analysis………………………………	125

Table 6.2. Results obtained by the analysis of the remaining 10 SNPs, by amplification and sequencing of the amplified product……………………………………………… 126

List of abbreviations and Units

ADC	Albumin, Dextrose and Catalase
AF	Acid-fast
AFLP	Amplified Fragment Length Polymorphism
AGID	Agar Gel Immunodiffusion
AIDS	Acquired Immunodeficiency Syndrome
aspB	Aspartate aminotransferase gene
ATCC	American Type Culture Collection
BACTEC	Automated blood culture system
BAL	Bronchoalveolar lavage
BHQ	Black Hole Quenche
Blast	Basic Local Alignment Search Tool
BS	Bronchial secretions
CEBAL	Centro de Biotecnologia Agricola e Agro-alimentar do Alentejo
CF	Complement Fixation
CFU	Colony-forming unit
Ct	Threshold cycle
Cy5.5	Cyanine 5.5
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
dGTP	Deoxyguanosine triphosphate
DNA	Deoxyribonucleic Acid
DnaJ	Heat shock protein 40 kD
dNTP	Deoxyribonucleotide triphosphate
DTH	Delayed-type hypersensitivity
dTTP	Deoxythymidine triphosphate

List of abbreviations (cont.)

EDTA	Ethylenediamine tetraacetic acid
ELISA	Enzyme-Linked Immunosorbent Assay
est	Alpha/beta hydrolase gene
EtBr	Ethidium bromide
EURL	European Union Reference Laboratory
EXT	External
F	Forward
FAM	6-carboxyfluorescein
FW	Forward
GC	Guanine and Cytosine
gnd1	Phosphogluconate dehydrogenase gene
gnd2	Phosphogluconate dehydrogenase gene
GPL	Glycopeptidolipids
groEL1	Heat shock protein gene
gyrA	DNA gyrase subunit A gene
gyrB	DNA gyrase subunit B gene
h	Allelic diversity
HEYM	Herrold egg yolk medium
HGDI	Hunter-Gaston Discriminatory Index
HPC	Hexadecylpyridinium chloride
HPLC	High-Performance Liquid Chromatography
Hsp40	Heat shock protein 40 kD
*hsp*65	65-kDa heat shock protein gene
IFN-γ	Interferon-gamma
IHMT	Instituto de Higiene e Medicina Tropical

List of abbreviations (cont.)

INIAV	Instituto Nacional de Investigação Agrária e Veterinária
INT	Internal
IS	Insertion Sequence
ITS	Internal Transcribed Spacer
JP	Japan
lipT	Lipoyltransferase 1 gene
LJ	Löwenstein–Jensen
LN	Lymph nodes
LOD	Limit of detection
LPS	Lipopolysaccharides
LSP	Large Sequence Polymorphism
M .a. avium	*Mycobacterium avium* subspecies *avium*
M .a. hominissuis	*Mycobacterium avium* subspecies *hominissuis*
M .a. paratuberculosis	*Mycobacterium avium* subspecies *paratuberculosis*
M .a. silvaticum	*Mycobacterium avium* subspecies *silvaticum*
M. avium	*Mycobacterium avium*
Maa	*Mycobacterium avium* subspecies *avium*
MAC	*Mycobacterium avium* Complex
Mah	*Mycobacterium avium* subspecies *hominissuis*
Map	*Mycobacterium avium* subspecies *paratuberculosis*
MAPMRI	*Mycobacterium avium* subspecies *paratuberculosis* Moredun Research Institute
Mas	*Mycobacterium avium* subspecies *silvaticum*
MATR	*Mycobacterium avium* Tandem Repeat
MDS	Multi-Dimensional Scaling
$MgCl_2$	Magnesium Chloride

List of abbreviations (cont.)

MGIT	Mycobacteria Growth Indicator Tube
MHC	Major Histocompatibility Complex
MIRU	Mycobacterial Interspersed Repetitive Units
MLSA	Multilocus Sequence Analysis
MLVA	Variable Number Tandem Repeat Analysis
MTBC	*Mycobacterium tuberculosis* Complex
n	Number
NCBI	National Center for Biotechnology Information
ND	Not determined
NGS	Next-generation sequencing
NTM	Nontuberculous mycobacteria
OADC	Oleic, Albumin, Dextrose and Catalase
OIE	World Organisation for Animal Health
P	Probe
PANTA	Polymyxin, amphotericin B, nalidixic acid, trimethoprim and azlocillin
PCR	Polymerase Chain Reaction
pepB	Peptidase B gene
PFGE	Pulsed-Field Gel Electrophoresis
PPD	Purified Protein Derivate
PPE	Pro-Pro-Glu
PRRs	Recognition receptors
PT	Portugal
qPCR	Real time PCR
R	Reverse
RAPD	Random Amplified Polymorphic DNA

List of abbreviations (cont.)

RBC	Red Blood Cells
rDNA	Ribosomal Deoxyribonucleic Acid
REA	Restriction Endonuclease Analysis
recF	Recombination protein F gene
RFLP	Restriction Fragment Length Polymorphism
rpoB	RNA polymerase beta subunit gene
rRNA	Ribosomal Ribonucleic Acid
RV	Reverse
*sig*F	Sigma factors gene
SNP	Single Nucleotide Polymorphisms
sodA	Superoxide dismutase gene
SSR	Short Sequence Repeats
STs	Sequence Types
Taq	*Thermus aquaticus*
TBE	Tris Borate EDTA
TE	Tris EDTA buffer
TET	Tetrachlorofluorescein
TLRs	Tol-like receptors
Tq	TaqMan probe
U	Units
UK	United Kigdom
UNL	Universidade Nova de Lisboa
VNTR	Variable Number Tandem Repeats
WGS	Whole Genome Sequencing
ZN	Ziehl-Neelsen

List of Units

µM	micromolar
µl	microliter
ml	milliliter
mm	millimetre
bp	base pair
ºC	degrees Celsius
g	gram
mM	milimolar
×g	gravity force
%	percentage
$msec^{-1}$	metres per second

Publications and communications

Articles published/accepted in international journals:

- **Leão C**, Goldstone RJ, Bryant J, McLuckie J, Inácio J, Smith DGE, Stevenson K. 2015. Novel SNP-based assay for genotyping Mycobacterium avium subsp. paratuberculosis. In press (doi: 10.1128/JCM.01958-15).

- **Leão C**, Canto A, Machado D, Santos-Sanches I, Couto I, Viveiros M, Inácio J, Botelho A. 2014. Relatedness of *Mycobacterium avium* subspecies *hominissuis* clinical isolates of human and porcine origins assessed by MLVA. Vet Microbiol 173:92-100 (doi:10.1016/j.vetmic.2014.06.027).

Articles published/accepted in national journals/magazines:

- **Leão C**, Amaro A, Santos-Sanches I, Inácio J, Botelho A. 2015. Paratuberculosis asymptomatic cattle as spillovers of *Mycobacterium avium* subsp. *paratuberculosis*: consequences for disease control. Rev Port Ciênc Vet 110: 69-73.
- **Leão C**, Inácio J, Botelho A. 2013. Highlights on *Mycobacterium avium* subsp. *paratuberculosis*: a fastidious yet insidious microorganism. Port Soc Microbiol Mag 2nd year, n. 3.
- **Leão C**, Inácio J, Albuquerque T, Botelho A. 2012. Paratuberculose. Newsletter port Lacticoop. 50 Years of Lacticoop. 250:10-11.

Articles submitted/in preparation:

- **Leão C**, Cruz C, Amaro A, Pinto C, Santos-Sanches I, McLuckie J, Watkins C, Stevenson K, Botelho A, Inácio J. Effectiveness of nested IS*900*-targeted real time PCR to detect *Mycobacterium avium* subsp. *paratuberculosis* in faeces. (submitted)
- **Leão C**, Botelho A, Martins E, Aguiar C, Rebelo I, Bexiga R. Presence of *Mycobacterium avium* subs. *paratuberculosis* DNA in milk used to feed calves in Portugal. (in preparation)
- Barbosa P, **Leão C**, Amaro A, Botelho A, Pinto C, Chaves I, Usié A, Stevenson K, Inácio J, Ramos AM. Genome wide comparison between *Mycobacterium avium* subsp. *paratuberculosis* non-pigmented and pigmented type C isolates. (in preparation)

Oral presentation:

- **Leão C**, Goldstone RJ, Bryant J, McLuckie J, Inácio J, Smith DGE, Stevenson K. Novel SNP-based analysis for the differentiation of isolates of *Mycobacterium avium* subsp. *paratuberculosis*. O-05.5. In Program and Book of Abstracts of the 12th International Colloquium on Paratuberculosis. June 22-26 2014 / Parma, Italy

Poster presentations:

- **Leão C**, Amaro A, Pinto C, Botelho A, Inácio J. Semi-nested probe-based real-time PCR for the detection of *Mycobacterium avium* subsp *paratuberculosis* in goats. P-02.24. In Program and Book of Abstracts of the 12th International Colloquium on Paratuberculosis. June 22-26 2014 / Parma, Italy
- **Leão C**, Amaro A, Botelho A, Inácio J. Screening of asymptomatic Portuguese cattle for the presence of *Mycobacterium avium* subsp. *paratuberculosis* in fecal samples. Poster. In VI Congresso da Sociedade Portuguesa de Ciências Veterinárias - CCV2014. April 3-5 2014 / Oeiras, Portugal.
- **Leão C**, Amaro A, Santos-Sanches I, Botelho A, Inácio J. (2013, December) Assessment of *Mycobacterium avium* subs. *paratuberculosis* faecal shedding in asymptomatic Portuguese cattle. P175. In Program and Book of Abstracts of the Portuguese Congress of Microbiology and Biotechnology' 2013. December 6-8 2013 / Aveiro, Portugal

http://microbiotec13.web.ua.pt/docs/e-book%20Microbiotec%202013%20Aveiro.pdf

- **Leão C**, Amaro A, Santos-Sanches I, Botelho A, Inácio J. Assessment of *Mycobacterium avium* subsp. *paratuberculosis* infection in asymptomatic Portuguese cattle using molecular and culture-based assays. P803. In Program and Book of Abstracts of the XVII World Buiatrics Congress 2012. June 3-8 2012 / Lisbon, Portugal.

http://www.admedic.pt/uploads/Programa-Cient-fico-XXVII-World-Buiatrics-Congress-2012.pdf

- **Leão C**, Inácio J, Sanches I, Botelho A. Genetic heterogeneity of *Mycobacterium avium* subspecies *hominissuis* isolated in Portugal. PS3:38. In Program and Book of Abstracts of the Microbiotec 2011, Portuguese Society of Microbiology and Biotechnology. December 1-3 2011 / Braga, Portugal.

https://microbiotec11.files.wordpress.com/2011/12/e-microbiotec11.pdf

- **Leão C**, Inácio J, Sanches I, Botelho A. Evaluation of *Mycobacterium avium* tandem repeat (MATR) typing of *M. avium* subsp. *hominissuis* strains with a high diversity of IS*1245*-RFLP profiles. P72. In Program and Book of Abstracts of the ESM. June 23-29 2011 / Lübeck, Alemanha.

Other poster presentations:

- Reis A, **Leão C**, Bandeira V, Fonseca C, Cunha MV. On the trail of *Mycobacterium avium* subsp. *paratuberculosis*. Poster. In Program and Book of Abstracts of the Frontiers in E3, cE3c Annual Meeting 2015, Faculdade de Ciências, Lisbon University. June 19-20 2015 / Lisbon, Portugal.

- Amaro A, Nunes M, Gomes J, **Leão C**, Inácio J, Vieira ML. Molecular investigation of *Borrelia burgdorferi* sensu lato in serum of several animal species from mainland Portugal. P31 In Program and Book of Abstracts of the EurNegVec. March 31- April 2 2015 / Izmir, Turkey.

http://www.eurnegvec.org/2ac_abstractbook.pdf

Chapter I

General Introduction

"Science knows no country,
because knowledge belongs to humanity, and
is the torch which illuminates the world."

Luis Pasteur

1.1. *Mycobacterium avium* Complex (MAC) and related Mycobacteria

Mycobacterium avium complex (MAC) is a group of microorganisms belonging to the Kingdom *Bacteria*, Phylum *Actinobacteria*, Order *Actinomycetales,* Suborder *Corynebacterineae*, Family *Mycobacteriaceae* and Genus *Mycobacterium* [1]. This Genus comprises organisms that are aerobic, nonmotile, known as acid-fast microorganisms and with a high level of Guanine and Cytosine (GC) genomic content (between 60 and 70%). Nowadays, more than 150 species are known to belong to this genus [2] being that the most relevant and commonly isolated in laboratory are members of *Mycobacterium tuberculosis* complex (MTBC) and *Mycobacterium avium* complex (MAC).

Traditionally species designations were determined based on phenotypic features and virulence, host range, source of isolation, pathogenicity, growth characteristics and nutrient requirements [3; 4; 5]. With the evolution of molecular methods era, species starts being classified based on phenotypic and genotypic assays. Relationship between MAC members have been assessed and so far it has been discovered that MAC include nine species of slow-growing mycobacteria and a wide subset of unclassified isolates considered as “MAC-others” with different preferences in host, pathogenicity and environment distribution [6]. The two MAC major species are *Mycobacterium intracellulare*, responsible for pulmonary infections in humans and *Mycobacterium avium* (*M. avium)*, causing diverse infections in a wide range of animals.

M. avium is now subdivided into four distinct subspecies: *Mycobacterium avium* subspecies *avium* (*M. a. avium, Maa*) causing avian tuberculosis (or tuberculosis-like infections) in birds, *Mycobacterium avium* subspecies *silvaticum* (*M. a. silvaticum, Mas*) causing avian tuberculosis (or tuberculosis-like infections) in wood pigeons, wild birds and deer, *Mycobacterium avium* subspecies *hominissuis* (*M. a. hominuissuis, Mah*) causing diverse infections in humans and pigs, and *Mycobacterium avium* subspecies *paratuberculosis* (*M. a. paratuberculosis, Map*) the responsible agent for the intestinal infection of ruminants known as paratuberculosis or Johne’s disease [6; 7; 8].

Other species belonging to MAC can also lead to human pulmonary infections like *Mycobacterium colombiense* [9], *Mycobacterium chimaera* [10], *Mycobacterium marseillense*, *Mycobacterium timonense* and *Mycobacterium boucherdurhonense* [11]. The “MAC-others” are also associated with human pulmonary and disseminated infections, *Mycobacterium vulneris* can cause lymphadenopathy and wound [12] and *Mycobacterium arosiense* is responsible for osteomyelitis [13].

1.2. Historical perspective

Historically, Robert Koch presented for the first time in 1882, the agent responsible for human and bovine tuberculosis. Nevertheless, in the early 1890s, based on growth characteristics it was evidenced that a different bacillus was causing avian tuberculosis and it was designated "*Bacillus tuberculosis gallinarum*" but widely known as *Mycobacterium avium* (*M. avium*) [5]. In 1895 Johne and Frothingham described for the first time the bacteria responsible for thickened intestinal mucosa and enlarged mesenteric lymph nodes of a cow. They concluded that the cause of disease was the bacterium that causes tuberculosis in birds and proposed the designation "*pseudotuberculosis enteritis*". In 1906 intradermal tuberculin test was developed by professor Bernhard Bang and it was revealed that different mycobacteria species originated different reactions to the test [14].

The "*pseudotuberculosis enteritis*" was isolated for the first time in 1912 by Twort and Ingram [15] who observed small colonies suspecting that bigger and older colonies of other mycobacteria called *Mycobacterium phlei* were supporting their growth by supplying culture medium with some essential nutrient. Twort renamed the small colonies "*Mycobacterium enteritidis chronicae pseudotuberculosae bovis*, Johne" [15]. Tuberculin test was adopted to skin testing and novel assays like complement fixation and agglutination tests were developed. In 1923 "*Mycobacterium enteritidis chronicae pseudotuberculosae bovis*, Johne" was officially renamed to *Mycobacterium paratuberculosis* (*M. paratuberculosis).* During the next decades infected animals with *M. paratuberculosis* have been described in some continents around the world [14]. In parallel *M. avium* was associated with infections in pigs and it was also isolated from different animals like cattle, sheep, deer, marsupials, primates and from other animals experimentally infected [5; 16].

Between the years 1940 and 1950 several atypical acid-fast agents were isolated from human with pulmonary diseases and novel microorganisms were described and named *Nocardia intracellularis* and *Battery bacillus*, latter considered as the same microorganism and renamed *Mycobacterium intracellulare* (*M. intracellulare*) [5]. As *M. avium* and *M. intracellulare* were almost indistinguishable by the methodologies used at that time the concept of *M. avium-intracellulare* complex was adopted [17].With the introduction of molecular methods the designation and characterization of *Mycobacterium avium* Complex members changed and it became possible to distinguish *M. avium* from *M. intracellulare*. Based on DNA-DNA hybridization analysis, pathogenicity and host range Thorel and colleagues (1990) [18] divided *M. avium* species in three subspecies: *Mycobacterium avium* subspecies *avium* (*Maa*), *Mycobacterium avium* subspecies *silvaticum* (*Mas*) and *Mycobacterium avium* subspecies *paratuberculosis* (*Map*), initially considered as a different species [5; 18].

Nucleic-acid typing assays identified genomic homologies between organisms with different phenotypic characteristics and revealed that MAC includes more taxa than *M. avium* and *M. intracellulare* species.

1.3. *Mycobacterium avium* Complex (MAC) major characteristics and classification

MAC members are known to be non-tuberculous mycobacteria causing diverse non-tuberculous diseases in different animal species and opportunistic infections in humans mainly in immunocompromised individuals [4; 19]. Members from this complex have a worldwide distribution and are commonly found in the environment comprising soil, diverse water sources, aerosols, protozoa and vegetation, remaining viable for several months. Mycobacteria from MAC have the ability to survive and proliferate under extreme conditions like low pH, starvation, hostile temperatures, chloride or ozone treatments and low oxygen levels [20]. Water sources are thought to be the major source of infection to humans. Nevertheless infected livestock and wildlife can also play an important role for transmission of MAC to human beings and to other animals (Figure 1.1).

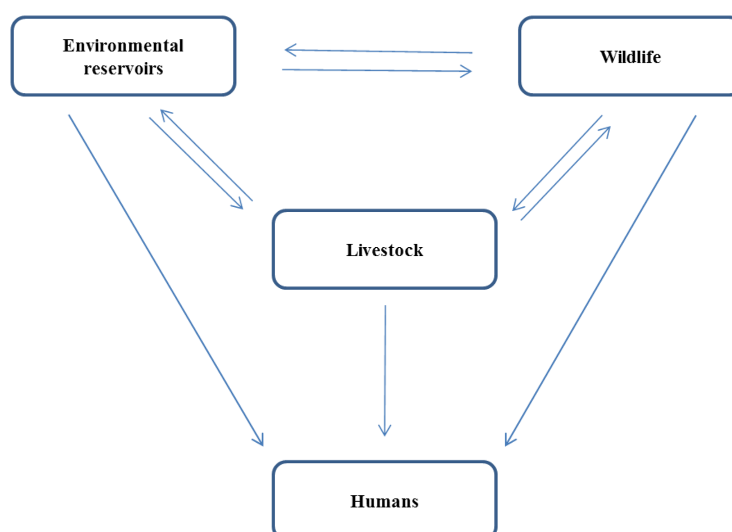

Figure 1.1. Illustration of the transmission routes of *Mycobacterium avium* Complex (MAC) members between environment, wildlife, livestock animals and humans.

Different biological and/or environmental sources can be used for isolation, in solid or liquid culture medium, of MAC members, that are characterized as slow growers with an incubation period ranging from 3-4 weeks to 6 months. Optimal growth conditions vary in a range of pH between 4.0 and 7.5 [20] and temperatures shifting from 20ºC to 42ºC [20; 21] depending on the organism. Colonies can present different morphologies, from smooth opaque or transparent to rough, and were considered in general as non-pigmented; however, some strains may present some yellow pigmentation [5; 19]. Mycobacterial strains are routinely stained by Ziehl-Neelsen (ZN) or Auramin-Rhodamin [22] procedures based on the mycolic acids of the cell wall that provides acid-fast (AF) characteristics to the organisms. Cell wall of the majority of MAC members is composed by three main layers, a first layer is constituted of soluble proteins, carbohydrate, lipids and insoluble macromolecules (arabinogalactan, peptidoglycan and mycolic acids) known as lipopolysaccharides (LPS). Second layer is constituted by glycopeptidolipids (GPL), found only in some organisms from this complex and assumed to be involved in biofilm formation capacity and antimicrobial resistance, and the third layer or external membrane by lipoarabinomannan (Figure 1.2) [3; 19; 20; 23; 24].

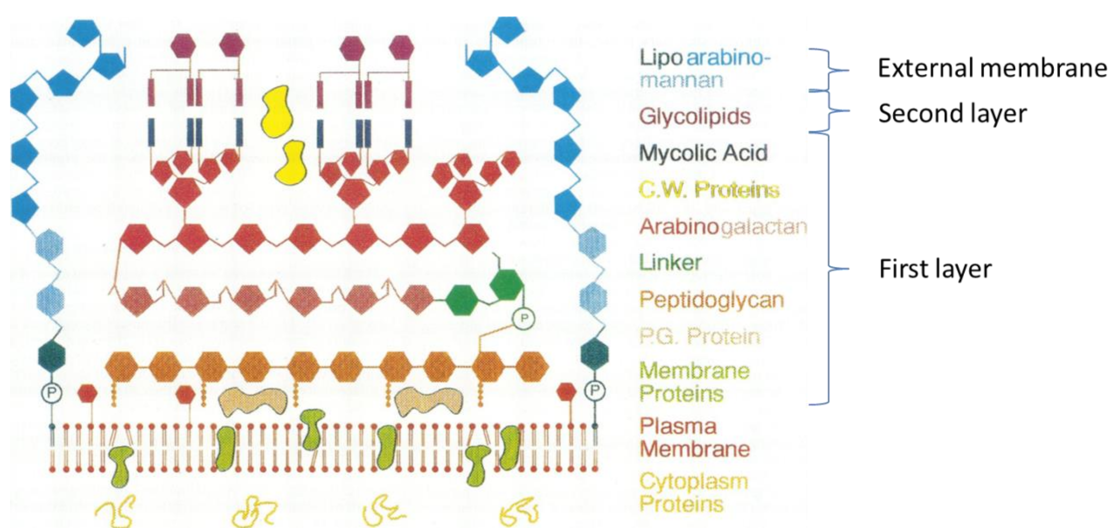

Figure 1.2. Schematic representation of the structure of mycobacteria cell wall composed by three main layers. (Adapted from [19]).

Mycobacteria are facultative intracellular pathogens with the ability to invade, reside and multiply within macrophages. LPS and lipoarabinomannan are important mycobacteria tools used for the invasion of macrophages by interaction with the pathogen recognition receptors (PRRs) and Toll-like receptors (TLRs) present in macrophage cell surface [1].

Once inside the macrophages, mycobacteria have the skill of obstructing the activation of cell, interfering with macrophage apoptosis and preventing the fusion of phagosome and lysosome by inhibiting the decrease of pH levels inside vacuole. This ability of mycobacteria is possible due to the upregulation of expression of certain genomic genes that leads to the characteristic of surveillance inside host cells. Once infected, macrophage starts producing and secreting chemokines, cytokines and their ligands leading to the start of immune response and inflammation procedure. Immune response is promoted by expression of major histocompatibility complex (MHC) class I and II that stimulate the activation of more macrophages, T-help cells, T-cytotoxic cells and Natural Killer cells. This process can lead to the elimination of pathogen and infection or to the formation of granulomas by containment of infection. Failure of host immune defence mechanisms leads to the spread of the infection and establishment of disease [1; 3; 19; 24; 25].

MAC members used to be serotyped into 28 serovars based on the analysis of serovar-specific GPLs. *Mycobacterium avium* (*M. avium*) was classified within serovars 1 to 6, 8 to 11 and 21 and *Mycobacterium intracellulare* (*M. intracellulare*) included serovars 7, 12 to 20, and 25 [7; 19; 26; 27]. However *Mycobacterium avium* subspecies *paratuberculosis* (*M. a. paratuberculosis, Map*) do not produce GPLs and it was not possible to serotype and differentiate these isolates from the remaining *M. avium* subspecies [5]. High-Performance Liquid Chromatography (HPLC) assay was also used in order to distinguish *M. avium* from *M. intracellulare* based on the analysis of a generated species-specific pattern obtained by mycolic acids analysis [28].

1.4. Post-genomic era of *Mycobacterium avium* subspecies

The 16S rRNA gene (16S rDNA) sequence is a highly conserved sequence in the genome, with rare variable regions in certain positions that started to be used to differentiate mycobacterial strains [29; 30]. Internal transcribed spacer (ITS) between 16S rDNA and 23S rDNA was thought to be a good target to differentiate MAC species by amplification and sequencing of the region but minor variations were observed [21; 31]. Because all *M. avium* subspecies have identical 16S rDNA sequence those methodologies are not appropriated for subspecies differentiation (Figure 1.3).

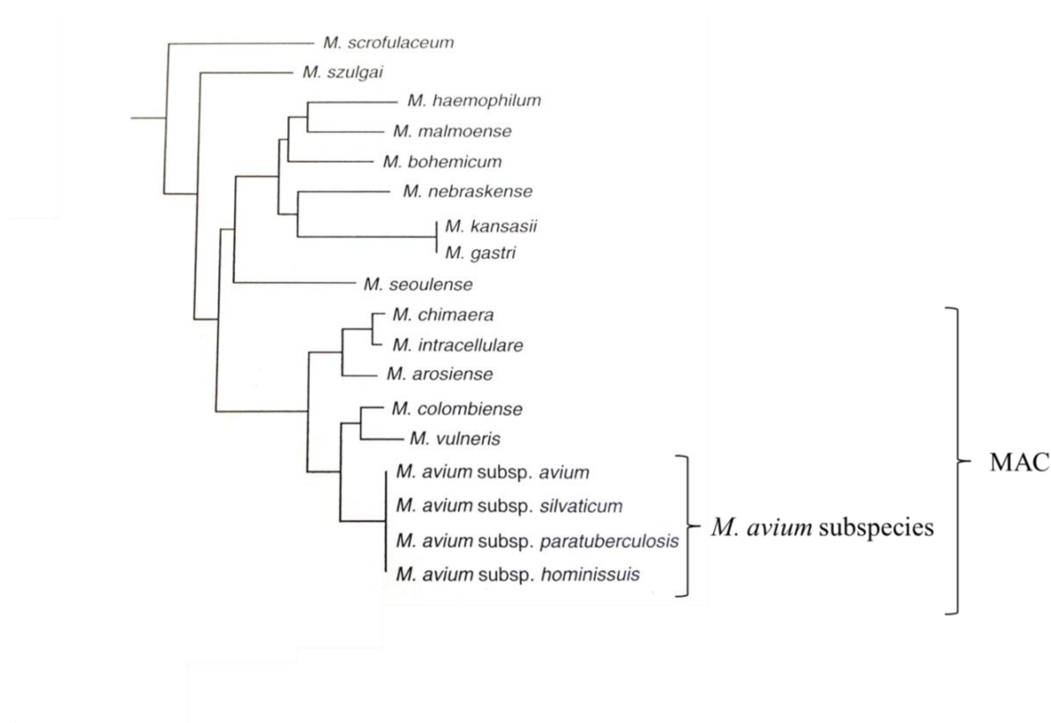

Figure 1.3. Representation of the phylogenetic tree of a subcluster of slow grow mycobacteria, including *Mycobacterium avium* Complex (MAC) members, generated by 16S rRNA gene sequences analysis (Adapted from [5]).

Other genomic regions can differentiate between MAC species, namely genes of 65-kDa heat shock protein (*hsp65*), RNA polymerase beta subunit (*rpoB*), Gyrase B (*gyrB*), and heat shock protein 40 kD (Hsp40 or dnaJ). Restriction fragment length polymorphism (RFLP) analysis of specific insertion sequences in the genome of MAC members have also been used [7; 21; 30; 32]. Nowadays additional and more reliable DNA-based tools are being developed and used in order to better understand the genetic and epidemiological traits of MAC members and the diseases they cause.

Availability of complete genome sequences from mycobacteria strains lead to the development of novel assays to distinguish and characterize different isolates. Two complete genomes from *M. avium* subspecies were sequenced and annotated for further studies, a *M. a. hominissuis* strain 104 with 5475491 bp [33], isolated from an Acquired Immunodeficiency Syndrome (AIDS) patient, and *M. a. paratuberculosis* strain K10 with 4829781 bp [34] isolated from a cow. Despite of a worldwide distribution, different host preferences and triggered infections, *Mycobacterium avium* subspecies share >98% of homology in the whole genome [6; 35]. Based on comparison of complete genomes some specific genomic markers were found to be useful for the identification and typing of *M. avium* subspecies and isolates from the same subspecies. The most important genomic regions for *M. avium* subspecies identification are Insertion Sequences (ISs), Large Sequence Polymorphisms (LSPs) and Single Nucleotide Polymorphisms (SNPs).

1.4.1. Insertion Sequences (IS)

An insertion sequence is a small transposable element found in the bacterial genome in diverse copy numbers, usually flanked by inverted repeat fragments.

There are at least four ISs described as the most important elements used to distinguish between *M. avium* subspecies: IS*1245*, IS*1311*, IS*901* and IS*900* (Table 1.1). IS*1245* [36] is an element with 1414 bp and can be found in *M. a. avium* (*Maa*), *M. a. silvaticum* (*Mas*) and *M. a. hominissuis* (*Mah*) in different copy numbers [6]. *Maa* and *Mas* have three copies of IS*1245* while *Mah* have more than seven copies in the genome [7]. IS*1311* [37] is an element with 1317 bp and shares 85% of homology with IS*1245*. This insertion sequence can be found in all the *M. avium* subspecies [6; 7]. IS*901* [38] have 1472 bp and is present in *M. a. avium* and *M. a. silvaticum* isolates. This element has 99% of similarity with another sequence designed IS*902* [39] only found in *Mas* strains but some authors consider that both correspond to the same element. IS*901*/IS*902* are known to have a very low mobility with very limited polymorphism [6; 7; 21]. The fourth principal insertion sequence is the IS*900* [40] found only in *M. a. paratuberculosis* (*Map*) isolates in 15 to 20 copies in the genome, depending of isolates. This element has 1451 bp and shares 60% of similarity with IS*901* [7]. This insertion sequence is routinely used in typing techniques to identify and discriminate between *Map* strains isolated from different hosts. However, isolates with IS*900*-like sequences have been described to be rarely found in environmental mycobacteria [6; 41].

Table 1.1. Characterization of each *M. avium* subspecies based on the presence or absence of the principal Insertion Sequences.

	Insertion Sequence				
***M. avium* subspecies**	**IS*1245***	**IS*1311***	**IS*901***	**IS*902***	**IS*900***
Maa	+	+	+	-	-
Mas	+	+	+	+	-
Mah	+	+	-	-	-
Map	-	+	-	-	+

Maa – *Mycobacterium avium* subspecies *avium*; *Mas* – *Mycobacterium avium* subspecies *silvaticum*; *Mah* – *Mycobacterium avium* subspecies *hominissuis*; *Map* – *Mycobacterium avium* subspecies *paratuberculosis*; + IS present; - IS absent

1.4.2. Large Sequence Polymorphisms (LSPs)

LSPs are large genomic fragments commonly associated with insertion sequence elements that can be present or absent in the genome [6].

Based on the analysis of complete genome of *M. a. hominissuis* 104 a whole-genome DNA microarray study was performed by Semret and colleagues (2004) [33] in order to represent the predicted coding sequences of the genome and to compare with other *M. avium* subspecies. This study revealed the presence of 14 LSPs (LSP^A1 to LSP^A14) unique in *Mycobacterium avium* subspecies *hominissuis* (*Mah*), where seven LSPs are simple genomic deletions or insertions and the remaining seven encompass a more complex combination of insertions and deletions indicating both vertical and horizontal acquisition of DNA [33]. Another similar study identified 24 LSP^As present in *Mah*, 13 already identified and 11 LSPs newly described, and an additional 18 LSP^Ps (MAP-1 to MAP-18 or LSP^P1 to LSP^P18) only present in *Mycobacterium avium* subspecies *paratuberculosis* (*Map*) were described [35; 42]. From those studies only three LSPs (LSP^A8, LSP^A18 and LSP^A20) were identified in *M. avium* subspecies except in *Map* genome and none of them can distinguish *Mycobacterium avium* subspecies *avium* (*Maa*) from *Mycobacterium avium* subspecies *silvaticum* (*Mas*). Deletion of the LSP^A8 is a specific marker for characterization of *Map* isolates and the absence of LSP^A17 is characteristic of *Maa* and *Mas* strains [6; 42]. LSP^P4, LSP^P11, LSP^P12, LSP^P14, LSP^P15 and LSP^P16 were reported to be specific and exclusive insertions for identification of *Map* isolates [42; 43].

1.4.3. Single Nucleotide Polymorphisms (SNPs)

A SNP is a DNA sequence variation in just a single nucleotide in the genome. This variation occurs more frequently in non-coding regions than in coding regions and is usually a consequence of natural selection adaptation. If a SNP is present in a coding region it can be classified as synonymous, if it doesn't change the aminoacid, or non-synonymous if it triggers an aminoacid change in the protein [44].

DNA polymorphisms in *Mycobacterium avium* (*M. avium*) isolates are mostly related with the presence of SNPs in the genome. Many studies have been reported based on the identification of SNPs in specific regions of the genome of *M. avium* subspecies [6]. Turenne and colleagues (2008) [4] reported a study with the analysis of SNPs in 10 genes (*recF*, *sodA*, *aspB*, *gnd1*, *lipT*, *pepB*, *gnd2*, *est*, *hsp65* and *groEL1*) by Multilocus Sequence Analysis (MLSA). These genes were amplified and the product was sequenced in order to identify regions with higher variability between subspecies and 205 polymorphisms were identified representing 33 Sequence Types (STs). *Mycobacterium avium* subspecies *avium* (*Maa*) was identified by having 6 STs, *Mycobacterium avium* subspecies *silvaticum* (*Mas*) 2 STs, *Mycobacterium avium* subspecies *hominissuis* (*Mah*) 18 STs and *M. a. paratuberculosis* 7 STs. Only four SNPs allow the distinction between *Maa* and *Mas,* and *Mah* presented the most

variability between all subspecies. Other studies have been used for differentiation of *Map* strains based on SNPs identification in different parts of the genome including *gyrA* and *gyrB* genes [45], PPE protein family genes [46], *hsp65* gene [47], IS*900* [48], and IS*1311* [49].

Genomic studies provided theories about phylogenetic evolution of *M. avium* subspecies based on LSPs, SNPs and genetic variations and recombination. Some hypotheses have been described suggesting *Mah* as an ancestor for the evolution of the remaining *M. avium* subspecies, once it has the longest and the most variable genome between subspecies. Proposed evolutionary model speculates two distinct lineages emerging from the ancestor *Mah*. One lineage originates *Maa* and *Mas* subspecies and the second lineage the *Map* strains (Figure 1.4) [4; 6; 43].

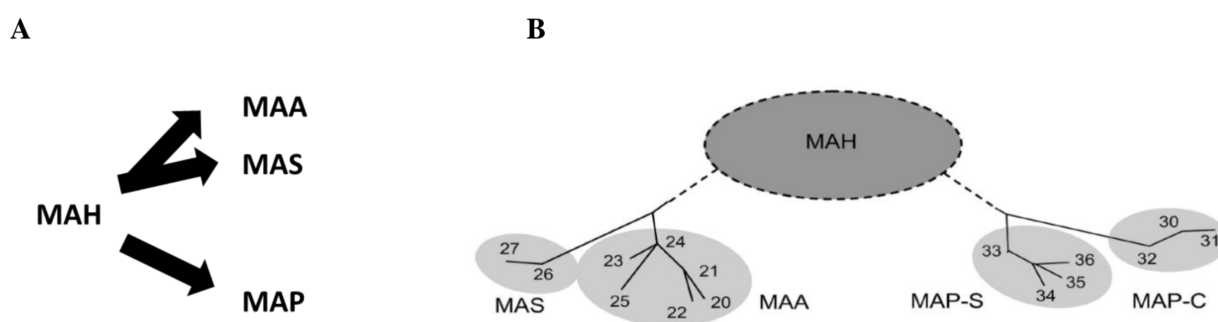

Figure 1.4. Illustrative representation of the phylogenetic evolution of *Mycobacterium avium* subspecies based on Large Sequence Polymorphisms (LSPs) and genetic variations and recombination (A, adapted from [6]) and from Single Nucleotide Polymorphisms (SNPs) analysis (B, adopted from [4]).

1.5. *Mycobacterium avium* subspecies typing tools

Different typing assays have been used for characterization and identification of clonal lineages of *Mycobacterium avium* subspecies, contributing to epidemiological studies of infectious diseases. Restriction fragment length polymorphism (RFLP); pulsed-field gel electrophoresis (PFGE) and variable number tandem repeats (VNTR) have been the most widely used assays. However, a few studies are also available reporting Amplified Fragment Length Polymorphism (AFLP) [50; 51; 52] and Random Amplified Polymorphic DNA Analysis (RAPD) [53; 54; 55].

1.5.1. Restriction Fragment Length Polymorphism (RFLP)

RFLP is a typing technique used to differentiate *Mycobacterium avium* (*M. avium*) isolates based on the restriction of all genomic DNA and, after transference of the obtained fragments by Southern-blot, the hybridization with specific probes designed for targeting, namely, Insertion Sequences (ISs). Extracted DNA from isolates is hydrolysed with a specific restriction enzyme that recognizes and cuts the DNA in specific regions. The obtained result is a set of bands with different sizes in a stained agarose gel. DNA fragments are transferred to a nylon membrane and hybridized with a specific labelled probe in order to visualize the specific bands. Results are revealed with observation of a profile characteristic of each analysed isolate. IS*1245* RFLP using PvuII restriction enzyme is a standardized method for the routinely differentiation of *M. avium* subspecies [56]. By the obtained IS*1245* RFLP patterns is possible to distinguish *Mycobacterium avium* subspecies *avium* (*Maa*), *Mycobacterium avium* subspecies *silvaticum* (*Mas*) and *Mycobacterium avium* subspecies *hominissuis* (*Mah*). *Maa* is characterized by presenting three bands, *Mas* one band and *Mah* isolates can be characterized by multiple and variable bands [21]. IS*901* RFLP had been standardized by Dvorska and colleagues (2003) [57] using the restriction enzyme PvuII or PstI with the identification of multiple RFLP patterns. In this study it was reported the association of high IS*901* copy numbers with bird RFLP profile and low IS*901* copy numbers associated with human/pig RFLP profiles, and different RFLP patterns were described to distinguish between virulent and non-virulent strains [57]. As *Mycobacterium avium* subspecies *paratuberculosis* (*Map*) does not have IS*1245* in the genome IS*1245* RFLP can't be used for *Map* typing. Therefore, IS*900* RFLP was proposed for standardization by Pavlik and colleagues (1999) [58] using the restriction enzyme PstI and BstEII with the identification of 28 different RFLP patterns.

1.5.2. Pulsed-field gel electrophoresis (PFGE)

PFGE requires entire DNA from isolates that is restricted with specific restriction enzymes and submitted to an electrophoresis separation using an electric field that periodically changes direction, originating a profile with different band sizes. For this assay, isolates are grown in liquid culture and cells are mixed with a low melting temperature agarose at 55ºC and cooled in molds in order to form agarose plugs of cells. These cells are then submitted to lysis and a subsequent digestion with specific restriction enzymes, in separate reactions, which cut the total DNA in several large fragments. Each restriction enzyme has specific digestion and electrophoretic conditions. Digested plugs are poured in agarose gel, sealed and the big size DNA fragments are separated by electrophoresis [44; 59]. Many studies had been reported with the use of this typing technique for the differentiation of clinical and environmental *Mycobacterium avium* (*M. avium*) isolates reporting the use of XbaI or AseI restriction enzymes [60; 61; 62]. PFGE had been especially used to distinguish

between *Mycobacterium avium* subspecies *paratuberculosis* (*Map*) strains using SnaBI and SpeI restriction enzymes with the observation of multiple PFGE profiles [59; 63; 64; 65]. A *Map* PFGE Database is available to data submission in order to find the correspondent PFGE profile of *Map* isolates (http://www.moredun.org.uk/research/research-@-moredun/diseases-of-the-gut/paratb-(johnes-disease)/mycobacteria-pfge-database).

1.5.3. Variable number tandem repeats (VNTR)

VNTR is a short region of the genome (10-100 bp), characterized as mini-satellites, repeated several times in tandem at defined loci of the genome. The assay is based on amplification of each locus with specific primers and observation of the band sizes in an agarose gel. Band size is proportional to the number of repeats of a locus. The obtained result is a numerical code composed by a representation of the number of copies in each specific locus allowing the comparison of results with other studies performed in other laboratories [44]. VNTR assay was initially associated with *Mycobacterium tuberculosis* typing with the identification of several specific regions in the genome named Mycobacterial Interspersed Repetitive Units (MIRU-VNTR) [66]. Other regions had been identified in *Mycobacterium avium* subspecies genome designated as Mycobacterium Avium Tandem Repeat (MATR-VNTR). A combination of VNTR, MIRU-VNTR and MATR-VNTR has been used for *Mycobacterium avium* subspecies characterization and for differentiation of strains from the same subspecies.

Based on the analysis of four MIRU-VNTR loci, Bull and colleagues (2003) [67] distinguished *Mycobacterium avium* subspecies *paratuberculosis* (*Map*) from other MAC members by the observation of the band pattern obtained by MIRU2 and MIRU4. Thibault and colleagues (2007) [68] described eight MIRU-VNTR loci, seven reported as new (32, 292, 7, 10, 25, 47 and 3) and one (senX3-regX3) already reported [67], to be useful for the characterization of *Map* and other *Mycobacterium avium* (*M. avium*) isolates. From 82 *M. avium* isolates, it was identified 30 MIRU-VNTR profiles and from 183 *Map* isolates 21 types were identified. In this study, MIRU32 and MIRU292 showed the highest allelic diversity for *Map* while for the remaining *M. avium* isolates it was X3 locus. In 2010, Pate and colleagues [69] typed *Mycobacterium avium* subspecies *avium* (*Maa*) and *Mycobacterium avium* subspecies *hominissuis* (*Mah*) isolates from animals and humans, with the eight MIRU-VNTR loci and identified 24 different patterns between 121 isolates. The highest allelic diversity locus observed for *Maa* and *Mah* was X3, as well. A study with the characterization of clinical isolates from AIDS patients from Japan was performed using MATR and MIRU-VNTR assays, with the analysis of 15 MATR loci (MATR1-MATR15) and the eight MIRU-VNTR loci. The most discriminative MATR loci were 7, 2 and 3, and MIRU loci were X3 and 292 [70]. Several other studies reported the utility of MIRU-VNTR and MATR-VNTR to characterize *Mah* and *Map* isolates

from environmental and clinical sources with diverse patterns identified [71; 72; 73; 74; 75; 76; 77; 78]. A web classification database “MAC INMV database” (http://mac-inmv.tours.inra.fr/) was generated based on the obtained profiles from the analysis of the eight MIRU-VNTR used for *M. avium* subspecies typing. Obtained genotypes are classified in INMV profiles and can be globally compared with other isolates. So far, 141 INMV profiles are described and available at “MAC INMV” database.

1.6. Diagnostic of *Mycobacterium avium* (*M. avium*) infections

Diagnostic of *M. avium* subspecies infections can be *ante mortem* or *post mortem*. *Ante mortem* diagnostic is assessed by clinical observation and non-invasive biological samples testing. *Post mortem* diagnosis is achieved by observation of pathognomonic lesions and histopathology.

1.6.1. Bacteriological diagnostic

The gold standard diagnostic method for *M. avium* subspecies is culture of the agent in specific culture medium with a specificity of 100%, but it can be expensive, technically difficult and time consuming requiring up to 6 months of incubation [79]. Different liquid and solid culture media can be used. Most commonly used liquid medium for culture of *M. avium* subspecies is Middlebrook 7H9 supplemented with 10% albumin-dextrose-catalase (ADC) or oleic-albumin-dextrose-catalase (OADC). Automatic systems like Mycobacteria Growth Indicator Tube (MGIT) systems can also be used. BACTEC MGIT 960 System is composed by liquid medium Middlebrook 7H9 supplemented with BBL MGIT PANTA Antibiotic Mixture (polymyxin, amphotericin B, nalidixic acid, trimethoprim and azlocillin) to avoid the growth of fast grower microorganisms [79; 80]. This automatic system enables the detection of emitted fluorescence by the consuming of oxygen present in culture medium. A fluorescent substance is present in a silicone matrix and is sensitive to the oxygen present in culture medium in large amounts acting as a quencher. With the growth of microorganisms oxygen is consumed and fluorescence increases allowing fluorescence detection by the BACTEC equipment (http://www.bd.com/ds/productCenter/MT-Bactec.asp).

Solid media with different compositions and supplements are mostly used for the first isolation and passage of *M. avium* subspecies. Herrold egg yolk medium (HEYM), Löwenstein–Jensen medium (LJ) and Coletsos media are egg based culture media. Middlebrook 7H10 and 7H11 and Dubos are synthetic culture media and can also be used [79; 80]. Depending of the suspicious *M.*

avium subspecies being tested addiction of supplements and/or additives are required to enhance the growth of the agents.

Despite the culture of the agent, bacterial diagnostic can be performed by direct microscopic observation of samples in order to identify the presence of mycobacteria and cellular alterations in lesions. For this identification, staining procedures are required for the observation of Acid Fast bacilli and host cells visualization.

1.6.2. Immunological and Serological diagnostic

Diagnosis of *M. avium* infections can also be routinely performed by Immunological and Serological methods.

Immunological diagnostic is based on Tuberculin test by intradermal inoculation of a purified protein derivate (PPD) and observation of the skin reaction 48-72 hours after the injection. A nodule to an oedema formation at the local of the test is suggestive of a positive result indicating the presence of a delayed-type hypersensitivity (DTH). Tuberculin test is not highly specific and can show cross reactivity with other mycobacteria.

Interferon-gamma (IFN-γ) release assay is other Immunological diagnostic test based on a quantitative detection of IFN-γ released by sensitised lymphocytes during an incubation period of 18-46 hours with a specific antigen. Results obtained with this assay are difficult to interpret due to the lack of standardization procedures [79].

Serological tests most commonly used for diagnostic of *M. avium* subspecies infection are enzyme-linked immunosorbent assay (ELISA), complement fixation (CF) and agar gel immunodiffusion (AGID). All these three tests are based on the detection of specific antibodies in serum of the animal using specific antigens. Presence of specific antibodies in serum is indicative of reaction from animal's immune system to identify and neutralize pathogens from the organism, being indicative of exposure of the animal to the agent.

ELISA is performed in 96-well plates allowing the analysis of multiple samples in the same test. Interaction of antibodies and antigen is enhanced in order to allow the formation of antibody-antigen complex. Excess of antigen and antibody is washed, a peroxidase-based reagent is added and a colorimetric reaction is formed. The colour is measured by spectrophotometer and is proportional to the amount of antibody in the reaction.

CF uses 96-well round-bottom microtitration plates and consists in interaction of specific antibodies with specific antigens and the interaction of the complement reagent with antibody-antigen complex leading to the lysis of red blood cells (RBC) and release of the haemoglobin originating a red coloration of solution in well, in case of a positive result.

AGID test is usually performed in a glass slide or petri dishes. Agarose gel is poured and some holes are made in agarose, one in the centre and the remaining around the centre. Antigen is added to the centre hole and samples around the antigen. After 24-48 hours of incubation the result is observed by the formation of a precipitate if the sample is positive. These tests have a variable range of sensitivity and specificity depending of the test, antigen and animal being tested [79].

1.6.3. Molecular diagnostic

Molecular approaches based on detection of specific nucleic acids are being used for the rapid detection and identification of the agent responsible for the infection. Specific genomic markers as insertion sequences (IS) or unique genomic elements are used as targets for the identification of *M. avium* subspecies. Standard polymerase chain reaction (PCR) and real time PCR have been the most widely eligible tools in order to increase sensitivity and decrease time of direct detection from biological samples.

Polymerase chain reaction (PCR) was firstly described in 1986 by Kary Mullis [81] and is considered as one of the most important discoveries in scientific research field allowing the advance in acid nucleic-based studies all over the world. Based on a reaction mixture composed by optimized concentrations of two "in silico" designed oligonucleotide primers (a forward and a reverse); deoxyribonucleotide triphosphate (dNTP) Solution Mix; Magnesium Chloride ($MgCl_2$); DNA polymerase and DNA template, it's possible to obtain thousands of copy numbers of a DNA fragment in order to identify a specific molecular target, revealed in a stained agarose gel.

Mullis and colleagues (1986) [81] also described for the first time the nested PCR, a reaction composed by two amplification steps with two different pairs of primers. First amplification step, using external primers, originates an amplified product that is used as DNA template in a second amplification with internal primers, increasing substantially the number of copies of a specific target. This nested strategy is especially useful when the amount of DNA target is suspected to be very low leading to the increase of sensitivity of the assay and beneficial in samples with large amount of PCR inhibitors because they are diluted in the second amplification step. However, nested PCR can present problems with false-positive results due to cross-contamination [82; 83]. For this reason, good laboratorial practices and the inclusion of multiple negative controls interleaved with samples are required in order to avoid cross-contamination issues.

A PCR can also be a singleplex if the amplification is performed just for one specific target, using a specific pair of primers, or multiplex by the amplification of multi-targets in the same reaction with multiple pairs of primers, each pair specific for each fragment.

Based on a PCR reaction, Restriction Endonuclease Analysis (REA) can also be performed for specific molecular targets where distinct band patterns can be obtained, after enzymatic digestion of the amplified products, in order to differentiate between bacterial species. PCR-REA was implemented for mycobacterial species identification being the 16S rDNA [84] and *hsp65* gene [85] the most widely used targets.

Commercially available tests such as AccuProbe® Culture Identification Tests (Gen-Probe), based on RNA ribosomal analysis [26] and INNO-LIPA® Mycobacteria V2 (Innogenetics), based on the hybridization of 16S-23S rDNA spacer regions with specific probes [86], can be currently used in diagnostic laboratories to identify mycobacterial species from a positive pure culture. However, these approaches cannot differentiate between *M. avium* subspecies.

Standard PCR had evolved to a kinetic analysis and named real time PCR, firstly described by Higuchi and colleagues (1993) [87]. This analysis consisted in a continuously monitoring of the PCR with a camera connected to a computer. A quantitative information and a real time progress observation of the amplification reaction was possible by the increasing fluorescence from ethidium bromide (EtBr) intercalated in double strand DNA. The amplification results were graphically presented at the computer, with the graphic curves being progressively formed all over the reaction cycles, reducing the analysis time comparing to standard PCR where the results are revealed in an agarose gel in an “endpoint” analysis [87; 88]. In 1996, a TaqMan probe quantitative real time PCR was described for the first time by Heid and colleagues [89]. Nowadays, these two real time PCR methodologies are still the most commonly used, being the TaqMan probe approach the most specific one. TaqMan probe assay uses a specific designed dually fluorophore-labeled (with a quencher at 3’end and a fluorophore at 5’end) DNA oligonucleotide and specific primers for the amplification of a DNA fragment of interest, while the intercalation of a fluorescent dye like EtBr or SybrGreen is not specific once it has the ability of binding with any double strand DNA [88]. Since the late 1990s real time PCR technology had progressed and currently several commercially available instrumentation and methodologies are accessible from a wide variety of companies.

TaqMan probe real time PCR offers more advantages than standard PCR including the sensitivity, reproducibility and specificity of the assay, conferred by the specific probe; faster analysis time and reduced contamination of samples, due to the non-manipulation of the amplified products for revealing the results; by using a standard curve it is possible to quantify the DNA target in a sample and it is possible to normalize the assay using an internal amplification control [88; 89]. However, real time PCR is very susceptible to inhibitors of amplification reaction present in biological samples [88]. For this reason, a critical step for all molecular diagnostic tools is the quality of DNA where a highly quality is required, being the extraction of nucleic acids from a biological sample a big challenge and a crucial step for the achievement of the assay.

1.7. *Mycobacterium avium* subspecies *hominissuis* (*Mah*)

Mycobacterium avium subspecies *hominissuis* (*M. a. hominissuis*, *Mah*) designation was the most recently proposed of the four subspecies in order to distinguish between mycobacteria that cause disease in humans and pigs from those found in birds [7; 21]. *Mah* strains had also been sporadically isolated from other animals including cattle, dog, deer and horses [6]. These mycobacteria are saprophyte and opportunistic pathogens causing tuberculosis lesions in lymph nodes of digestive tract of pigs and it can be isolated mainly from animal organs and from peat used as a feeding supplement.

It is thought that pigs get infected during the intake of contaminated peat and that the mycobacteria penetrate through mucosa of mouth to gastrointestinal tract and reaches the head and mesenteric lymph nodes via lymphatic drainage. However, mycobacteria do not spread into other lymph nodes or organs [8; 90].

Mah strains are associated with opportunistic infections in mammals and humans being considered as the most important and dangerous member of *Mycobacterium avium* Complex (MAC) affecting immunocompromised individuals, mainly patients with Acquired Immunodeficiency Syndrom (AIDS), and can also cause pulmonary infections and cervical lymphadenitis usually in children [6; 8; 19]. Humans usually get infected with *Mah* by respiratory tract and the mycobacteria slowly reach other organs or intestinal epithelium causing disseminated infections [24].

The amoeba, soil and the water are considered as natural reservoirs of *Mah* with the ability of remaining as viable cells for up to 26 months being the water assumed as the main source of infection for humans.

Diagnostic of *Mah* infections can be performed using intradermal tuberculin test [91]; IFN-γ assays [92]; collection and histopathology of tissue from lesions (lymph nodes) or pulmonary extracts (from humans) in order to identify Acid-Fast (AF) bacilli and culture of the agent in solid (Middlebrook 7H10; HEYM; LJ with pyruvate; Coletsos) or liquid culture medium (Middlebrook 7H9 supplemented with ADC), at 24 and 45ºC for up to three months. Identification of isolates is routinely achieved by molecular tools based on IS*1245* and IS*901* [91; 93; 94; 95].

Mah infections can be treated with a combination of antimicrobial drugs including clarithromycin or azithromycin, ethambutol and rifabutin or rifampin [96; 97; 98] but no vaccines are available.

Some studies had reported the relationship between human, environmental and pig isolates suggesting a controversial issue of possible common infection source or transmission between humans and pigs, however this hypothesis have never been confirmed [73; 99; 100; 101; 102]. These studies were based on typing assays like RFLP, PFGE, MIRU-VNTR and MATR-VNTR where

distinct patterns can be obtained and compared between isolates in order to evaluate characterization of isolates and epidemiological traits.

In a study conducted by Mijs and colleagues (2002) [21] *Mah* isolates were characterized to belong to serotypes 8 and 9, by having an IS*1245* RFLP profile highly polymorphic and having a specific G signature at 16S-23S rDNA ITS that differ from the bird type isolates [21]. *Mah* is the *M. avium* subspecies with the most variable genome leading to a large number of different patterns obtained by typing techniques. This genetic characteristic difficults the correlation between isolates and their origin supporting that this agent is widespread in the environment with permanent contact with animals without developing disease. However, when the natural environmental conditions and animals' immunological system are changed the equilibrium between the agent and the animal is broken and the pathogen can cause disease [94].

1.8. *Mycobacterium avium* subspecies *paratuberculosis (Map)*

Mycobacterium avium subspecies *paratuberculosis (M. a. Paratuberculosis*, *Map)* is the causative agent of paratuberculosis, or Johne's disease, a chronic granulomatous enteritis infecting a wide range of animals, causing disease especially in ruminants, camelids, rabbits and hares [6; 103; 104; 105; 106]. *Map* infection has also become a highly contentious issue regarding its potential implication in the etiology of human inflammatory bowel disease (IBD) known as Crohn disease, with the agent or its DNA being detected mainly in diseased tissues [107; 108]. Humans are thought to be exposed to the agent mainly by the food chain with *Map* being found in milk, milk products and meat [109; 110; 111].

During the past decades, paratuberculosis was considered as one of the most important disease for worldwide cattle and livestock industry, due to its considerable economic impact triggered by a progressive and fatal weight loss of the animals, reduction of productivity, infertility and a diminution of milk production [109; 112; 113]. It is estimated that more than 50% of the dairy cattle in Europe and more than 68% in United States are infected with *M. paratuberculosis* [114].

The transmission of *Map* usually occurs by the fecal-oral route during the animals first months of life, being calves under 6 months of age the most susceptible due to their immature immune systems, with an infective dose about 50 to 10^3 CFU per calf [104; 115]. Clinical signs of disease usually appear after a long period of incubation, between two to five years triggered by stressful conditions such as giving birth and overcrowding. Nevertheless, some animals exposed to the agent may be

infected but successfully contain or clear infection and some animals can also remain asymptomatic but excreting the agent contribution to the spread of the disease [104; 109].

The infection has been divided into four stages according to clinical signs, faecal shedding of the agent and immunological response:

Stage I) silent infection, with no clinical signs, animals can shed minimal quantities of the agent in their faeces;

Stage II) subclinical stage, without visible signs of the disease but antibodies can be detected by an altered immune response. Cattle harbour high concentrations of *M. paratuberculosis* cells in their intestinal tissues, contaminating the environment and serving as source of infection to other animals;

Stage III) clinical stage, the animal presents visible signs of weight loss, aqueous diarrhoea and positive results on serological tests and faecal culture. The infection disseminates to several extra-intestinal sites like supra-mammary, pulmonary, hepatic and lymph nodes;

Stage IV) advanced clinical stage, evident signs are observed with an abundant excretion of *M. paratuberculosis* in faeces [116]. At this point most animals are sent to slaughter; otherwise death occurs as a consequence of dehydration and cachexia [41; 116].

In small ruminants like sheep, the clinical signs of paratuberculosis are limited to chronic weight loss from two years of age. Most animals usually succumb to disease at three to five years of life.

Presently, there are no approved drugs for the prevention or treatment of paratuberculosis. Vaccination does not totally protect animals from infection nor prevent the dissemination of *Map* in the environment and also have the disadvantage of interfering with the tuberculin test used in the control of bovine tuberculosis [79; 117].

The detection and control of paratuberculosis is difficult due to the absence of clinical signs in infected animals from a few months to several years, to the lack of reliable diagnostic methods for the initial stages of the disease and to the fastidious growth of the agent in artificial culture media. Biological samples like milk, blood, colostrum, lymph nodes and intestinal tissues are commonly used for the detection of *Map*.

Different and complementary diagnostic approaches can be used:

i) anatomopathological diagnosis, with the observation of specific pathognomonic thickening and corrugation lesions in tissue samples like ileum, intestine (visualization of plaques) and lymph nodes (normally enlarged and oedematous);

ii) immunodiagnostics by tuberculin skin tests, IFN-γ, ELISA, CF and AGID;

iii) bacteriological diagnosis, considered the "gold standard" method, involving the culture and isolation of *Map* in solid and liquid medium and automatic systems, from biological samples;

iv) molecular diagnostics, based on the detection of specific *Map* nucleic acid sequences [41; 79; 103].

The introduction of molecular diagnostic techniques has contributed to a faster, sensitive and specific detection and genotyping of *Map* in animal biological samples (Table 1.2). The multi-copy IS*900* and the single-copy *F57* elements are the most important *Map*-specific genomic targets used in molecular assays [103; 104; 118; 119; 120]. Nevertheless, improvements are still needed to enhance detection sensitivities and to clear complex samples from PCR inhibitors.

Table 1.2. Commonly used molecular techniques to identify and genotype *Mycobacterium avium* subsp. *paratuberculosis*.

	Method	Target	References
Identification	PCR/qPCR	IS*900* (15-20 copies)	[121]
		F*57* (1 copy)	[122]
		IS*Mav2* (3 copies)	[123]
		IS*Map02* (6 copies)	[124]
Genotyping	RFLP	IS*900*	[125]
	PFGE	IS*900*	[65]
	PCR-REA	IS*1311*	[126]
	PCR	SNPs	[127]
		LSPs	[128]
		MIRU-VNTR	[67]

PCR – Polymerase Chain Reaction; qPCR – quantitative Polymerase Chain Reaction or real-time PCR; RFLP - Restriction Fragment Length Polymorphism; PFGE - Pulsed Field Gel Electrophoresis; PCR-REA – Restriction Endonuclease Analysis PCR-based; SNPs - Single-nucleotide polymorphisms; LSPs - Large Sequence Polymorphisms; MIRU-VNTR - Mycobacterial Interspersed Repetitive Units- Variable Number of Tandem Repeats.

Map is one of the smallest known mycobacteria (0.5 × 1.5 µm), with a thin morphology and usually found in clumps [104]. For *in vitro* growth, *Map* requires the addition of exogenous iron-chelating mycobactin and due to its extremely low growth rate it may need more than six months of incubation to develop rough circular, 1-2 mm diameter, off-white or yellow colonies.

According to its growth characteristics, host preference and pathogenicity, *Map* was classified in two major types: Type C (type II) and Type S (type I/III) [103]. More recently, a third *Map* type had also been reported from USA and Indian isolates as Bison Type (Table 1.3).

Type C (Type II) is the most commonly isolated *Map* type, showing no host preference once it can infect a wide range of animals including livestock, wildlife and even humans. However, the initial Type C designation was associated with the bacterial isolation from cattle samples. Phenotypically, these isolates are characterized by being relatively easy to isolate depending on the initial inoculum and from the culture medium used, taking about 4-6 weeks to produce visible colonies (Table 1.3) [129].

Type S (Type I/III) is mostly isolated from sheep and goats, suggesting host preference for these two species. However, Type S strains have been isolated from other animals especially from cattle [104]. Phenotypically, Type S isolates are more difficult to isolate, grow slowly and have some special needs for culture medium. HEYM is not the best choice for this type of strains, sodium pyruvate and antimicrobials like ampicillin or vancomycin can have an inhibitory effect on growth. However, addition of egg yolk to culture medium revealed to be benefic for Type S strains. Middlebrook 7H11 and LJ both supplemented with mycobactin are the recommended culture media for this *Map* type. These strains can take about 16-52 weeks to produce visible colonies that vary from white to yellow pigmented (Table 1.3) [129].

In 2002, Stevenson and colleagues [65] proposed the designation of Type I and Type II based on RFLP analysis in order to reduce the confusion about *Map* host preference. Type I was attributed for strains with apparent host preference to sheep and goats, initially designated as Type S, and with a slower growing time. Type II was the designation for strains with faster growing time and with a broad host range including humans, known as Type C. In 2005, several Type S strains isolated from goat samples were tested by PFGE and based on the obtained profiles a Type III strain designation was proposed for an intermediate type between Type I and Type II strains [64]. Type S pigmented strains were initially considered as Type I and non-pigmented Type S strains as Type III. However, nowadays it is known that both types can comprise both pigmented and non-pigmented isolates [63; 129]. So far, there only has been reported one study with the isolation of a pigmented strain from a cow, from Scotland, after 8-10 weeks of incubation suggesting to be a Type C strain, based on host and growth characteristics [130]. Unfortunately, the strain is not available for genotyping.

Despite the high genetic homology between *Map* strain types differences in the genome had been identified allowing the characterization of strains. Genomic insertions, deletions [131], inversions and duplications in the genome were identified between Type C and Type S strains [129]. Typing tools based on insertion sequences as IS*900* and IS*1311* can be also used as well as MIRU-VNTR and Short Sequence Repeats (SSR) [132; 133]. Specific SNPs at *gyrA* and *gyrB* genes [45]; at

IS*900* [48] and at the whole genome level [129; 134] were reported to be useful for strain differentiation. Some LSPs are present and/or absent in each *Map* strain type genome: LSP^{A}4-II, MAV-14, LSP^{A}18 and GPL cluster were identified to be present in Type S strains and absent from Type C strains while LSP^{A}20 and deletion 2 are present in Type C strains and absent in Type S strains [43].

The Bison Type is a more recent designation for strains that can be isolated from different hosts from USA but showing different characteristics from Type C and Type S. Bison Type do not require sodium pyruvate in culture medium and are more difficult to culture than Type C taking about 18-20 weeks to grow (Table 1.3). By IS*1311* RFLP typing it is possible to distinguish between Type C, Type S and Bison Type based on a T-C base variation at 233 bp position [49]. Type C has one or more copies with a T or a C, Type S has all copies with a C and Bison Type has a T in all copies of *IS1311* [49; 129], however, this classification had presented some issues [134] . Indian Bison Type was a proposed designation for isolates from India that presented different genotypes from Bison USA strains [133]. Recently, whole genome sequencing studies identified that Indian Bison Type is a sub-lineage of Type C and Bison Type strains from USA were mixed with Type C strains across the phylogenetic group [129; 134].

Table 1.3. Summary of the principal characteristics of *Mycobacterium avium* subsp. *paratuberculosis* strain Types.

Type	Host	Growing time	Growing Characteristics	collonies
Type C (Type II)	Livestock, specially cattle; wildlife and humans	4-6 weeks	Relatively easy to isolate	Non-pigmented
Type S (Type I/III)	Preference by sheep and goats	16-52 weeks	More difficult to isolate, sodium piruvateand ampicillin and vancomicyn can have inhibitory effect, egg yolk can be benefic.	Pigmented and non-pigmented
Type B	Multiple hosts	18-20 weeks	Do not require sodium piruvate	Non-pigmented

1.9. Objectives and outline

Isolation, identification and characterization of *Mycobacterium avium* (*M. avium*) subspecies have presented a big challenge for researchers due to its unique characteristics of growth, requiring specific media and a long incubation period, overlapping of phenotypic and genotypic traits and distinct host infections. Novel and more reliable tools are needed in order to better characterize these mycobacteria allowing the progress of epidemiological studies. Molecular approaches based on the amplification of specific fragments of the genome have the advantage of decreasing the testing time contributing to the rapid identification and characterization of fastidious microorganisms. A critical step for molecular assays is the quality of nucleic acids especially when extracted from biological samples like faeces and milk because of the presence of large amounts of amplification inhibitors. Reliable and robust DNA extraction methodologies and amplification systems are thus in high demand.

The two general main objectives of this work were: i) to assess the epidemiological traits of *Mycobacterium avium* subspecies *hominissuis* (*M. a. hominissuis*) and unveil the potential relationship between human and animal isolates; ii) to develop and implement rapid and accurate molecular tools for the direct detection and characterization of *Mycobacterium avium* subspecies *paratuberculosis* (*M. a. paratuberculosis*) from difficult matrices such as faeces and milk. These two *M. avium* subspecies were selected for this work due to their relevance in veterinary and public health research fields. *M. a. hominissuis* is a zoonotic agent causing diverse infections in humans with an increasing implication on public health while *M. a. paratuberculosis* is a challenging pathogen due to the difficulty on its laboratorial diagnostics and worldwide economic impact caused by the progressive debilitating condition of infected animals.

This work was mainly developed at the Instituto Nacional de Investigação Agrária e Veterinária (INIAV, I.P.), Portugal, and at the Moredun Research Institute, UK, with the cooperation of The Research Unit on Applied Molecular Biosciences (UCIBIO), Departamento de Ciências da Vida, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal, of Unidade de Micobactérias, Instituto de Higiene e Medicina Tropical (IHMT/UNL), Portugal and of NEIKER, Tecnalia, Spain. The specific objectives of the work are summarized below:

- To unravel the epidemiological links between human and porcine isolates of *Mycobacterium avium* subsp. *hominissuis* in Portugal;
- To assess the presence of *Mycobacterium avium* subsp. *paratuberculosis* in ruminant samples using culture-based and molecular approaches;
- To develop rapid and accurate PCR-based assays for the direct *ante mortem* detection of *Mycobacterium avium* subsp. *paratuberculosis* in biological samples;
- To develop a novel SNP-based assay for genotyping *Mycobacterium avium* subsp. *paratuberculosis*;

- To characterize selected isolates of *Mycobacterium avium* subsp. *paratuberculosis* with whole-genome sequencing approaches.

This dissertation is composed by seven chapters which are mostly based on two published manuscripts, one manuscript recently submitted for publication, and three manuscripts in preparation.

Chapter I corresponds to a general introduction of *Mycobacterium avium* Complex describing the members of the complex, their major characteristics, diagnostic and typing tools. A more detailed description of the studied *Mycobacterium avium* subspecies (*Mycobacterium avium* subspecies *hominissuis - Mah* and *Mycobacterium avium* subspecies *paratuberculosis- Map*) and respective infections is also provided.

Chapter II describes the assessment of the relationship between human and porcine *Mycobacterium avium* subspecies *hominissuis* strains, most of them isolated during an outbreak occurred between 2004 and 2008 in Portugal, based on an MLVA assay with the evaluation of 20 MIRU-VNTR and MATR-VNTR loci. This work was published in the Veterinary Microbiology, Elsevier (doi:10.1016/j.vetmic.2014.06.027).

Chapter III describes the evaluation of the presence of *Mycobacterium avium* subspecies *paratuberculosis* in faecal samples of asymptomatic dairy cattle from the northeast region of Portugal by conventional PCR and culture of the agent. Clonal lineages were assessed by the typing of the isolates by a MIRU-VNTR assay. This work was published in the “Revista Portuguesa de Ciências Veterinárias”, SPCV.

Chapter IV describes the development of a nested real-time PCR and the optimization of DNA extraction methodologies for faecal and milk samples from cattle, goats and sheep allowing a rapid and sensitive detection of *Mycobacterium avium* subspecies *paratuberculosis* from live animals. This work was split into two manuscripts, one regarding the development of the PCR method and analysis of faecal samples and the other regarding the analysis of milk samples, both in preparation.

Chapter V describes the development of a novel typing tool for the characterization of *Mycobacterium avium* subspecies *paratuberculosis* based on SNPs analysis, by the amplification and restriction or sequencing of an amplified fragment containing a SNP in a specific position of the

genome. This work was accepted for publication at Journal of Clinical Microbiology, ESM (doi: 10.1128/JCM.01958-15).

Chapter VI describes the isolation and molecular characterization by a SNPs-based assay of a rare *Mycobacterium avium* subspecies *paratuberculosis* Type C pigmented strain isolated from a goat faecal sample from Portugal. A manuscript describing this work and a whole genome sequencing analysis is being prepared for submission.

Chapter VII evidences the most important results and the main conclusions of the work. Final considerations and future perspectives are also highlighted.

1.10. References

1. **Rue-Albrecht K, Magee DA, Killick KE, Nalpas NC, Gordon SV, MacHugh DE**. 2014. Comparative Functional Genomics and the Bovine Macrophage Response to Strains of the *Mycobacterium* Genus. Front Immunol doi: 10.3389/fimmu.2014.00536.
2. **Biet F, Boschiroli ML**. 2014. Non-tuberculous mycobacterial infections of veterinary relevance. Res Vet Sci **97**:S69–S77.
3. **Rastogi N, Legrand E, Sola C**. 2001. The mycobacteria: an introduction to nomenclature and pathogenesis. Rev Sci Tech Off Int Epiz **20**:21-54.
4. **Turenne CY, Collins DM, Alexander DC, Behr MA**. 2008. *Mycobacterium avium* subsp. *paratuberculosis* and *M. avium* subsp. *avium* are independently evolved pathogenic clones of a much broader group of *M. avium* organisms. J Bacteriol **190**:2479–2487.
5. **Turenne CY, Alexander DC**. 2010. *Mycobacterium avium* Complex. *In* Behr MA, Collins DM (eds) Paratuberculosis: Organism, Disease, Control. CAB International, Cambridge, MA, pp 60-61.
6. **Rindi L, Garzelli C**. 2014. Genetic diversity and phylogeny of *Mycobacterium avium.* Infect Genet Evol **21**:375–383.
7. **Turenne CY, Wallace R Jr, Behr MA**. 2007. *Mycobacterium avium* in the postgenomic era. Clin Microbiol Rev **20**:205–229.
8. **Shin SJ, Lee BS, Koh W, Manning EJB, Anklam KA, Sreevatsan S, Lambrecht RS, Collins MT**. 2010. Efficient Differentiation of *Mycobacterium avium* Complex Species and Subspecies by Use of Five-Target Multiplex PCR. J Clin Microbiol **48**:4057–4062.
9. **Murcia MI, Tortoli E, Menendez MC, Palenque E, Garcia MJ**. 2006. *Mycobacterium colombiense* sp. nov., a novel member of the *Mycobacterium avium* complex and description of MAC-X as a new ITS genetic variant. Int J Syst Evol Microbiol **56**:2049–2054.
10. **Tortoli E, Rindi L, Garcia MJ, Chiaradonna P, Dei R, Garzelli C, Kroppenstedt RM, Lari N, Mattei R, Mariottini A, Mazzarelli G, Murcia MI, Nanetti A, Piccoli P, Scarparo C**. 2004. Proposal to elevate the genetic variant MAC-A, included in the *Mycobacterium avium* Complex, to species rank as Mycobacterium chimaera sp. nov. Int J Syst Evol Microbiol **54**:1277–1285.
11. **Salah IB, Cayrou C, Raoult D, Drancourt M**. 2009. *Mycobacterium marseillense* sp. nov., *Mycobacterium timonense* sp. nov. and *Mycobacterium bouchedurhonense* sp. nov., members of the *Mycobacterium avium* complex. Int J Syst Evol Microbiol **59**:2803–2808.
12. **Van Ingen J, Boeree MJ, Kosters K, Wieland A, Tortoli E, Dekhuijzen PNR, Soolingen DV**. 2009. Proposal to elevate *Mycobacterium avium* complex ITS sequevar MAC-Q to *Mycobacterium vulneris* sp. nov. Int J Syst Evol Microbiol **59**:2277–2282.
13. **Bang D, Herlin T, Stegger M, Andersen AB, Torkko P, Tortoli E, Thomsen VO**. 2008. *Mycobacterium arosiense* sp. nov., a slowly growing, scotochromogenic species causing osteomyelitis in an immunocompromised child. Int J Syst Evol Microbiol **58**:2398–2402.
14. **Manning EJB, Collins MT**. 2010. History of paratuberculosis. *In* Behr MA, Collins DM (eds) Paratuberculosis: Organism, Disease, Control. CAB International, Cambridge, MA, pp 1-9.
15. **Twort F, Ingram GLY**. 1912. A method for isolating and cultivating the Mycbacterium enteritidis chronicae pseudotuberculosae bovis, Jonhe, and some experiments on the preparation of a diagnostic vaccine for pseudo-tuberculous enteritis of bovines. Proceedings of the Royal Society of London. Series B **84**:517-542.
16. **Feldman WH**. 1939. Types of tubercle bacille in lesions of Garbage-Fed swine. Am J Public Health **19**:1231-1238.
17. **Meissner G, Schroder KH, Amadio GE, Anz W, Chaparas S, Engel HWB, Jenkins PA, Kappler W, Kleeberg HH, Kubala E, Kubin SM, Lauterbach GD, Lind A, Magnusson M, Mikova Z, Pattyn GSR, Schaefer WB, Stanford JL, Tsukamura M, Wayne LG, Willers I, Wolinskyls E**. 1974. A co-operative numerical analysis of nonscoto- and nonphotochromogenic slowly growing mycobacteria. J Gen Microbiol **83**:207-235.

18. **Thorel M-F, Krichevsky M, Lévy-frebault VV**. 1990. Numerical taxonomy of mycobactin-dependent mycobacteria, emended description of *Mycobacterium avium*, and description of *Mycobacterium avium* subsp. *avium* subsp. nov., *Mycobacterium avium* subsp. *paratuberculosis* subsp. nov., and *Mycobacterium avium* subsp. *silvaticum* subsp. nov. Int J Syst Bacteriol **40**: 254-260.
19. **Inderlied CB, Kemper CA, Bermudez LEM.** 1993. The *Mycobacterium avium* Complex. Clin Microbiol Rev **6**: 266-310.
20. **Biet F, Boschiroli ML, Thorel MF, Guilloteau LA**. 2005. Zoonotic aspects of *Mycobacterium bovis* and *Mycobacterium avium-intracellulare* complex (MAC). Vet Res **36**:411–436.
21. **Mijs W, Haas P, Rossau R, Laan TVD, Rigouts L, Portaels F, van Soolingen D**. 2002. Molecular evidence to support a proposal to reserve the designation *Mycobacterium avium* subsp. *avium* for bird-type isolates and '*M. avium* subsp. *hominissuis*' for the human/porcine type of *M. avium*. Int J Syst Evol Microbiol **52**:1505–1518.
22. **Bird BB, Madison BM**. 2000. Use of fluorochrome staining for detecting acid-fast mycobacteria. In Current Laboaratory Practice Series. Department of Health and Human Services. Centers for Diseases Control and Prevention, Atlanta, pp 1-17.
23. **Laurent J-P, Hauge K, Burnside K, Cangelosi G**. 2003. Mutational Analysis of Cell Wall Biosynthesis in Mycobacterium avium. J Bacteriol **185**: 5003–5006.
24. **Ignatov D, Kondratieva E, Azhikina T**, Apt A. 2012. *Mycobacterium avium*-triggered diseases: pathogenomics. Cell Microbiol doi:10.1111/j.1462-5822.2012.01776.x.
25. **Dhama K, Mahendran M, Tiwari R, Singh SD, Kumar D, Singh S, Sawant PM**. 2011. Tuberculosis in Birds: Insights into the *Mycobacterium avium* Infections. Vet Med Int doi:10.4061/2011/712369.
26. **Saito H, Tomioka H, Sato K, Tasaka H, Dawson DJ**. 1990. Identification of Various Serovar Strains of *Mycobacterium avium* Complex by Using DNA Probes Specific for *Mycobacterium avium* and *Mycobacterium intracellulare*. J Clin Microbiol **28**: 1694-1697.
27. **Wayne LG, Good RC, Tsang A, Butler R, Dawson D, Groothuis D, Gross W, Hawkins J, Kilburn J, Kubin M, Schroder KH, Silcox VA, Smith C, Thorel M-F, Woodley C, Yakrus M A**. 1993. Serovar determination and molecular taxonomic correlation in *Mycobacterium avium*, *Mycobacterium intracellulare*, and *Mycobacterium scrofulaceum*: a cooperative study of the international working group on mycobacterial taxonomy. Int J Syst Bacteriol **43**: 482-489.
28. **Butler WR, Thibert L, Kilburn JO**. 1992. Identification of *Mycobacterium avium* Complex strains and some similar species by High-Performance Liquid Chromatography. J Clin Microbiol **30**:2698-2704.
29. **Rogall T, Flohr T, Böttger EC**. 1990. Differentiation of *Mycobacterium* species by direct sequencing of amplified DNA. J Gen Microbiol **136**:191 5-1920.
30. **Tortoli E**. 2003. Impact of genotypic studies on mycobacterial taxonomy: the new mycobacteria of the 1990s. Clin Microbiol Rev **16**: 319–354.
31. **Frothingham R, Wilson KH**. 1993. Sequence-based differentiation of strains in the *Mycobacterium avium* Complex. J Bacteriol **175**: 2818-2825.
32. **Kasai H, Ezaki T, Harayama S**. 2000. Differentiation of phylogenetically related slowly growing mycobacteria by their gyrB sequences. J Clin Microbiol **38**:301–308.
33. **Semret M, Zhai G, Mostowy S, Cleto C, Alexander D, Cangelosi G, Cousins D, Collins DM, van Soolingen D, Behr MA**. 2004. Extensive genomic polymorphism within *Mycobacterium avium*. J Bacteriol **186**: 6332–6334.
34. **Li L, Bannantine JP, Zhang Q, Amonsin A, May BJ, Alt D, Banerji N, Kanjilal S, Kapur V**. 2005. The complete genome sequence of *Mycobacterium avium* subspecies *paratuberculosis*. Proc Natl Acad Sci USA **102**:12344–12349.
35. **Wu C, Glasner J, Collins M, Naser S, Talaat AM**. 2006. Whole-Genome plasticity among *Mycobacterium avium* Subspecies: insights from comparative genomic hybridizations. J Bacteriol **188**: 711–723.

36. **Guerrero C, Bernasconi C, Burki D, Bodmer T, Telenti A**. 1995. A novel insertion element from *Mycobacterium avium*, IS1245, is a specific target for analysis of strain relatedness. J Clin Microbiol **33**: 304–307.
37. **Roiz MP, Palenque E, Guerrero C, Garcia MJ**. 1995. Use of restriction fragment length polymorphism as a genetic marker for typing *Mycobacterium avium* strains. J Clin Microbiol **33**: 1389–1391.
38. **Kunze ZM, Portaels F, Mcfadden JJ**. 1992. Biologically distinct subtypes of *Mycobacterium avium* differ in possession of Insertion Sequence IS901. J Clin Microbil **30**: 2366-2372.
39. **Moss MT, Malik ZP, Tizarde MLV, Green EP, Sanderson JD, Hermon-Taylor J**. 1992. IS902, an insertion element of the chronic-enteritis-causing *Mycobacterium avium* subsp. *silvaticum*. J Gen Microbiol **138**:139-145.
40. **Green EP, Tizard MLV, Moss MT, Thompson J, Winterbourne DJ, McFadden J, Hermon-Taylor J**. 1989. Sequence and characteristics of IS900, an insertion element identified in a human Crohn's disease isolate of *Mycobacterium paratuberculosis*. Nucleic Acids Res **17**: 9063-9073.
41. **Englund S, Bölske G, Johansson KE**. 2002. An IS900-like sequence found in a *Mycobacterium* sp. other than *Mycobacterium avium* subsp. *paratuberculosis*. FEMS Microbiol. Lett. **209**: 267 - 271.
42. **Semret M, Alexander DC, Turenne CY, Haas P de, Overduin P, van Soolingen D, Cousins D, Behr MA**. 2005. Genomic polymorphisms for *Mycobacterium avium* subsp. *paratuberculosis* diagnostics. J Clin Microbiol **43**: 3704–3712.
43. **Alexander DC, Turenne CY, Behr MA**. 2009. Insertion and deletion events that define the pathogen *Mycobacterium avium* subsp. *paratuberculosis*. J Bacteriol **191**: 1018–1025.
44. **Jagielski T, Ingen J van, Rastogi N, Dziadek J, Mazur PK, Bielecki J.** 2014. Current methods in the molecular typing of *Mycobacterium tuberculosis* and other mycobacteria. Biomed Res Int Article ID 645802.
45. **Castellanos E, Aranaz A, Romero B, de Juan L, Alvarez J, Bezos J, Rodriguez S, Stevenson K, Mateos A, Dominguez L**. 2007. Polymorphisms in *gyrA* and *gyrB* genes among *Mycobacterium avium* subspecies *paratuberculosis* Type I, II, and III isolates. J Clin Microbiol **45**:3439-3442.
46. **Griffiths TA, Rioux K, De Buck J**. 2008. Sequence polymorphisms in a surface PPE protein distinguish types I, II, and III of *Mycobacterium avium* subsp. *paratuberculosis*. J Clin Microbiol **46**:1207-1212.
47. **Turenne CY, Semret M, Cousins DV, Collins DM, Behr MA**. 2006. Sequencing of hsp65 distinguishes among subsets of the *Mycobacterium avium* complex. J Clin Microbiol **44**:433–440.
48. **Castellanos E, Aranaz A, de Juan L, Álvarez J, Rodríguez S, Romero B, Bezos J, Stevenson K, Mateos A, Domínguez L**. 2009. Single nucleotide polymorphisms in the IS*900* sequence of *Mycobacterium avium* subsp. *paratuberculosis* are strain type specific. J Clin Microbiol **47**:2260–2264.
49. **Whittington RJ, Marsh IB, Whitlock RH**. 2001. Typing of IS*1311* polymorphisms confirms that bison (*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 **15**:139–145.
50. **Motiwala AS, Strother M, Amonsin A, Byrum B, Naser SA, Stabel JR, Shulaw WP, Bannantine JP, Kapur V, Sreevatsan S.** 2003. Molecular epidemiology of *Mycobacterium avium* subsp. *paratuberculosis*: evidence for limited strain diversity, strain sharing, and identification of unique targets for diagnosis. J Clin Microbiol **41**:2015–2026.
51. **Kiehnbaum LA, Amonsin A, Wells SJ, Kapur V**. 2005. Amplified fragment length polymorphism to detect clonal diversity and distribution of *Mycobacterium avium* subspecies *paratuberculosis* in selected Minnesota dairy cattle. J Vet Diagn Invest **17**:311–315.
52. **Pfaller SL, Aronson TW, Holtzman AE, Covert TC**. 2007. Amplified fragment length polymorphism analysis of *Mycobacterium avium* complex isolates recovered from southern California. J Med Microbiol **56**:1152–1160.

53. **Ramasootaa P, Chansiripornchaia N, KaÈlleniusc G, Hoffnere SE, Svenson SB.** 2001. Comparison of *Mycobacterium avium* complex (MAC) strains from pigs and humans in Sweden by random amplified polymorphic DNA (RAPD) using standardized reagentes. Vet Microbiol **78**:251-259.
54. **Glawischnig W, Steineck T, Spergser J**. 2006. Glawischnig et al, 2006 _ Infections caused by *Mycobacterium avium* subspecies *avium*, *hominissuis,* and *paratuberculosis* in free-ranging red deer (*Cervus elaphus hippelaphus*) in Austria, 2001-2004. J Wildl Dis **42**:724–731.
55. **Christianson S, Wolfe J, Soualhine H, Sharma MK**. 2012. Comparison of repetetive-seqence-based polymerase chain reaction with random amplified polymorphic DNA analysis for rapid genotyping of nontuberculosis mycobacteria. Can J Microbiol **58**: 953-964.
56. **van Soolingen D, Bauer J, Ritacco V, Leao SC, Pavlik I, Vincent V, Rastogi N, Gori A, Bodmer T, Garzelli C, Garcia MJ**. 1998. *IS*1245 restriction fragment length polymorphism typing of *Mycobacterium avium* isolates: proposal for standardization. J Clin Microbiol **36**:3051–3054.
57. **Dvorska L, Bull TJ, Bartos M, Matlova L, Svastova P, Weston RT, Kintr J, Parmova I, van Soolingen D, Pavlik I**. 2003. A standardised restriction fragment length polymorphism (RFLP) method for typing *Mycobacterium avium* isolates links *IS*901 with virulence for birds. J Microbiol Meth **55**:11–27.
58. **Pavlik I, Horvathova A, Dvorska L, Bartl J, Svastova P, du Maine R, Rychlik I**. 1999. Standardisation of restriction fragment length polymorphism analysis for *Mycobacterium avium* subspecies *paratuberculosis*. J Microbiol Meth **38**:155–167.
59. **Sevilla I, Garrido JM, Geijo M, Juste RA.** 2007. Pulsed-field gel electrophoresis profile homogeneity of *Mycobacterium avium* subsp. *paratuberculosis* isolates from cattle and heterogeneity of those from sheep and goats. BMC Microbiol **7**:18.
60. **Álvarez J, García I G, Aranaz A, Bezos J, Romero B, de Juan L, Mateos A, Gómez-Mampaso E, Domínguez L**. 2008. Genetic diversity of *Mycobacterium avium* isolates recovered from clinical samples and from the environment: molecular characterization for diagnostic purposes. J Clin Microbiol **46**:1246–1251.
61. **Mazurek GH, Hartman S, Zhang Y, Brown BA, Hector JSR, Murphy D, Wallace RJ jr.**. 1993. Large DNA restriction fragment polymorphism in the *Mycobactenium avium-M. intracellulare* Complex: a potential epidemiologic tool. J Clin Microbiol **31**: 390-394.
62. **Pestel-Caron M, Graff G, Berthelot G, Pons J, Lemeland J**. 1999. Molecular analysis of *Mycobacterium avium* isolates by using pulsed-field gel electrophoresis and PCR. J Clin Microbiol **37**: 2450–2455.
63. **Biet F, Sevilla IA, Cochard T, Lefrançois LH, Garrido JM, Heron I, Juste RA, McLuckie J, Thibault VC, Supply P, Collins D, Behr MA, Stevenson K**. 2012. Inter and intra-subtype genotypic differences that differentiate *Mycobacterium avium* subsp. *paratuberculosis* strains. BMC Microbiol **12**:264.
64. **de Juan L, Mateos A, Domínguez L, Sharp JM, Stevenson K**. 2005. Genetic diversity of *Mycobacterium avium* subspecies *paratuberculosis* isolates from goats detected by pulsed-field gel electrophoresis. Vet Microbiol **106**:249–257.
65. **Stevenson K, Hughes VM, de Juan L, Inglis NF, Wright F, Sharp JM**. 2002. Molecular characterization of pigmented and non-pigmented isolates of *Mycobacterium avium* subspecies *paratuberculosis*. J Clin Microbiol **40**:1798–1804.
66. **Supply P, Magdalena J, Himpens S, Locht C**. 1997. Identification of novel intergenic repetitive units in a mycobacterial two-component system operon. Mol Microbiol **26**:991–1003.
67. **Bull TJ, McMinn EJ, Sidi-Boumedine K, Skull A, Durkin D, Neild P, Rhodes G, Pickup R, Hermon-Taylor J**. 2003. Detection and verification of *Mycobacterium avium* subsp. *paratuberculosis* in fresh ileocolonic mucosal biopsy specimens from individuals with and without Crohn's disease. J Clin Microbiol **41**:2915–2923.
68. **Thibault VC, Grayon M, Boschiroli ML, Hubbans C, Overduin P, Stevenson K, Gutierrez M C, Supply P, Biet F**. 2007. New variable-number tandem-repeats markers for typing *Mycobacterium avium* subsp. *paratuberculosis* and *M. avium* strains: comparison with

*IS*900 and *IS*1245 restriction fragment length polymorphism typing. J Clin Microbiol **45**:2404-2410.

69. **Pate M, Kušar D, Zŏlnir-Dovč M, Ocepek M**. 2010. MIRU–VNTR typing of *Mycobacterium avium* in animals and humans: Heterogeneity of *Mycobacterium avium* subsp. *hominissuis* versus homogeneity of *Mycobacterium avium* subsp. *avium* strains. Res Vet Sci **91**:376-81.
70. **Inagaki T, Nishimori K, Yagi T, Ichikawa K, Moriyama M, Nakagawa T, Shibayama T, Uchiya K, Nikai T, Ogawa K**. 2009. Comparison of a variable number tandem-repeat (VNTR) method for typing *Mycobacterium avium* with mycobacterial interspersed repetitive-unit-VNTR and IS1245 restriction fragment length polymorphism typing. J Clin Microbiol **47**:2156–2164.
71. **Castellanos E, Aranaz A, De Buck J**. 2010. Rapid Identification and Differentiation of *Mycobacterium avium* subspecies *paratuberculosis* Types by Use of Real-Time PCR and High-Resolution Melt Analysis of the MAP1506 Locus. J Clin Microbiol **48**:1474–1477.
72. **Fujita K, Ito Y, Hirai T, Maekawa K, Imai S, Tatsumi S, Niimi A, Iinuma Y, Ichiyama S, Mishima M**. 2012. Genetic relatedness of *Mycobacterium avium-intracellulare* complex isolates from patients with pulmonary MAC disease and their residential soils. Clin Microbiol Infect **19**:537-541.
73. **Iwamoto T, Nakajima C, Nishiuchi Y, Kato T, Yoshida S, Nakanishi N, Tamaru A, Tamura Y, Suzuki Y, Nasu M**. 2012. Genetic diversity of *Mycobacterium avium* subsp. *hominissuis* strains isolated from humans, pigs, and human living environment. Infect Genet Evol **12**:846-852.
74. **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.
75. **Radomski N, Thibault VC, Karoui C, de Cruz K, Cochard T, Gutiérrez C, Supply P, Biet F, Boschiroli ML**. 2010. Determination of genotypic diversity of *Mycobacterium avium* subspecies from human and animal origins by mycobacterial interspersed repetitive-unit-variable-number tandem-repeat and *IS*1311 restriction fragment length polymorphism typing methods. J Clin Microbiol **48**:1026-1034.
76. **Romano MI, Amadio A, Bigi F, Klepp L, Etchechoury I, Llana MN, Morsella C, Paolicchi F, Pavlik I, Bartos M, Leão SC, Cataldi A**. 2005. Further analysis of VNTR and MIRU in the genome of *Mycobacterium avium* complex, and application to molecular epidemiology of isolates from South America. Vet Microbiol **110**:221–237.
77. **Stevenson K, Àlvarez J, Bakker D, Biet F, de Juan L, Denham S, Dimareli Z, Dohmann K, Gerlach G-F, Heron I, Kopecna M, May L, Pavlik I, Sharp JM, Thibault VC, Willemsen P, Zadoks R, Greig A**. 2009. Occurrence of *Mycobacterium avium* subspecies *paratuberculosis* across host species and European countries with evidence for transmission between wildlife and domestic ruminants. BMC Microbiol **9**:212.
78. **Tatano Y, Sano C, Yasumoto K, Shimizu T, Sato K, Nishimori K, Matsumoto T, Yano S, Takeyama H, Tomioka H**. 2012. Correlation between variable-number tandem-repeat-based genotypes and drug susceptibility in *Mycobacterium avium* isolates. Eur J Clin Microbiol Infect Dis **31**:445-454.
79. **OIE**. 2014. Paratuberculosis (Johne's Disease). In OIE Terrestrial Manual (World Organisation for Animal Health), Chapter 2.1.11.
80. **Whittington R**. 2010. Cultivation of *Mycobacterium avium* subsp. *paratuberculosis*. *In* Behr MA, Collins DM (eds) Paratuberculosis: Organism, Disease, Control. CAB International, Cambridge, MA, pp 244-266.
81. **Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H**. 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harbor Symp Quant Biol **51**:263-273.
82. **Neonakis IK, Gitti Z, Krambovitis E, Spandidos DA**. 2008. Molecular diagnostic tools in mycobacteriology. J Microbiol Meth **75**:1-11.

83. **Bölske G, Herthnek**. Diagnosis of Paratuberculosis by PCR. *In* Behr MA, Collins DM (eds) Paratuberculosis: Organism, Disease, Control. CAB International, Cambridge, MA, pp 267-283
84. **Vaneechoutte M, de Beenhouwer H, Claeys G, Verschraegen G, de Rouck A, Paepe N, Elaichouni A, Portaels F**. 1993. Identification of *Mycobacterium Species* by using amplified ribosomal DNA restriction analysis. J Clin Microbiol **31**:2061-2065.
85. **Telenti A, Marchesi F, Balz M, Bally F, Botrger EC, Bodmer T**. 1993. Rapid identification of mycobacteria to the species level by polymerase chain reaction and restriction enzyme analysis. J Clin Microbiol **31**:175-178.
86. **Tortoli E, Nanetti A, Piersimoni C, Cichero P, Farina C, Mucignat G, Scarparo C, Bartolini L, Valentini R, Nista D, Gesu G, Tosi CP, Crovatto M, Brusarosco G.** 2001. Performance assessment of new multiplex probe assay for identification of mycobacteria. J Clin Microbiol **39**: 1079–1084
87. **Higuchi R, Fockler C, Dollinger G, Watson R**. 1993. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. Biotechnol **11**:1026-1030.
88. **Valasek MA, Repa JJ**. 2005. The power of real-time PCR. Adv Physiol Educ **29**:151–159.
89. **Heid CA, Stevens J, Livak KJ, Williams PM**. 1996. Real time quantitative PCR. Genome Res **6**:986-994.
90. **Matlova L, Dvorska L, Ayele WY, Bartos M, Amemori T, Pavlik I**. 2005. Distribution of *Mycobacterium avium* Complex isolates in tissue samples of pigs fed peat naturally contaminated with mycobacteria as a supplement. J Clin Microbiol **43**:1261–1268.
91. **Agdestein A, Johansen TB, Polaček V, Lium B, Holstad G, Vidanović D, Aleksić-Kovačević S, Jørgensen A, Žultauskas J, Nilsen SF Djønne B**. 2011. Investigation of an outbreak of mycobacteriosis in pigs. BMC Vet Res **7**:63.
92. **Faldyna M, Göpfert E, Kudlackova H, Stepanova H, Kaevska M, Slana I, Pavlik I**. 2012. Usability of a gamma interferon release assay in the diagnosis of naturally infected pigs with *Mycobacterium avium* subspecies *hominissuis*. J Vet Diagn Invest **24**:376–378.
93. **Álvarez J, Castellanos E, Romero B, Aranaz A, Bezos J, Rodríguez S, Mateos A, Domínguez L, de Juan L**. 2011. Epidemiological investigation of a *Mycobacterium avium* subsp. *hominissuis* outbreak in swine. Epidemiol Infect **139**:143–148.
94. **Domingos D, Amado A, Botelho A**. 2009. IS1245 RFLP analysis of strains of *Mycobacterium avium* subspecies *hominissuis* isolated from pigs with tuberculosis lymphadenitis in Portugal. Vet Rec **164**:116-120.
95. **Kaevska M, Slana I, Kralik P, Reischl U, Orosova J, Holcikova A, Pavlik I.** 2011. "*Mycobacterium avium* subsp. *hominissuis*" in neck lymph nodes of children and their environment examined by culture and triplex quantitative real-time PCR. J Clin Microbiol **49**:167–172.
96. **Griffith DE, Aksamit T, Brown-Elliott BA, Catanzaro A, Daley C, Gordin F, Holland SM, Horsburgh R, Huitt G, Iademarco MF, Iseman M, Olivier K, Ruoss S, von Reyn CF, Wallace RJ Jr., Winthrop K**. 2007. An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. Am J Respir Crit Care Med **175**:367–416.
97. **Shafran SD, Singer J, Zarowny DP, Philips P, Salit MD, Walmsley SL, Fong IW, Gill MJ, Rachlis AR, Lalonde RG, Fanning MM, Tsoukas CM**. 1996. A comparison of two regimens for the treatment of *Mycobacterium avium* Complex bacteremia in AIDS: rifabutin, ethambutol, and clarithromycin versus rifampin, ethambutol, clofazimine, and ciprofloxacin. Canadian HIV Trials Network Protocol 010 Study Group. N Engl J Med **335**:377-383.
98. **Sim YS, Park HY, Jeon K, Suh GY, Kwon OJ, Koh W**. 2010. Standardized combination antibiotic treatment of *Mycobacterium avium* Complex lung disease. Yonsei Med J **51**:888-894.
99. **Johansen TB, Olsen I, Jensen MR, Dahle UR, Holstad G, Djonne B**. 2007. New probes used for *IS*1245 and *IS*1311 restriction fragment length polymorphism of *Mycobacterium avium* subsp. *avium* and *Mycobacterium avium* subsp. *hominissuis* isolates of human and animal origin in Norway. BMC Microbiol **7**:14.

100. **Möbius P, Lentzsch P, Moser I, Naumann L, Martin G, Kohler H**. 2006. Comparative macrorestriction and RFLP analysis of *Mycobacterium avium* subsp. *avium* and *Mycobacterium avium* subsp. *hominissuis* isolates from man, pig, and cattle. Vet Microbiol **117**:284–291.
101. **Pate M, Žolnir-Dovč M, Krt B, Ocepek M**. 2008. *IS*1245 RFLP-based genotyping study of *Mycobacterium avium* subsp. *hominissuis* isolates from pigs and humans. Comp Immun Microbiol Infect Dis **31**:537–550.
102. **Tirkkonen T, Pakarinen J, Rintala E, Ali-Vehmas T, Marttila H, Peltoniemi OAT, Mäkinen J**. 2010. Comparison of variable-number tandem-repeat markers typing and *IS*1245 restriction fragment length polymorphism fingerprinting of *Mycobacterium avium* subsp. *hominissuis* from human and porcine origins. Acta Vet Scand **52**:21.
103. **Timms VJ, Gehringer MM, Mitchell HM, Daskalopoulos G, Neilan BA**. 2011. How accurately can we detect *Mycobacterium avium* subsp. *paratuberculosis* infection? J Microbiol Meth **85**:1-8.
104. **Salem M, Heydel C, El-Sayed A, Ahmed SA, Zschöck M, Baljer G**. 2012. *Mycobacterium avium* subspecies *paratuberculosis*: an insidious problem for the ruminant industry. Trop Anim Health Prod **45**:351-366.
105. **Carta T, Álvarez J, Pérez de la Lastra JM, Gortázar C**. 2013. Wildlife and paratuberculosis: A review. Res Vet Sci **94**:191-197.
106. **Salgado M, Monti G, Sevilla I, Manning E**. 2014 Association between cattle herd *Mycobacterium avium* subsp. *paratuberculosis* (MAP) infection and infection of a hare population. Trop Anim Health Prod **46**:1313-1316.
107. **Grant I R**. 2005. Zoonotic potential of *Mycobacterium avium* ssp. *paratuberculosis*: the current position. J Appl Microbiol **98**:1282–1293.
108. **Juste RA, Perez V**. 2011. Control of Paratuberculosis in Sheep and Goats. Vet Clin Food Anim **27**:127–138.
109. **Atreya R, Bülte M, Gerlach GF, Goethe R, Hornef MH, Köhler H, Meens J, Möbius P, Roeb E, Weiss S**, on behalf of the ZooMAP Consortium. 2014. Facts, myths and hypotheses on zoonotic nature of *Mycobacterium avium* subspecies *paratuberculosis*. Int J Med Microbiol **304**:858-867.
110. **Donaghy JA, Johnston J, Rowe MT.** 2010. Detection of *Mycobacterium avium* ssp. *paratuberculosis* in cheese, milk powder and milk using *IS*900 and f57- based qPCR assays. J Appl Microbiol **110**:479-489.
111. **Slana I, Paolicchi F, Janstova B, Navratilova P, Pavlik I**. 2008. Detection methods for *Mycobacterium avium* subsp. *paratuberculosis* in milk and milk products: a review. Vet Med **53**:283-306.
112. **Hanifian S, Khani S, Barzegari A, Shayegh J**. 2013. Quantitative real-time PCR and culture examination of *Mycobacterium avium* subsp. *paratuberculosis* at farm level. Vet Microbiol **162**:160-165.
113. **Sting R, Hrubenja M, Mand J, Seemann G, Salditt A, Waibel S**. 2014. Detection of *Mycobacterium avium* subsp. *paratuberculosis* in faeces using different procedures of pre-treatment for real-time PCR in comparison to culture. Vet J **199**:138-142.
114. **Manning EJB, Collins MT.** 2010. Epidemiology of paratuberculosis. *In* Behr MA, Collins DM (eds) Paratuberculosis: Organism, Disease, Control. CAB International, Cambridge, MA, pp 22-28.
115. **Laurin EL, Chaffer M, McClure JT, McKenna LB, Keefe GP**. 2014. The association of detection method, season, and lactation stage on identification of fecal shedding in *Mycobacterium avium* ssp. *paratuberculosis* infection dairy cows. J Dairy Sci **98**:211-220.
116. **Fecteau M-E, Whitlock RH**. 2010. Paratuberculosis in cattle. *In* Behr MA, Collins DM (eds) Paratuberculosis: Organism, Disease, Control. CAB International, Cambridge, MA, pp 144-156.
117. **Lisle GW de**. 2010. Ruminant aspects of Paratuberculosis vaccination. *In* Behr MA, Collins DM (eds) Paratuberculosis: Organism, Disease, Control. CAB International, Cambridge, MA, pp 344-352.

118. **Vansnick E, De Rijk P, Vercammen F, Geysen D, Rigouts L, Portaels F**. 2004. Newly developed primers for the detection of *Mycobacterium avium* subspecies *paratuberculosis*. Vet Microbiol **100**:197-204.
119. **Castellanos E, de Juan L, Domínguez L, Aranaz A**. 2012. Progress in molecular typing of *Mycobacterium avium* subspecies *paratuberculosis*. Res Vet Sci **92**:169-179.
120. **Gillardoni LR, Paolicchi FA, Mundo SL**. 2012. Bovine paratuberculosis: a review of the advantages and disadvantages of different diagnostic tests. Rev Argent Microbiol **44**:201-215.
121. **Vary PH, Anderson PR, Green E, Hermon-Taylor J, McFadden JJ.** 1990. Use of highly specific DNA probes and the polymerase chain reaction to detect *Mycobacterium paratuberculosis* in Johne's disease. J Clin Microbiol **28**:933-7.
122. **Coetsier C, Vannuffel P, Blondeel N, Denef JF, Cocito C, Gala JL.** 2000. Duplex PCR for differential identification of *Mycobacterium bovis*, *M. avium*, and *M. avium* subsp. *paratuberculosis* in formalin-fixed paraffin-embedded tissue from cattle. J Clin Microbiol **38**:3048–3054.
123. **Strommenger B, Stevenson K, Gerlach GF**. 2001. Isolation and diagnostic potential of *IS*Mav2, a novel insertion sequence-like element from *Mycobacterium avium* subspecies *paratuberculosis*. FEMS Microbiology Letters **196**:31-37.
124. **Stabel JR, Bannantine JP**. 2005. Development of a nested PCR method targeting a unique multicopy element, *IS*Map02, for detection of *Mycobacterium avium* subsp. *paratuberculosis* in faecal samples. J Clin Microbiol **43**:4744-4750.
125. **Collins DM, Gabric DM, Lisle GW de**. 1990. Identification of two groups of *Mycobacterium paratuberculosis* strains by restriction endonuclease analysis and DNA hybridization. J Clin Microbiol **28**:1591–1596.
126. **Marsh I, Whittington R, Cousins D.** 1999. PCR-restriction endonuclease analysis for identification and strain typing of *Mycobacterium avium* subsp. *paratuberculosis* and *M. avium* subsp. *avium* based on polymorphisms in *IS*1311. Mol Cell Probes **13**:115-126.
127. **Marsh IB, Whittington RJ**. 2007. Genomic diversity in *Mycobacterium avium*: single nucleotide polymorphisms between the S and C strains of *M. avium* subsp. *paratuberculosis* and with *M. a. avium*. Mol Cell Probes **21**:66–75.
128. **Semret M, Turenne CY, Haas P de, Collins DM, Behr MA**. 2006. Differentiating host-associated variants of *Mycobacterium avium* by PCR for detection of large sequence polymorphisms. J Clin Microbiol **44**:881–887.
129. **Stevenson K**. 2015. Genetic diversity of *Mycobacterium avium* subspecies *paratuberculosis* and the influence of strain type on infection and pathogenesis: a review. Vet Res **46**:64.
130. **Watt JAA**. 1954. Johne's disease in a bovine associated with the pigmented strain of *Mycobacterium johnei*. Vet Rec **66**:387.
131. **Marsh IB, Bannantine JP, Paustian ML, Tizard ML, Kapur V, Whittington RJ**. 2006. Genomic comparison of *Mycobacterium avium* subspecies *paratuberculosis* sheep and cattle strains by microarray hybridization. J Bacteriol **188**:2290–2293.
132. **Thibault VC, Grayon M, Boschiroli ML, Willer E, Allix 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.
133. **Sevilla I, Singh SV, Garrido JM, Aduriz G, Rodríguez S, Geijo MV, Whittington RJ, Saunders V, Whitlock RH, Juste RA**. 2005. Molecular typing of *Mycobacterium avium* subsp. *paratuberculosis* strains from different hosts and regions. Sci Tech Rev **24**:1061–1066.
134. **Bryant JM, Thibault VC, Smith DGE, McLuckie J, Heron I, Sevilla IA, Biet F, Harris SR, Maskell DJ, Bentley SD, Parkhill J, Stevenson K**. 2015. Phylogenomic exploration of the relationships between strains of *Mycobacterium avium* subspecies *paratuberculosis.* BMC Genomics In press.

Chapter II

Relatedness of *Mycobacterium avium* subspecies *hominissuis* clinical isolates of human and porcine origins assessed by Multiple-Locus Variable number tandem repeat Analysis (MLVA)

Célia Leão, Ana Canto, Diana Machado, Ilda Santos-Sanches, Isabel Couto, Miguel Viveiros, João Inácio, Ana Botelho

Manuscript (doi:10.1016/j.vetmic.2014.06.027) published with minor modifications in 2014 in the Veterinary Microbiology Journal vol. 173 pages 92–100

"Com autorização do editor e sujeitos aos direitos de cópia impostos pelo mesmo "

Relatedness of *Mycobacterium avium* subspecies *hominissuis* clinical isolates of human and porcine origins assessed by Multiple-Locus Variable number tandem repeat Analysis (MLVA)

2.1. Abstract

Mycobacterium avium subsp. *hominissuis* (*Mah*) is an important opportunistic pathogen, infecting humans and animals, notably pigs. Several methods have been used to characterize *Mah* strains. Restriction fragment length polymorphisms (RFLP) and pulsed-field gel electrophoresis (PFGE) typing techniques have been used as standard methods but are technically demanding. In contrast, the analysis of variable-number tandem-repeats (VNTR) *loci* is a simpler, affordable and highly reliable PCR-based technique, allowing a numerical and reproductive digitalization of typing data. In this study, the analysis of *Mycobacterium avium* tandem repeats (MATRs) *loci* was adapted to evaluate the genetic diversity of epidemiological unrelated *Mah* clinical strains of human (n = 28) and porcine (n = 69) origins, collected from diverse geographical regions across mainland Portugal. These *Mah* isolates were found to be genetically diverse and genotypes are randomly distributed across the country. Some of the human strains shared identical VNTR profiles with porcine isolates. Our study shows that the VNTR genotyping using selected MATR *loci* is a useful analysis technique for assessing the genetic diversity of *Mah* isolates from Portugal. This typing method could be successfully applied in other countries towards the implementation of a worldwide open-access database of MATR-VNTR profiles of *Mah* isolates, allowing a better assessment of the global epidemiology traits of this important pathogenic species.

Keywords

Nontuberculous mycobacteria; *Mycobacterium avium* complex (MAC); *Mycobacterium avium* subsp. *hominissuis*; *Mycobacterium avium* tandem repeats; MATR-VNTR typing; opportunistic pathogens

2.2. Introduction

Mycobacterium avium subspecies *hominissuis* (*Mah*) is a member of the *Mycobacterium avium* complex (MAC) that comprises a heterogeneous group of nontuberculous mycobacteria (NTM) [1]. Members of MAC are ubiquitous in the environment and are opportunistic pathogens associated with human and animal diseases. *Mah* is the most common NTM infecting AIDS patients, especially in developed countries, and is often associated with pulmonary disease, localized soft-tissue infections, lymphadenitis and cystic fibrosis in immunocompetent individuals [2; 3; 4]. *Mah* has also the ability to cause generalized tuberculosis in poultry and wild birds and localized granulomatous lesions in the lymph nodes of the digestive tract of mammals, especially pigs [5]. The detection of these lesions is frequent during meat inspections in abattoirs, reducing the value of the carcasses and leading to potential market restrictions and relevant economic impacts.The routes of *Mah* infections have not yet been clearly identified and human-to-human transmission has not been clearly demonstrated. Most studies report that clinical *Mah* isolates are genetically diverse and that both humans and animals may acquire the infection from common environmental sources [3; 5; 6], such as water distribution systems [7]. However, in some occasions, a potential direct transmission of *Mah* between pigs and humans cannot be excluded, since the genetic profiles of some strains isolated from both hosts have revealed great similarity [6; 8; 9].

Several methods have been used to characterize *Mah* strains and namely to clarify the sources of infection and routes of transmission: restriction fragment length polymorphisms (RFLP) analysis using IS*1245*-targeted probes [5; 6; 10], pulsed-field gel electrophoresis (PFGE) [6; 7; 11] and mycobacterial interspersed repetitive units - variable-number tandem-repeats (MIRU-VNTR) *loci* analysis [12; 13; 14] approach that is generally denominated Multiple-Locus Variable number tandem repeat Analysis (MLVA). RFLP and PFGE typing techniques have been used as standard methods but are technically demanding, hard to standardize and time consuming, requiring large amounts of highly purified DNA [13]. The existence of *Mah* strains harboring low numbers of IS*1245* copies can also reduce the discriminatory power of RFLP analysis [11]. In contrast, MLVA is an affordable and highly reliable PCR-based technique, allowing a numerical and reproducible digitalization of typing data and enabling a more effective comparison of inter-laboratory results. Individual VNTR *loci* appear to be highly stable in the genome of MAC members even when testing the isolates after several generations, by repeated passages *in vitro* or *in vivo* [15; 16; 17]. Furthermore, the analysis of specific *Mycobacterium avium* tandem repeats (MATR) *loci* revealed an excellent discriminatory power for typing MAC isolates [15] and is being increasingly and successfully used to characterize human, animal and environmental isolates [15; 18; 19; 20; 21; 22; 23; 24; 25].

In this study, we used a MLVA method based on a combination of MIRU and MATR *loci*, to elucidate the genetic diversity of *Mah* clinical strains of human and porcine origins from diverse Portuguese geographical regions, and clarify the epidemiological traits of the strains causing infections in humans and animals.

2.3. Material and Methods

2.3.1. Reference strain

The reference strain *Mycobacterium avium* subsp. *avium* (ATCC 25291^{T}) was used as a PCR amplification control.

2.3.2. Clinical isolates

A total of 97 *Mah* isolates were characterized. Twenty eight strains were isolated from human samples (blood, sputum, urine, bronchial secretions, biopsy and bronchoalveolar lavage fluids) collected in four hospitals from the Lisbon Health Region between 2005 and 2011. Isolation was carried out in MGIT tubes (modified Middlebrook 7H9 broth) for the BACTEC MGIT 960 system (Becton-Dickinson Diagnostic Instrument Systems, Towson, Md., USA), and identification of *M. avium* Complex was carried out using Accuprobe hybridization probes (Gen-Probe Inc., San Diego, California, USA) or Genotype CM (Hain Lifescience) based on methods routinely employed in the Mycobacteriology Laboratories of Instituto de Higiene e Medicina Tropical (IHMT/UNL Lisbon, Portugal) [26]. These strains are maintained at IHMT, UNL.

Sixty seven strains from porcine lymph nodes and two strains from wild boar lung and lymph nodes, previously isolated between 2004 and 2006, during a tuberculous lymphadenitis outbreak that occurred in mainland Portugal [5] were analyzed. The *Mah* isolates were cultured at 37 ºC in Middlebrook 7H9 broth supplemented with ADC (Albumin, Dextrose and Catalase). These strains are maintained at Instituto Nacional de Investigação Agrária e Veterinária (INIAV, IP).

2.3.3. DNA extraction

The DNAs of the 28 isolates from human samples were extracted using QIAamp DNA mini kit (QIAGEN) as described by the manufacturer. DNA extraction of the 69 animal isolates was performed by a simple boiling method. Briefly, 500µL of the culture were centrifuged at 2000 × g for 5 minutes at room temperature, the pellets were suspended in 100 µL of Tris EDTA (TE) buffer pH 8

(10 mM Tris.Cl, 1 mM EDTA) and incubated for 45 minutes at 95ºC. After the bacterial inactivation, the culture was centrifuged at 2000 × g for 1 minute at room temperature, and the supernatant containing the DNA was transferred to a new tube and used as template for PCR reactions.

2.3.4. Confirmation of *Mah* isolates

The correct assignment of the isolates to *Mah* was performed as previously described [5], using a duplex PCR assay based on the specific sequences IS*901* (1108 bp amplicon, [27]) and IS*1245* (427 bp amplicon, [28]), and a single PCR assay based on the IS*901* flanking region FR300 (300 bp amplicon [27]). Some of the animal isolates used in this study have also previously been characterized by IS*1245*/RFLP by Domingos and colleagues (2009) [5].

2.3.5. Selection of VNTR *loci*

An extended panel of 20 VNTR *loci* (MATR-1, MATR-2, MATR-3, MATR-4, MATR-5, MATR-6, MATR-7, MATR-8, MATR-9, MATR-11, MATR-12, MATR-13, MATR-14, MATR-15, MATR-16, MIRU-3, MIRU-7, MIRU-25, MIRU-32 and MIRU-47), including the 15 *loci* referred by Inagaki and colleagues (2009) [15] and five MIRU-VNTR *loci* according to Thibault and colleagues (2007) [12], were tested with 21 *Mah* strains (Table 2.1), to evaluate the discriminatory capability (h) of each *loci* (Table 2.2). Based on an allelic diversity index $h \geq 0.45$, a subset of six *loci* (MATR-3, MATR-6, MATR7, MATR8, MATR11 and MATR15) was selected to type all the 28 human strains and the remaining 48 animal strains.

The amplification of MATR and MIRU *loci* was performed as previously described by Inagaki et al. (2009) [15] and Thibault et al. (2007) [12], respectively, with minor modifications. PCR reactions were carried out in 25 µl containing 200 µM of each deoxynucleotide triphosphate (Promega), 2.0 mM of $MgCl_2$ (Promega), 0.5 µM of each primer (oligonucleotide sequences described by Inagaki et al., 2009 [15] and Thibault et al., 2007 [12]), 1 U of GoTaq® DNA polymerase, 1× of the respective buffer (Promega), and 1 µl of extracted DNA solution (with the exception of the amplification of MATR-11 *locus* where 5 µL of template DNA were used). Amplification was performed in a MJmini™ Thermocycler (BioRad) with an initial step at 95 ºC for 10 minutes, followed by 38 cycles at 98 ºC for 10 seconds, 68 ºC for 30 seconds and 72 ºC for 1 minute, ending with a step at 72 ºC for 7 minutes, with the exception of the amplification of MATR-11 *locus* for which an annealing temperature of 64 ºC was used. The amplified products were electrophoretically analyzed in a 2% (w/v) agarose gel in 1× Tris-Borate-EDTA (TBE) buffer stained with ethidium bromide. Gel electrophoresis images were acquired with an UV Transilluminator UVP M-20. Each agarose gel contained the DNA ladder IV (Bioline) and the *Mycobacterium avium* subsp. *avium* (ATCC 25291^T)

control amplicons, which were used as size references. Size of the amplified DNA bands was assigned using the software GELCOMPAR II (Applied Maths). Copy number of the repeated consensus for each VNTR *locus* was attributed based on the respective amplicon size (Table 2.3), *in silico* analysis and PCR amplified products of the reference strain *Mycobacterium avium* subsp. *avium* (ATCC 25291^T) (Table 2.3) and according to Inagaki et al. (2009) [15] and Thibault et al. (2007) [12].

Table 2.1. MATR/MIRU allelic profile for a subset of 21 Portuguese *Mycobacterium avium* subsp. *hominissuis* isolates of porcine origins[(1)]

	MATR-1	**MATR-2**	**MATR-3**	**MATR-4**	**MATR-5**	**MATR-6**	**MATR-7**	**MATR-8**	**MATR-9**	**MATR-11**	**MATR-12**	**MATR-13**	**MATR-14**	**MATR-15**	**MATR-16**	**MIRU-3**	**MIRU-7**	**MIRU-25**	**MIRU-47**	**MIRU-32**
P 8222/5/2005	1	0	6	1	1	1	5	2	2	2	3	2	4	2	3	1	1	3	2	na
P 4409/0/2005	1	2	6	0	1	1	3	2	2	4	3	2	2	2	3	1	1	3	2	na
P 10294/0/2005	2	0	6	1	1	1	5	3	2	2	3	2	2	2	3	1	1	2	2	na
WB 10178/2/2006	2	0	6	1	1	1	5	3	2	2	na	2	4	2	3	1	1	2	2	na
P 2446/2/2005	2	0	6	1	1	1	5	3	2	4	3	2	4	2	3	1	1	2	2	na
P 9395/3/2005	2	2	2	1	1	3	3	4	1	5	3	2	4	4	3	1	1	2	2	na
P 9078/0/2005	2	2	2	1	2	1	3	4	1	6	3	2	4	4	3	1	1	2	2	na
P 9276/0/2005	2	2	2	1	2	1	3	4	2	5	3	2	4	4	3	1	1	2	2	na
P 4212/0/2005	2	2	2	1	2	1	3	4	2	6	3	2	4	4	3	1	1	2	2	na
WB 10178/1/2006	2	2	2	1	2	3	3	4	2	2	3	2	4	3	2	1	1	2	2	na
P 9395/1/2005	2	2	2	1	2	3	3	4	2	6	na	2	4	3	2	1	1	2	2	na
P 9271/4/2005	2	2	2	1	2	3	4	1	1	5	3	2	2	4	2	1	1	4	3	na
P 9080/4/2005	2	2	5	0	2	1	3	2	2	6	3	2	4	4	3	1	1	2	2	na
P 9066/0/2005	2	2	5	0	2	1	3	4	2	6	3	2	4	4	3	1	1	2	2	na
P 1314/0/2005	2	2	5	0	2	3	3	4	2	6	3	2	4	4	3	1	1	2	2	na
P4205/0/2005	2	2	5	1	2	1	3	4	2	2	3	2	4	4	3	1	1	2	2	na
P 4413/0/2005	2	2	5	2	2	3	5	4	2	3	na	2	4	2	3	1	1	3	2	na
P 9080/1/2005	2	2	6	0	2	1	3	4	2	2	3	2	4	4	3	1	1	2	2	na
P 10801/9/2006	2	2	6	0	2	3	3	1	2	4	1	2	4	4	2	1	1	2	3	na

	MATR-1	MATR-2	MATR-3	MATR-4	MATR-5	MATR-6	MATR-7	MATR-8	MATR-9	MATR-11	MATR-12	MATR-13	MATR-14	MATR-15	MATR-16	MIRU-3	MIRU-7	MIRU-25	MIRU-47	MIRU-32
P 8222/6/2005	2	2	6	1	1	1	5	2	2	4	3	2	2	2	3	1	1	3	2	na
P 9273/2/2005	2	2	6	2	2	3	3	4	2	2	3	1	4	4	3	1	1	2	2	na

[1]According to Inagaki and colleagues (2009) [15] (MATR) and Thibault and colleagues (2007) [12] (MIRU); na – no amplification.

Table 2.2. MATR/MIRU allelic distribution and diversity for a subset of 21 Portuguese *Mycobacterium avium* subsp. *hominissuis* isolates of porcine origins.

Locus[1,2]	No. of isolates with a specific number of tandem repeats (allele)							Allelic diversity (h)
	0	1	2	3	4	5	6	
MATR-1	0	2	19	0	0	0	0	0.13
MATR-2 (≈MIRU-292)	4	0	17	0	0	0	0	0.27
MATR-3 (≈MIRU-X3)	0	0	7	0	0	5	9	0.63
MATR-4	6	15	0	0	0	0	0	0.38
MATR-5	0	6	15	0	0	0	0	0.38
MATR-6	0	13	0	8	0	0	0	0.45
MATR-7	0	0	0	14	1	6	0	0.45
MATR-8	0	2	4	3	12	0	0	0.59
MATR-9 (≈MIRU-10)	0	3	18	0	0	0	0	0.21
MATR-11	0	0	7	1	4	5	6	0.70
MATR-12[3]	0	1	17	0	0	0	0	0.05
MATR-13	0	1	20	0	0	0	0	0.05
MATR-14	0	0	4	0	17	0	0	0.27
MATR-15	0	0	7	2	12	0	0	0.53
MATR-16	0	0	4	17	0	0	0	0.27
MIRU-3	0	21	0	0	0	0	0	0.00
MIRU-7	0	21	0	0	0	0	0	0.00
MIRU-25	0	0	17	4	0	0	0	0.27
MIRU-47	0	0	19	2	0	0	0	0.13

[1]According to Inagaki and colleagues (2009) [15] (MATR) and Thibault and colleagues (2007) [12] (MIRU); [2]An additional *locus* MIRU-32 was tested but amplification failed for most isolates; [3]Amplification has failed for three of the 21 isolates.

Table 2.3. VNTR *loci* analyzed in this study, respective amplicon sizes and number of tandem repeats.

	Number of tandem repeats and respective estimated amplicon size (bp)							Positive control (ATCC 25291^T)	
Locus	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**Amplicon (bp)**	**No of tandem repeats[1]**
MATR-1	229	282	334	387	440	493	546	282	1
MATR-2	200	250	300	353	406	459	512	300	2
MATR-3	195	248	301	378	400	460	506	378	3
MATR-4	168	221	274	327	380	433	486	274	2
MATR-5	133	191	249	307	365	423	481	307	3
MATR-6	223	269	326	384	418	475	532	269	1
MATR-7	220	277	338	391	448	505	562	277	1
MATR-8	110	160	220	280	334	391	448	220	2
MATR-9	325	380	435	490	545	600	655	435	2
MATR-11	282	337	392	437	500	559	612	392	2
MATR-12	368	425	482	542	600	657	714	424	1
MATR-13	233	290	347	403	459	515	571	290	1
MATR-14	215	273	330	384	447	505	536	273	1
MATR-15	194	251	308	365	422	479	536	308	2
MATR-16	263	322	381	418	477	536	595	418	3
MIRU-3	154	181	208	235	262	289	316	208	2
MIRU-7	169	191	203	225	247	269	291	191	1
MIRU-25	176	234	292	350	408	466	524	234	1
MIRU-47	116	151	186	217	252	287	322	219	3

[1]Based in NCBI-GenBank genome sequence with accession number ACFI01000000

2.3.6. Data analysis

The VNTR allelic diversity index (h), of the several *loci*, was determined as described by Selander and colleagues (1986) [29], using the formula:

$$h = 1 - \sum x_i^2 \left[\frac{n}{(n-1)}\right]$$

where x_i is the frequency of the i-th allele at the *locus* and n is the number of isolates.

A lower triangular matrix of normalized Manhattan distances of *Mah* isolates was created based on the respective VNTR allele profiles [15; 23]. The normalized Manhattan distance between strain X and strain Y was determined using the formula:

$$\frac{\sum_{n=1}^{N} |X_n - Y_n|}{N}$$

where X_n and Y_n are the number of repeat consensus units in the n-th VNTR *locus* (of a total of N *locus*). The distance matrix was used as input for constructing dendrograms using the Fitch-Margoliash algorithm implemented in PHYLIP package (version 3.69). The genotypic diversity was calculated using the Hunter-Gaston Discriminatory Index (HGDI), according to Hunter and Gaston (1988) [30], using the formula:

$$HGDI = 1 - \left[\frac{1}{N(N-1)} \sum_{j=1}^{s} n_j (n_j - 1)\right]$$

where N is the total number of isolates, s the number of typing groups obtained and n_j the number of isolates belonging to the j-th typing group.

A multi-dimensional scaling (MDS) analysis plot was obtained for comparing populations of *Mah* strains from different origins as described by Iwamoto and colleagues (2012) [21], with minor modifications. Briefly, the genetic distances among the different *Mah* populations were calculated by the following function using 6 VNTR *loci* data:

$$D_{PQ} = \sqrt{\sum_{k=1}^{6}\sum_{i=1}^{n}(P_{ki} - Q_{ki})^2}$$

where k represents a VNTR *locus* (out of 6 selected *locus* analyzed in this study), n is the highest copy number of tandem repeat units at *locus* k, P_{ki} is the proportion of strains with the i-th number of tandem repeat units at *locus* k of population P, and Q_{ki} is the proportion of strains with the i-th number of tandem repeat units at *locus* k of population Q. A matrix of genetic distances was created by applying the above function to all combinations of different *Mah* populations (e.g. populations of *Mah* isolates of human and porcine origins). This matrix was normalized, by dividing all genetic distance values by the highest one, converted in a similarity matrix (where genetic similarity values = 1 - normalized genetic distance values) and used for the MDS analysis implemented in NTSYSpc Version 2.02h software (Applied Biostatistics Inc.). The resulting MDS plot represent distinct *Mah* populations as black circles, and the relative genetic similarity (or distance) among populations can be visualized as the distance between the circles.

2.4. Results

2.4.1. VNTR allelic diversity

The VNTR profiles of a subset of 21 *Mah* strains (Table 2.1), originally selected for showing diverse IS*1245*-RFLP profiles [5], were used to estimate the VNTR allelic diversity for all *loci*. The amplification of *locus* MIRU-32 failed for most isolates (Tables 2.1 and 2.2). Among the remaining 19 *loci*, the highest allelic diversity indexes were observed for MATR-3, MATR-6, MATR-7, MATR-8, MATR-11 and MATR-15. The lowest allelic diversity was observed for MATR-1, MATR-12, MATR-13, MIRU-3, MIRU7 and MIRU-47. Seventeen of the 21 *Mah* strains presented distinct MLVA profiles when using these six combined VNTR *loci*. The two pairs of strains (P 9078/0/2005 and P 4212/0/2005; P 10294/0/2005 and WB 10178/2/2006) that showed identical profiles also presented almost identical patterns when using the full set of MATR/MIRU *loci* (Table 2.1).

2.4.2. Genetic diversity of *Mah* isolates

The six most discriminatory *loci* were therefore selected for characterization of an additional 48 animal and 28 human isolates. VNTR allelic profiles, using these *loci*, were obtained for a total of

97 *Mah* clinical isolates from Portugal (Figure 2.1). Fifty three different allelic profiles were found, revealing high genetic heterogeneity, in concordance with previous results using an IS*1245* RFLP typing method [5]. The Hunter-Gaston Discriminatory Index (HGDI) was 0.972, considering the 97 isolates and was estimated as 0.968 and 0.962, respectively, considering only the *Mah* human isolates (18 allelic profiles in 28 isolates) and the porcine isolates (39 allelic profiles in 69 isolates). A dendrogram showing the relatedness of the *Mah* isolates was constructed based on the allelic profiles data set of the six selected VNTR *loci* (Figure 2.1). No clear correlation was observed between the main clusters and the geographical origin, the host (humans or animals) or the biological sample of origin of the *Mah* isolates (Figure 2.1). Identical 6 *loci* VNTR allele profiles were observed in some isolates from pigs (clusters A, B, D, F, G, J, K, L and P in Figure 2.1), although the respective geographic origins were usually different. In contrast, a few isolates collected from distinct pigs in the same farm usually showed distinct profiles: groups of isolates 8222/2, 8222/4 and 8222/6; 9080/1 and 9080/4; and 9395/1 and 9395/3 (Figure 2.1). Several human isolates shared identical profiles and, in most of these cases, strains were isolated in the same hospitals and sometimes in different years (clusters A, C, E, H, M, N and O in Figure 2.1), suggesting a common and persistent source of hospital infection. In a few cases, human isolates shared identical VNTR profiles with porcine isolates (clusters A, C, F and P in Figure 2.1), suggesting a close relatedness between these two groups of strains.

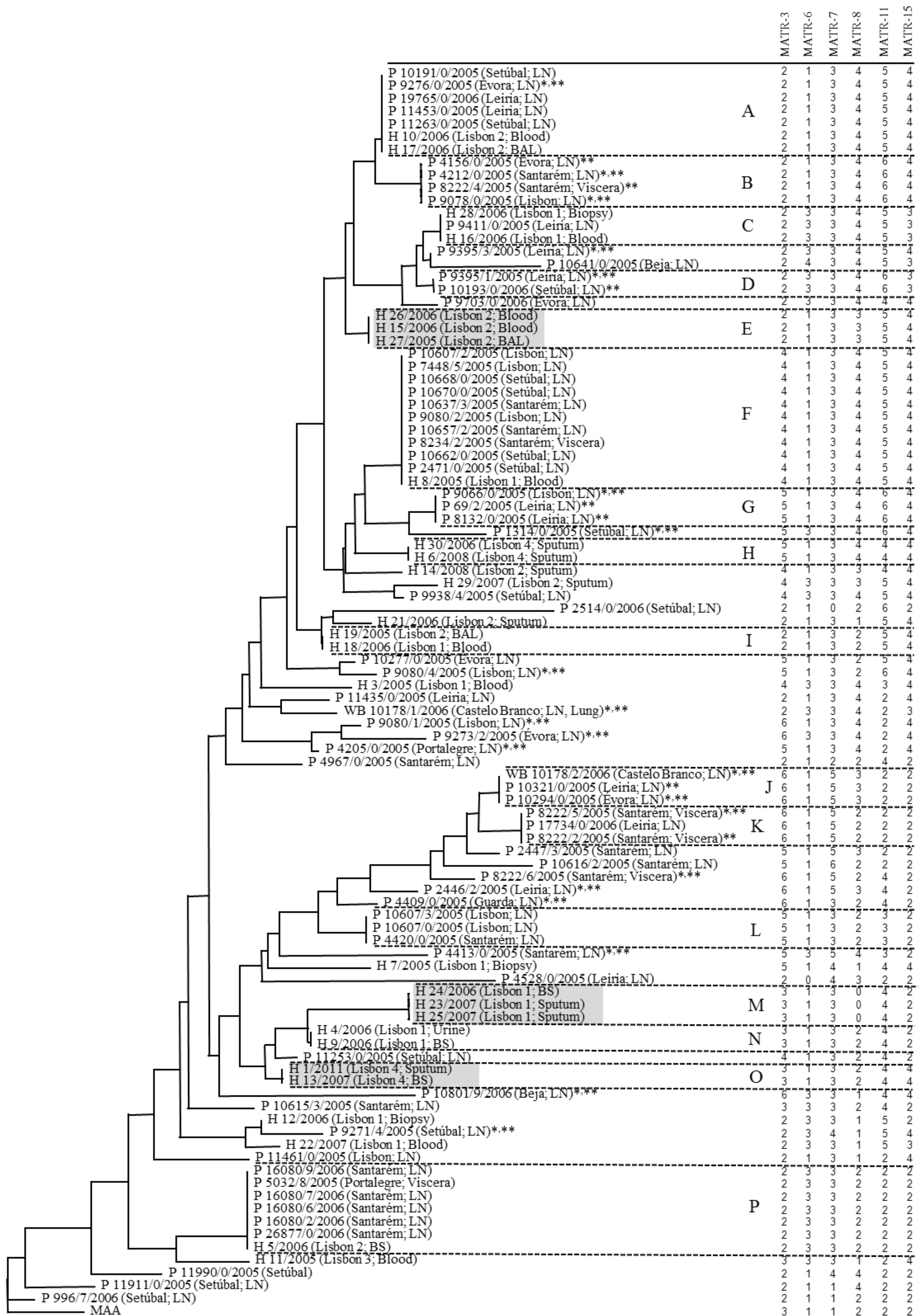
MATR-3 MATR-6 MATR-7 MATR-8 MATR-11 MATR-15
P 10191/0/2005 (Setúbal; LN) 2 1 3 4 5 4
P 9276/0/2005 (Évora; LN)*,** 2 1 3 4 5 4
P 19765/0/2006 (Leiria; LN) 2 1 3 4 5 4
P 11453/0/2005 (Leiria; LN) 2 1 3 4 5 4
P 11263/0/2005 (Setúbal; LN) 2 1 3 4 5 4
H 10/2006 (Lisbon 2; Blood) 2 1 3 4 5 4
H 17/2006 (Lisbon 2; BAL) 2 1 3 4 5 4
A
P 4156/0/2005 (Évora; LN)** 2 1 3 4 6 4
P 4212/0/2005 (Santarém; LN)*,** 2 1 3 4 6 4
P 8222/4/2005 (Santarém; Viscera)** 2 1 3 4 6 4
P 9078/0/2005 (Lisbon; LN)*,** 2 1 3 4 6 4
B
H 28/2006 (Lisbon 1; Biopsy) 2 3 3 4 5 3
P 9411/0/2005 (Leiria; LN) 2 3 3 4 5 3
H 16/2006 (Lisbon 1; Blood) 2 3 3 4 5 3
C
P 9395/3/2005 (Leiria; LN)*,** 2 3 3 4 5 4
P 10641/0/2005 (Beja; LN) 2 4 3 4 5 3
P 9395/1/2005 (Leiria; LN)*,** 2 3 3 4 6 3
P 10193/0/2006 (Setúbal; LN)** 2 3 3 4 6 3
D
P 9703/0/2006 (Évora; LN) 2 3 3 4 4 4
H 26/2006 (Lisbon 2; Blood) 2 1 3 3 5 4
H 15/2006 (Lisbon 2; Blood) 2 1 3 3 5 4
H 27/2005 (Lisbon 2; BAL) 2 1 3 3 5 4
E
P 10607/2/2005 (Lisbon; LN) 4 1 3 4 5 4
P 7448/5/2005 (Lisbon; LN) 4 1 3 4 5 4
P 10668/0/2005 (Setúbal; LN) 4 1 3 4 5 4
P 10670/0/2005 (Setúbal; LN) 4 1 3 4 5 4
P 10637/3/2005 (Santarém; LN) 4 1 3 4 5 4
P 9080/2/2005 (Lisbon; LN) 4 1 3 4 5 4
P 10657/2/2005 (Santarém; LN) 4 1 3 4 5 4
P 8234/2/2005 (Santarém; Viscera) 4 1 3 4 5 4
P 10662/0/2005 (Setúbal; LN) 4 1 3 4 5 4
P 2471/0/2005 (Setúbal; LN) 4 1 3 4 5 4
H 8/2005 (Lisbon 1; Blood) 4 1 3 4 5 4
F
P 9066/0/2005 (Lisbon; LN)*,** 5 1 3 4 6 4
P 69/2/2005 (Leiria; LN)** 5 1 3 4 6 4
P 8132/0/2005 (Leiria; LN)** 5 1 3 4 6 4
G
P 1314/0/2005 (Setúbal; LN)*,** 5 3 3 4 6 4
H 30/2006 (Lisbon 4; Sputum) 5 1 3 4 4 4
H 6/2008 (Lisbon 4; Sputum) 5 1 3 4 4 4
H
H 14/2008 (Lisbon 2; Sputum) 4 1 3 3 4 4
H 29/2007 (Lisbon 2; Sputum) 4 3 3 3 5 4
P 9938/4/2005 (Setúbal; LN) 4 3 3 4 5 4
P 2514/0/2006 (Setúbal; LN) 2 1 0 2 6 2
H 21/2006 (Lisbon 2; Sputum) 2 1 3 1 5 4
H 19/2005 (Lisbon 2; BAL) 2 1 3 2 5 4
H 18/2006 (Lisbon 1; Blood) 2 1 3 2 5 4
I
P 10277/0/2005 (Évora; LN) 5 1 3 2 5 4
P 9080/4/2005 (Lisbon; LN)*,** 5 1 3 2 6 4
H 3/2005 (Lisbon 1; Blood) 4 3 3 4 3 4
P 11435/0/2005 (Leiria; LN) 2 1 3 4 2 4
WB 10178/1/2006 (Castelo Branco; LN, Lung)*,** 2 3 3 4 2 3
P 9080/1/2005 (Lisbon; LN)*,** 6 1 3 4 2 4
P 9273/2/2005 (Évora; LN)*,** 6 3 3 4 2 4
P 4205/0/2005 (Portalegre; LN)*,** 5 1 3 4 2 4
P 4967/0/2005 (Santarém; LN) 2 1 2 2 4 2
WB 10178/2/2006 (Castelo Branco; LN)*,** 6 1 5 3 2 2
P 10321/0/2005 (Leiria; LN)** 6 1 5 3 2 2
P 10294/0/2005 (Évora; LN)*,** 6 1 5 3 2 2
J
P 8222/5/2005 (Santarém; Viscera)*,** 6 1 5 2 2 2
P 17734/0/2006 (Leiria; LN) 6 1 5 2 2 2
P 8222/2/2005 (Santarém; Viscera)** 6 1 5 2 2 2
K
P 2447/3/2005 (Santarém; LN) 5 1 5 3 2 2
P 10616/2/2005 (Santarém; LN) 5 1 6 2 2 2
P 8222/6/2005 (Santarém; Viscera)*,** 6 1 5 2 4 2
P 2446/2/2005 (Leiria; LN)*,** 6 1 5 3 4 2
P 4409/0/2005 (Guarda; LN)*,** 6 1 3 2 4 2
P 10607/3/2005 (Lisbon; LN) 5 1 3 2 3 2
P 10607/0/2005 (Lisbon; LN) 5 1 3 2 3 2
P 4420/0/2005 (Santarém; LN) 5 1 3 2 3 2
L
P 4413/0/2005 (Santarém; LN)*,** 5 3 5 4 3 2
H 7/2005 (Lisbon 1; Biopsy) 5 1 4 1 4 4
P 4528/0/2005 (Leiria; LN) 2 0 4 3 2 2
H 24/2006 (Lisbon 1; BS) 3 1 3 0 4 2
H 23/2007 (Lisbon 1; Sputum) 3 1 3 0 4 2
H 25/2007 (Lisbon 1; Sputum) 3 1 3 0 4 2
M
H 4/2006 (Lisbon 1; Urine) 3 1 3 2 4 2
H 9/2006 (Lisbon 1; BS) 3 1 3 2 4 2
N
P 11253/0/2005 (Setúbal; LN) 4 1 3 2 4 2
H 1/2011 (Lisbon 4; Sputum) 3 1 3 2 4 4
H 13/2007 (Lisbon 4; BS) 3 1 3 2 4 4
O
P 10801/9/2006 (Beja; LN)*,** 6 3 3 1 4 4
P 10615/3/2005 (Santarém; LN) 3 3 3 2 4 2
H 12/2006 (Lisbon 1; Biopsy) 2 3 3 1 5 2
P 9271/4/2005 (Setúbal; LN)*,** 2 3 4 1 5 4
H 22/2007 (Lisbon 1; Blood) 2 3 3 1 5 3
P 11461/0/2005 (Lisbon; LN) 2 1 3 1 2 4
P 16080/9/2006 (Santarém; LN) 2 3 3 2 2 2
P 5032/8/2005 (Portalegre; Viscera) 2 3 3 2 2 2
P 16080/7/2006 (Santarém; LN) 2 3 3 2 2 2
P 16080/6/2006 (Santarém; LN) 2 3 3 2 2 2
P 16080/2/2006 (Santarém; LN) 2 3 3 2 2 2
P 26877/0/2006 (Santarém; LN) 2 3 3 2 2 2
H 5/2006 (Lisbon 2; BS) 2 3 3 2 2 2
P
H 11/2005 (Lisbon 3; Blood) 3 3 3 1 2 4
P 11990/0/2005 (Setúbal) 2 1 4 4 2 2
P 11911/0/2005 (Setúbal; LN) 2 1 1 4 2 2
P 996/7/2006 (Setúbal; LN) 2 1 1 2 2 2
MAA 3 1 1 2 2 2
0.1

Figure 2.1. Dendrogram constructed from the analysis of VNTR profiles (6 selected *loci*) including 28, 67 and 2 *Mycobacterium avium* subsp. *hominissuis* strains of human (H), pig (P) and wild boar (WB) origins, respectively.

The VNTR allele profile is shown for all *Mah* isolates. The year of isolation and strains code is indicated for each isolate and the Portuguese district of origin and the respective sample source is indicated inside parenthesis (BAL – Bronchoalveolar lavage; LN – Lymph nodes; BS – Bronchial secretions; Lisbon 1, 2, 3 and 4 refer to four different hospitals). Strains also tested for the remaining 14 VNTR *loci* analyzed in this study are indicated by an asterisk. Animal strains previously typed by IS*1245*-RFLP [5] are signaled with two asterisks. Clusters of isolates showing the same VNTR profile are indicated by a letter (from A to P). Isolates from the same hospital, but from distinct years and patients, showing 100% similarity are showed on a grey box. The dendrogram was created from a lower triangular distance matrix file by Fitch-Margoliash algorithm implemented in PHYLIP package (the Manhattan distance scale is indicated at the bottom). *Mycobacterium avium* subsp. *avium* (*Maa*) ATCC25291^T is included as reference strain.

2.4.3. Multi-dimensional scaling analysis

A two-dimensional scaling (MDS) plot was constructed in order to visually compare the *Mah* populations of human and porcine origins from Portugal, comprising the total of 97 isolates tested with the 6 selected VNTR *loci* (Figure 2.2). A large dataset containing the VNTR allelic profiles (for the same *loci*) of 258 *Mah* isolates from humans and pigs from Japan [21] was additionally included in our analysis. On the MDS plot, the Portuguese *Mah* populations of human and porcine origins and the Japanese porcine isolates were located close together, indicating a high degree of relatedness. The more divergent *Mah* populations infecting humans in Japan were located more distantly in the plot. Noteworthy, a few cases of identical VNTR allelic profiles were observed between *Mah* isolates from Portugal and Japan: the human isolates H15, H26 and H27 (Cluster E in Figure 2.1), and H4 and H9 (Cluster N) presented a profile identical to that of Japanese porcine isolates; isolates P996/7/2006, P10616/2/2005, H7 and isolates of cluster F (Figure 2.1) presented a profile identical to that of Japanese human isolates.

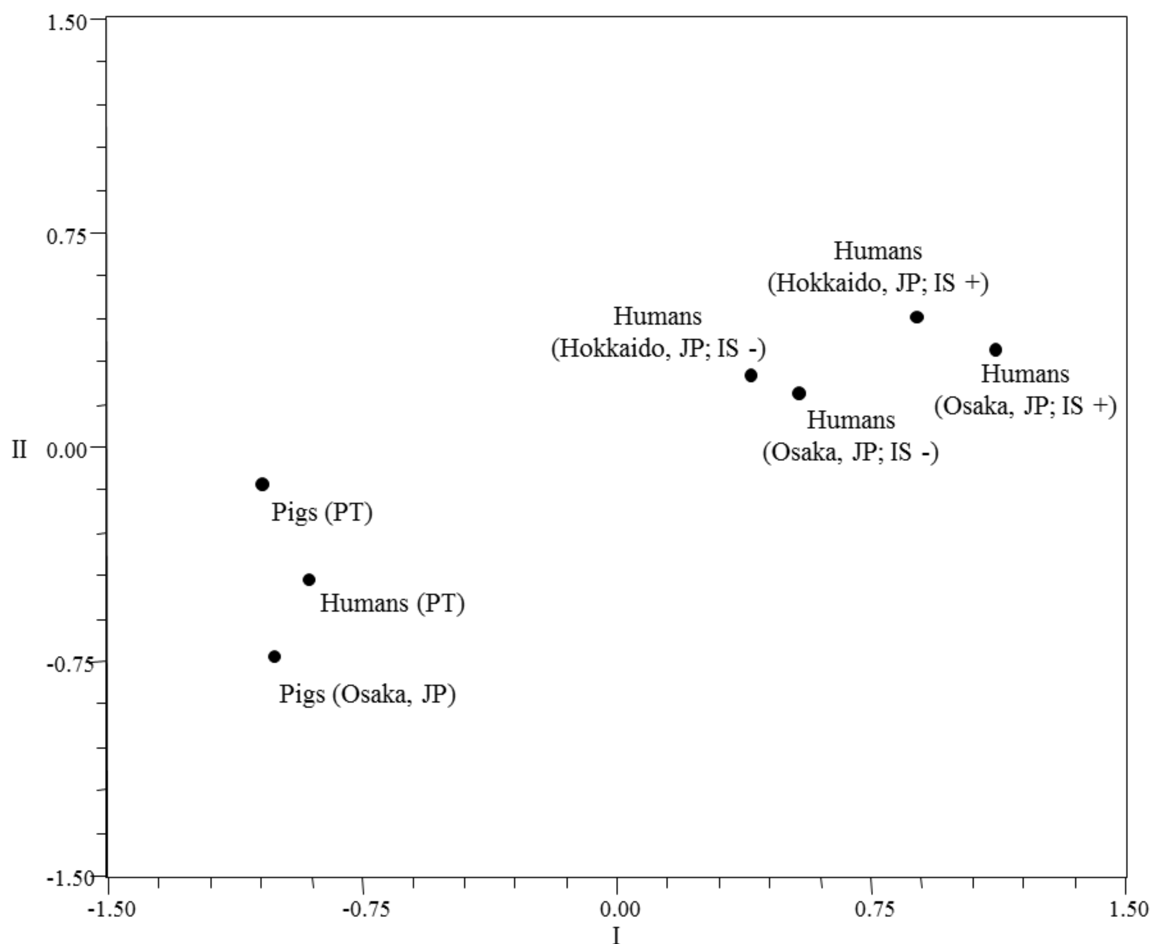

Figure 2.2. Relatedness of *Mycobacterium avium* subsp. *hominissuis* populations of different sources (humans and pigs) and regions (Portugal and Japan) plotted in a multi-dimensional scaling graph based on the genetic distances of the respective VNTR profiles (including the analysis of the six selected *loci* MATR-3, MATR-6, MATR-7, MATR-8, MATR-11 and MATR-15).

PT - Portugal; JP - Japan; IS- and IS+ refers to the absence or presence, respectively, of the IS*901*-like (IS*Mav6*) insertion sequence in the Japanese clinical isolates from humans. VNTR typing data of Japanese isolates were obtained from Iwamoto and colleagues (2012) [21].

2.5. Discussion

Mycobacterium avium subsp. *hominissuis* is an important opportunist pathogen, infecting humans and other animals, notably pigs. Clinical and environmental *Mah* strains are genetically heterogeneous but related genotypes are shared between humans and their living environments, or between humans and pigs [6; 8; 9; 10]. Several molecular methods have been used to unravel the epidemiological traits and sources of infection of *Mah* [7; 9; 12; 13; 14]. Based on the analysis of a set of 15 *Mycobacterium avium* tandem repeats (MATR)-VNTR *loci* Inagaki and colleagues (2009) [15]

described a typing approach for MAC isolates. This technique presents a discriminatory power similar or even higher than the more frequently used IS*1245*-RFLP typing and MIRU-VNTR analysis [15]. The MATR-VNTR typing analysis is being increasingly and successfully used to study the genetic relatedness between human, porcine and/or environmental MAC isolates [20; 21], mainly from Japan, among other applications [18; 22; 23]. In this work we used a MLVA method, mostly based in the analysis of MATR *loci*, to assess the genetic diversity of epidemiological unrelated *Mah* clinical strains of human and animal origins, from diverse geographical regions of mainland Portugal.

We initially selected a set of 15 MATR [15] and 5 MIRU [12; 13] *loci* for our typing analysis. Three of the selected MATR markers are located at the same *loci* of the MIRU-VNTR typing approach: MATR-2 ≈ MIRU-292, MATR-3 ≈ MIRU-X3 and MATR-9≈ MIRU-10. The highest VNTR allelic diversity indexes were observed for MATR-3, MATR-6, MATR-7, MATR-8, MATR-11 and MATR-15. The discriminatory hierarchy of MATR-VNTR *loci* seems to change between *Mah* populations of distinct geographical regions. For example, the most discriminatory *loci* between Japanese *Mah* isolates were MATR-2, MATR-3 and MATR-7 [21]. Considering only the Japanese porcine *Mah* isolates, since the Japanese human isolates are highly divergent (see below), the most discriminatory *loci* were MATR-3 and MATR-8 (similarly to our study), and MATR-16 (which revealed to be much less discriminative among the Portuguese isolates). Therefore, an eventual simplification of the VNTR typing approach for application in large-scale studies in different geographic regions, by reducing the number of PCR-tested *loci* while keeping a high discriminatory power, must take into consideration the average diversity patterns of local *Mah* populations. Using this concept, we selected the six most discriminatory VNTR *loci* among *Mah* isolates from Portugal (MATR-3, MATR-6, MATR-7, MATR-8, MATR-11 and MATR-15) to characterize a larger collection of isolates of human and porcine origins.

The MLVA (with the 6 selected *loci*) showed that the *Mah* clinical isolates of human and porcine origins from Portugal are genetically very heterogeneous, yielding a HGDI of 0.972, which may reflect their widespread geographical origins. The 69 strains of porcine origins were collected from nine Portuguese districts while the 28 strains from human samples were collected from four hospitals of two districts. No clear or strict correlation was observed between specific allelic MLVA signatures and the respective geographical, host and biological sample sources of the *Mah* isolates. Some pig isolates showed identical VNTR allelic profiles but the respective geographic origins were usually different, and isolates from the same farms presented distinct profiles. Overall, our data suggest that *Mah* isolates are not herd-specific. These results corroborate a previous study from our team, where several porcine *Mah* isolates (some also used in this study) were typed by using an IS*1245*-RFLP approach (Figure 2.1) [5]. Most isolates presenting distinct VNTR profiles also displayed different IS*1245*-RFLP banding patterns, as assessed by Domingos and colleagues (2009) [5]. From the several clusters of porcine isolates with identical VNTR profiles (Figure 2.1), for which

we also have IS*1245*-RFLP typing data, one group (cluster D: strains P9395/1/2005 and P10193/0/2006) also shared the same IS*1245*-RFLP banding pattern. Nevertheless, comparison of results from the two typing systems must be done with caution since the RFLP analysis targets the whole genome, while MLVA is based on specific minisatellite regions.

Recently, Eisenberg and colleagues (2012) [31] reported that an increasing number of cases of reproductive disorders and generalized mycobacteriosis in several pig farms were caused by one single virulent *Mah* strain. However, our study and most of the molecular epidemiology studies of *Mah* isolates in other European countries report a high genetic heterogeneity, environmental ubiquity and lack of correlation between genotype groupings and geographical or host (human vs. pig) origins, suggesting that common environmental sources are the most probable origin of infections for both pigs and humans [1; 7; 9; 13; 14].

A few human *Mah* strains isolated in the same hospital, sometimes in different years and from different biological samples, shared identical VNTR profiles in our study (Figure 2.1). Noteworthy, Álvarez and colleagues (2008) [11] reported the isolation of the same *Mah* clone from all environmental sources and from most of the biological samples tested from a Spanish hospital, suggesting that these biological samples were most probably contaminated by the hospital-inhabitant *Mah* clone (from the water distribution system). Moreover, it is worth noting that some of the human isolates also shared identical VNTR profiles with porcine isolates in our study, suggesting their close relatedness.

On the MDS analysis, the Portuguese *Mah* populations of human and porcine origins show a considerable degree of relatedness, when compared to the relatedness between the Japanese human and porcine isolates, reinforcing the possibility that there is a common source of *Mah* infection for pigs and humans in Portugal. Noteworthy, the Japanese porcine isolates also showed a higher degree of relatedness with the Portuguese isolates than with the *Mah* populations infecting humans in Japan. This observation corroborates a previous finding of Iwamoto and colleagues (2012) [21], in a similar MDS analysis using MIRU-VNTR markers, that Japanese pig isolates are more closely related to European than to Japanese human isolates. Previous studies also reported the apparent singularity of the prevalent Japanese human-infecting *Mah* strains, which harbor an unusual IS*901*-like (IS*Mav6*) insertion sequence [21; 32]. The relatedness of global pig *Mah* isolates may suggest the occurrence of common infectious sources for pigs at the global level, such as bedding materials or feed, and/or a global distribution of this pathogen through the international import/export markets of the animals.

Our study shows that MLVA is a useful technique for global evaluation of the genetic diversity of *Mah* human and porcine isolates from Portugal. In this aspect it gives comparable information about the general population structure of *Mah* isolates from a certain region. The *Mah*

population in Portugal is genetically diverse and distinct genotypes are randomly distributed across the country.

The MLVA method could be successfully applied in other countries towards the implementation of a worldwide open-access database of VNTR profiles of *Mah* isolates, allowing a better assessment of the global epidemiology traits of this pathogenic species. An eventual simplification of the VNTR typing approach, by reducing the number of PCR-tested *loci*, must take into consideration the local *Mah* populations.

It has been assumed that both humans and pigs seem to be infected by the pool of environment-inhabiting opportunistic *Mah* strains rather than by specifically or potentially high virulent clones. However, reports addressing this issue are very scarce.

Therefore, comparative analysis using Portuguese *Mah* environmental strains, including from soil or water sources, as well as infection isolates of different hosts (diverse animal species and humans) are envisaged to ascertain if environmental strains are the main source of infections.

Acknowledgements

Célia Leão and Diana Machado are a recipients of PhD grants from "Fundação para a Ciência e a Tecnologia" SFRH/BD/62469/2009 and SFRH/BD/65060/2009, respectively. This work was partially funded by the project PTDC/CVT/111634/2009 from "Fundação para a Ciência e a Tecnologia". We are indebted to all the clinicians and laboratories from the collaborating hospitals that helped in this work.

Author's contribution

CL contributed to the experimental work of growing and characterization of animal isolates, DNA extraction, VNTR PCRs, data analysis and writing the manuscript. AC collaborated with the MIRU-VNTR PCRs. DM isolated and characterized human isolates and revised the manuscript. ISS, IC, MV, JI and AB contributed to the designing of the study and revised the manuscript. All authors have read and approved the final manuscript.

2.6. References

1. **Turenne CY, Wallace R Jr, Behr MA**. 2007. *Mycobacterium avium* in the postgenomic era. Clin Microbiol Rev **20**:205–229.
2. **Inderlied C, Kemper C, Bermudez L**. 1993. The *Mycobacterium avium* Complex. Clin Microbiol Rev **6**:266-310.
3. **Despierres L, Cohen-Bacrie S, Richet H, Drancourt M**. 2012. Diversity of *Mycobacterium avium* subsp. *hominissuis* mycobacteria causing lymphadenitis, France. Eur J Clin Microbiol Infect Dis **31**:1373-1379.
4. **Ignatov D, Kondratieva E, Azhikina T, Apt A**. 2012. *Mycobacterium avium*-triggered diseases: pathogenomics. Cell Microbiol **14**:808-818.
5. **Domingos D, Amado A, Botelho A. 2009**. IS*1245* RFLP analysis of strains of *Mycobacterium avium* subspecies *hominissuis* isolated from pigs with tuberculosis lymphadenitis in Portugal. Vet Rec **164**:116-120.
6. **Möbius P, Lentzsch P, Moser I, Naumann L, Martin G, Köhler H**. 2006. Comparative macrorestriction and RFLP analysis of *Mycobacterium avium* subsp. *avium* and *Mycobacterium avium* subsp. *hominissuis* isolates from man, pig and cattle. Vet Microbiol **117**:284–291.
7. **Álvarez J, Castellanos E, Romero B, Aranaz A, Bezos J, Rodríguez S, Mateos A, Domínguez L, de Juan L.** 2011. Epidemiological investigation of a *Mycobacterium avium* subsp. *hominissuis* outbreak in swine. Epidemiol Infect **139**:143–148.
8. **Johansen TB, Olsen I, Jensen MR, Dahle UR, Holstad G, Djonne B**. 2007. New probes used for IS*1245* and IS*1311* restriction fragment length polymorphism of *Mycobacterium avium* subsp. *avium* and *Mycobacterium avium* subsp. *hominissuis* isolates of human and animal origin in Norway. BMC Microbiol **7**:14.
9. **Tirkkonen T, Pakarinen J, Rintala E, Ali-Vehmas T, Marttila H, Peltoniemi OA, Mäkinen J**. 2010. Comparison of variable-number tandem-repeat markers typing and IS*1245* restriction fragment length polymorphism fingerprinting of *Mycobacterium avium* subsp. *hominissuis* from human and porcine origins. Acta Vet Scand **52**:21.
10. **Pate M, Zolnir-Dovc M, Krt B, Ocepek M**. 2008. IS*1245* RFLP-based genotyping study of *Mycobacterium avium* subsp. *hominissuis* isolates from pigs and humans. Comp Immunol Microbiol Infect Dis **31**:537-550.
11. **Álvarez J, García IG, Aranaz A, Bezos J, Romero B, de Juan L, Mateos A, Gómez-Mampaso E, Domínguez L**. 2008. Genetic diversity of *Mycobacterium avium* isolates recovered from clinical samples and from the environment: molecular characterization for diagnostic purposes. J Clin Microbiol **46**:1246-1251.
12. **Thibault VC, Grayon M, Boschiroli ML, Hubbans C, Overduin P, Stevenson K, Gutierrez MC, Supply P, Biet** F. 2007. New variable-number tandem-repeats markers for typing *Mycobacterium avium* subsp. *paratuberculosis* and *M. avium* strains: comparison with IS*900* and IS*1245* restriction fragment length polymorphism typing. J Clin Microbiol **45**:2404-2410.
13. **Radomski N, Thibault VC, Karoui C, de Cruz K, Cochard T, Gutiérrez C, Supply P, Biet F, Boschiroli ML**. 2010. Determination of genotypic diversity of *Mycobacterium avium* subspecies from human and animal origins by mycobacterial interspersed repetitive-unit-variable-number tandem-repeat and IS*1311* restriction fragment length polymorphism typing methods. J Clin Microbiol **48**:1026-1034.
14. **Pate M, Kušar D, Zolnir-Dovč M, Ocepek M**. 2011. MIRU-VNTR typing of *Mycobacterium avium* in animals and humans: heterogeneity of *Mycobacterium avium* subsp. *hominissuis* versus homogeneity of *Mycobacterium avium* subsp. *avium* strains. Res Vet Sci **91**:376-381.
15. **Inagaki T, Nishimori K, Yagi T, Ichikawa K, Moriyama M, Nakagawa T, Shibayama T, Uchiya K, Nikai T, Ogawa K**. 2009. Comparison of a variable number tandem-repeat (VNTR) method for typing *Mycobacterium avium* with mycobacterial interspersed repetitive-

unit-VNTR and IS*1245* restriction fragment length polymorphism typing. J Clin Microbiol **47**:2156–2164.

16. **Dauchy FA, Dégrange S, Charron A, Dupon M, Xin Y, Bébéar C, Maugein J**. 2010. Variable-number tandem-repeat markers for typing *Mycobacterium intracellulare* strains isolated in humans. BMC Microbiol **10**:93.
17. **Ichikawa K, Yagi T, Inagaki T, Moriyama M, Nakagawa T, Uchiya K, Nikai T, Ogawa K**. 2010. Molecular typing of *Mycobacterium intracellulare* using multilocus variable-number of tandem-repeat analysis: identification of *loci* and analysis of clinical isolates. Microbiol **156**:496-504.
18. **Kikuchi T, Watanabe A, Gomi K, Sakakibara T, Nishimori K, Daito H, Fujimura S, Tazawa R, Inoue A, Ebina M, Tokue Y, Kaku M, Nukiwa T**. 2009. Association between mycobacterial genotypes and disease progression in *Mycobacterium avium* pulmonary infection. Thorax **64**:901-907.
19. **Kikuchi T, Kobashi Y, Hirano T, Tode N, Santoso A, Tamada T, Fujimura S, Mitsuhashi Y, Honda Y, Nukiwa T, Kaku M, Watanabe A, Ichinose M**. 2014. *Mycobacterium avium* genotype is associated with the therapeutic response to lung infection. Clin Microbiol Infect **20**:256-262.
20. **Fujita K, Ito Y, Hirai T, Maekawa K, Imai S, Tatsumi S, Niimi A, Iinuma Y, Ichiyama S, Mishima M**. 2012. Genetic relatedness of *Mycobacterium avium-intracellulare* complex isolates from patients with pulmonary MAC disease and their residential soils. Clin Microbiol Infect **19**:537-541.
21. **Iwamoto T, Nakajima C, Nishiuchi Y, Kato T, Yoshida S, Nakanishi N, Tamaru A, Tamura Y, Suzuki Y, Nasu M**. 2012. Genetic diversity of *Mycobacterium avium* subsp. *hominissuis* strains isolated from humans, pigs, and human living environment. Infect Genet Evol **12**:846-852.
22. **Taga S, Niimi M, Kurokawa K, Nakagawa T, Ogawa K**. 2012. A case of environmental infection with pulmonary *Mycobacterium avium* complex disease from a residential bathroom of a patient suggested by variable-number tandem-repeat typing of *Mycobacterium avium* tandem repeat *loci*. Kekkaku **87**:409-414 (*Japanese, abstract in English*).
23. **Tatano Y, Sano C, Yasumoto K, Shimizu T, Sato K, Nishimori K, Matsumoto T, Yano S, Takeyama H, Tomioka H**. 2012. Correlation between variable-number tandem-repeat-based genotypes and drug susceptibility in *Mycobacterium avium* isolates. Eur J Clin Microbiol Infect Dis **31**:445-454.
24. **Rindi L, Buzzigoli A, Medici C, Garzelli C**. 2013. High phylogenetic proximity of isolates of *Mycobacterium avium* subsp. *hominissuis* over a two decades-period. Infect Genet Evol **16**:99-102.
25. **Muwonge A, Oloya J, Kankya C, Nielsen S, Godfroid J, Skjerve E, Djønne B, Johansen TB**. 2014. Molecular characterization of *Mycobacterium avium* subspecies *hominissuis* isolated from humans, cattle and pigs in the Uganda cattle corridor using VNTR analysis. Infect Genet Evol **21**:184-91.
26. **Couto I, Machado D, Viveiros M, Rodrigues L, Amaral** L. 2010. Identification of nontuberculous mycobacteria in clinical samples using molecular methods: a 3-year study. Clin Microbiol Infect **16**:1161-1164.
27. **Kunze ZM, Portaels F, Mcfadden JJ**. 1992. Biologically distinct subtypes of *Mycobacterium avium* differ in possession of insertion sequence IS*901*. J Clin Microbiol **30**:2366-2372.
28. **Guerrero C, Bernasconi C, Burki D, Bodmer T, Telenti A**. 1995. A novel insertion element from *Mycobacterium avium*, IS*1245*, is a specific target for analysis of strain relatedness. J Clin Microbiol **33**:304–307.
29. **Selander RK, Caugant DA, Ochman H, Musser JM, Gilmour MN, Whittam TS**. 1986. Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. Appl Environ Microbiol **51**:873–884.
30. **Hunter PR, Gaston MA**. 1988. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. J Clin Microbiol **26**:2465–2466.

31. **Eisenberg T, Volmer R, Eskens U, Moser I, Nesseler A, Sauerwald C, Seeger H, Klewer-Fromentin K, Möbius P**. 2012. Outbreak of reproductive disorders and mycobacteriosis in swine associated with a single strain of *Mycobacterium avium* subspecies *hominissuis*. Vet Microbiol **14**:69-76.
32. **Ichikawa K, Yagi T, Moriyama M, Inagaki T, Nakagawa T, Uchiya K, Nikai T, Ogawa K**. 2009. Characterization of *Mycobacterium avium* clinical isolates in Japan using subspecies-specific insertion sequences, and identification of a new insertion sequence, IS*Mav6*. J Med Microbiol **58**:945-950.

Chapter III

Search for *Mycobacterium avium* subsp. *paratuberculosis* in Portuguese asymptomatic cattle

Célia Leão, Ana Amaro, Ilda Santos-Sanches, João Inácio, Ana Botelho

Manuscript published with minor modifications in 2015 in "Revista Portuguesa de Ciências Veterinárias" vol. 110 pages 69-73

"Com autorização do editor e sujeitos aos direitos de cópia impostos pelo mesmo "

Search for *Mycobacterium avium* subsp. *paratuberculosis* in Portuguese asymptomatic cattle

3.1. Abstract

Mycobacterium avium subsp. *paratuberculosis* (*Map*) is the causative agent of paratuberculosis, one of the most important diseases in cattle worldwide, imposing a relevant economic impact for the livestock industry. Paratuberculosis is a chronic intestinal granulomatous infection manifested by a progressive and fatal weight loss, significant decrease of milk production, infertility, oedema and diarrhoea. However, animals can remain asymptomatic for two to five years shedding the agent in faeces leading to the spread of the disease. Paratuberculosis is considered an underdiagnosed disease in Portugal and the real prevalence in cattle is unknown. The aim of this study was to assess the presence of *Map* in apparently healthy and asymptomatic Portuguese cattle. Faecal samples from 24 bovines were analysed. The samples were screened for the presence of *Map* by an IS*900*-targeted PCR assay and by culture in specific media. The isolates were confirmed to be acid-fast bacilli by auramine-rhodamine staining, and further identified as *M. avium* subsp. *paratuberculosis* by the presence of the IS*900* and F*57* elements in their genomes. Further characterization of the isolates was performed by a Multiple Loci VNTR Analysis (MLVA) approach. From the 24 faecal samples 22 were IS*900*-PCR positive and from these 12 yielded positive *Map* cultures. The 12 *Map* isolates shared an identical MLVA profile, also corresponding to the INMV2 genotype. This is the first study reporting the isolation, identification and typing of *Map* from Portuguese cattle. Paratuberculosis may be more widespread in Portugal than initially expected and asymptomatic animals are shedding the agent in their faeces, perpetuating the cycle of infection.

Keywords

Paratuberculosis; Johne's disease; asymptomatic cattle; IS*900*-PCR; MIRU-VNTR; *Mycobacterium avium* subsp. *paratuberculosis*

3.2. Introduction

Mycobacterium avium subsp. *paratuberculosis* (*Map*) is the causative agent of paratuberculosis, or Johne's disease, a chronic granulomatous enteritis affecting a wide range of animals, especially ruminants [1]. Other animals like rabbits, foxes, birds, domestic and wildlife animals can also be infected [2; 3]. Paratuberculosis is considered a widespread and very relevant disease in the livestock industry due to its considerable economic impact [4]. *Map* infection has also become a highly contentious issue regarding its potential implication in the etiology of human inflammatory bowel disease, known as Crohn's disease [5; 6]. *Map* is an intracellular pathogen that infects and multiplies inside macrophages in the animal's gut. The transmission of *Map* usually occurs by the faecal-oral route during the animal's first months of life, being newborn animals the most susceptible. The disease may be triggered by stressful conditions such as giving birth and overcrowding. The clinical signs of the disease appear between the first two to five years of life, after a long period without clinical signs. They include the diminution of milk production, infertility, oedema and progressive and fatal weight loss with diarrhoea, initially intermittent but becoming progressively more severe. Nevertheless, some animals exposed to the agent may be infected but successfully contain or clear infection while others can remain asymptomatic without developing the disease [3; 7]. The diagnosis of paratuberculosis is difficult and time consuming due to the characteristics of the agent and is also influenced by the disease stage. Culture remains the gold standard method for the detection of *Map* in biological samples, requiring the supply of an exogenous mycobactin source, and may take up to six months due to the extreme fastidious growth of this agent. The introduction of molecular diagnostic techniques has contributed to more rapid, sensitive and specific detection and characterization of *Map* [2]. The multi-copy IS*900* and the single-copy F57 elements are the most important *Map*-specific genomic targets used in molecular assays [8].

Paratuberculosis is considered an underdiagnosed disease in Portugal and the real prevalence in cattle is not known. Since shedding can contribute to the silent maintenance of the infection cycle of *Map*, the aim of this study was to preliminarily appraise the presence of this agent in faeces of asymptomatic cattle from the North of Portugal, an important dairy farming region.

3.3. Materials and Methods

3.3.1. Faecal samples

Twenty four faecal samples, collected from four farms, "Barcelos" (n = 7), "Póvoa de Varzim" (n = 10), "Vila do Conde" (n = 5) and "Vila Nova de Famalicão" (n = 2) in the North of Portugal, from cattle five to ten years old and without clinical signs of disease, were analyzed. The samples were stored at -20ºC until they were processed for culture and PCR assays.

3.3.2. Culture and identification of isolates

Faecal samples were prepared for culture as described in the OIE Terrestrial Manual (2014) [1], with minor modifications. Briefly, 20 mL of sterile distilled water were added to 1 g of faeces and stirred at room temperature for 30 minutes. After settling for 30 minutes, 5 mL of the uppermost suspension were transferred to a new tube containing 20 mL of 0.9% hexadecylpyridinium chloride (HPC), inverted several times and allowed to stand undisturbed at room temperature for 18 hours. After this period, the sediment was carefully transferred to a new tube, washed with 10 mL of sterile distilled water and centrifuged at 900 × g for 30 minutes at room temperature. The pellet was resuspended in 500 µL of sterile distilled water and volumes of 100 µL were inoculated onto slopes of Herrold's egg yolk medium (HEYM) with and without Mycobactin J. The incubation was performed at 37 ºC for 6 months. Isolated colonies were confirmed to be acid-fast bacilli by auramine-rhodamine staining as described by [9]. Isolates were identified as *Map* by IS*900*-PCR [10] and nested *F57*-PCR [11], using DNA extracted by a simple boiling method. For this purpose, one colony was transferred to a microtube containing 100 µL of TE buffer (Tris EDTA, buffer pH 8, 10 mM Tris.Cl, 1 mM EDTA), vortexed and centrifuged at 2000 × g for 5 minutes at room temperature. The pellet was resuspended again in 100 µL of TE buffer and incubated for 45 minutes at 95 ºC. After bacterial inactivation, the culture was centrifuged at 2000 × g for 1 minute and the supernatant containing the DNA was transferred to a new microtube and used as template for PCR reactions.

3.3.3. DNA extraction and IS*900*-targeted PCR amplification

Total DNA from faecal samples was extracted with QIAamp® DNA Stool mini kit (Qiagen), stool pathogen detection protocol, as described by the manufacturer. The evaluation of the presence of *Map* in faecal samples was performed using a standard IS*900*-PCR assay, described by [10]. The reaction was carried out in a final volume of 25 µl containing 250 µM of each deoxynucleotide triphosphate (Promega), 2.0 mM of $MgCl_2$ (Promega), 0.5 µM of each primer (Primer 90, 5'-GTT CGG GGC CGT CGC TTA GG-3'; and Primer 91, 5'-GAG GTC GAT CGC CCA CGT GA-3'), 1 U

of GoTaq® DNA polymerase and 1× of the respective buffer (Promega), and 2.5 µl of the extracted DNA solution. The amplification was performed in a MJmini™ Thermocycler (BioRad) with an initial step at 95 ºC for 5 minutes, followed by 35 cycles at 95 ºC for 1 minute, 60ºC for 1 minute and 72 ºC for 1 minute, ending with a step at 72 ºC for 3 minutes. DNA from *M. avium* subsp. *paratuberculosis* (ATCC 19698^T) and ultrapure sterilized water were used as positive and negative controls of amplification, respectively. The amplified products were electrophoretically analyzed in a 1.5% (w/v) agarose gel in 1× Tris-Borate-EDTA (TBE) buffer stained with GelRed®. Gel electrophoresis images were acquired with an UV Transilluminator UVP M-20.

3.3.4. Characterization of *Map* isolates by MLVA

A set of ten VNTR *loci* was selected for the molecular discrimination of *Map* isolates: VNTR-3, VNTR-7, VNTR-10, VNTR-47, MIRU-2, MIRU-3 (alias X3), VNTR-25, VNTR-32, VNTR-259 and VNTR-292 [12; 13]. Briefly, PCR reactions were carried out in a total volume of 25 µl containing 1× buffer (Promega), 200 µM of each deoxynucleotide triphosphate (Promega), 2.0 mM of $MgCl_2$ (Promega), 1 µM of each forward and reverse primers (for each *locus* separately; see Table 3.1), 2.5 U of GoTaq® DNA polymerase and 1 µl of extracted DNA solution. Amplification was performed in a MJmini™ Thermocycler (BioRad) with an initial step at 95 ºC for 10 minutes, followed by 38 cycles at 98 ºC for 10 seconds, 60ºC for 30 seconds and 72 ºC for 1 minute, ending with a step at 72 ºC for 7 minutes. DNA from *M. avium* subsp. *paratuberculosis* (ATCC 19698^T) was used as positive control and amplicon size reference. PCR products were analyzed in a 2.5% (w/v) agarose gel in 1× Tris-Borate-EDTA (TBE) buffer, stained with GelRed®, at 45 V for 2 hours, using a 100 bp molecular ladder to estimate the size of the PCR amplicons. Allele calling tables available by [13] and the INMV database (http://mac-inmv.tours.inra.fr) were used to assign the hypothetical size of each PCR product, for *loci* VNTR-3, VNTR-7, VNTR-10, VNTR-47, MIRU-2, MIRU-3, VNTR-25, VNTR-32, VNTR-259 and VNTR-292, to their correspondent tandem repeat copy number.

Table 3.1. Sequences of the primers used for the MLVA characterization of the isolates

Locus	**Primer forward (5'-3')**	**Primer reverse (5'-3')**
VNTR-3	CAT ATC TGG CAT GGC TCC AG	ATC GTG TTG ACC CCA AAG AAA T
VNTR-7	GAC AAC GAA ACC TAC CTC GTC	GTG AGC TGG CGG CCT AAC
VNTR-10	GAC GAG CAG CTG TCC GAG	GAG AGC GTG GCC ATC GAG
VNTR47	CGT TGC GAT TTC TGC GTA GC	GGT GAT GGT CGT GGT CAT CC
MIRU-2	GAA CGA AGA TCC TGG GAC TG	CGA CGA CGA ACA CCT CAA C
MIRU-3 (alias X3)	AAC GAG AGG AAG AAC TAA GCC G	TTA CGG AGC AGG AAG GCC AGC GGG
VNTR-25	GTC AAG GGA TCG GCG AGG	TGG ACT TGA GCA CGG TCA T
VNTR-32	CCA CAG GGT TTT TGG TGA AG	GGA AAT CCA ACA GCA AGG AC
VNTR-259	GGG TGT GGA GCT ACG ACT TC	GAG CTG CTT GAC CAG GTG AT
VNTR-292	CTT GAG CAG CTC GTA AAG CGT	GCT GTA TGA GGA AGT CTA TTC ATG G

3.4. Results

Of the 24 faecal samples from asymptomatic Portuguese bovines analysed, 22 were IS*900*-PCR positive (2 from "Vila Nova de Famalicão", 4 from "Vila do Conde", 7 from "Barcelos" and 9 from "Póvoa de Varzim") (Table 2), exhibiting a *Map*-specific 400 bp amplified product. Twelve cultures were obtained after 60 days incubation, from the 24 inoculated samples (1 from "Barcelos", 4 from "Vila do Conde" and 7 from "Póvoa de Varzim") (Table 3.2). All 12 culture positive samples yielded IS*900*-PCR positive results. The 12 isolates were confirmed to be acid-fast bacilli by auramine-rhodamine staining and were identified as *Map* by PCR, showing the specific 400 bp and 424 bp products for the *IS*900 and *F57* PCR-amplified genomic targets, respectively. The isolates were further characterized by a ten *loci* MLVA analysis approach, showing that all shared the same profile (Table 3.2), also identical to the control *Map* strain ATCC 19698^{T}. When considering only the eight VNTR *loci* set analysed by Thibault *et al*. (2007) [12] (VNTR 292, MIRU 3, VNTR 25, VNTR 47, VNTR 3, VNTR 7, VNTR 10 and VNTR 32), the *Map* isolates were shown to belong to the INMV2 type (Table 3.2), similarly to the control strain ATCC 19698^{T}. When using the six *loci* set proposed by Castellanos *et al*. (2010) [13] (MIRU 2, MIRU 3, VNTR 25, VNTR 32, VNTR 292 and VNTR 259), the respective profile 323832 was not found among their 70 *Map* Spanish isolates.

Table 3.2. Results obtained for the culture-based and molecular detection of *Map* in bovine faecal samples

Geographical origin	Number of samples	Number of *IS*900-PCR positive samples	Number of culture positive samples	MLVA profile[a]
Barcelos	7	7	1	3233222832
Póvoa do Varzim	10	9	7	3233222832
Vila do Conde	5	4	4	3233222832
Vila Nova de Famalicão	2	2	0	3233222832
Total	24	22	12	

[a]Number of tandem repeat copies in the order VNTR292 - MIRU3 - VNTR25 - VNTR47 - VNTR3 - VNTR7 - VNTR10 - VNTR32 - MIRU2 - VNTR259; The profile corresponding to the first eight *loci* match the *Map* INMV2 type, according to Thibault *et al*. (2007) [12].

3.5. Discussion

The worldwide herd prevalence of paratuberculosis is estimated to be 7 to 40%, based on serological monitoring tests [2]. However, an accurate estimation of the prevalence of *Map* in cattle is difficult since most infected animals are asymptomatic, the diagnosis in the early stages of disease is difficult and the animals with clinical signs of decreased milk production can be slaughtered without confirmation of *Map* infection [14]. To date, only a few studies have tried to assess the prevalence of paratuberculosis in Portugal, namely in the cattle population, where this disease probably runs under-diagnosed. Ferreira and colleagues (2002) [15] reported 4.8-7.0% of bovines serologically positive to *Map* in the "Alentejo" region, South of Portugal, with 13-25% of herds positive for the disease. In another more recent study, anti-*Map* antibodies were detected in 2.3% of milk samples collected from 5294 milking cows from the Northern region of Portugal, corresponding to 45.9% of infected herds/farms [16]. A few other studies are available regarding the seroprevalence of paratuberculosis in small ruminants in several regions of Portugal [15; 17; 18; 19; 20; 21; 22; 23]. The prevalence of paratuberculosis at the flock/herd level may be high, with values ranging from 47% to 67% [20; 23]. However, the serological detection of *Map* was not always correlated with the presence of clinical signs of disease in animals, noteworthy in sheep [18]. *Map* was also previously isolated in Portugal from the mesenteric lymph nodes of wild boars (*Sus scrofa*) with granulomatous lymphadenitis [24],

from kidney samples of wild red deer (*Cervus elaphus*) [25] and from Eurasian otters (*Lutra lutra*) [26]. The agent was also detected by PCR-based methods in tissues of domestic pigs [27]. There are, presently, in Portugal no reliable data about shedding of *Map* in faeces.

Detection of *Map* in faeces by IS*900*-PCR is an efficient and rapid method when compared with the conventional culture-based assays, which take up 8 to 16 weeks to obtain a result. However, bacteriological culture is the gold standard methodology for the diagnosis of paratuberculosis and the isolation of the agent is required if further studies are intended. In this study 24 faecal samples from asymptomatic Portuguese bovines from the North of Portugal were analysed. Twenty two samples were found to be IS*900*-PCR positive (91.7%) while only 12 *Map* isolates were obtained from them (50%). Our preliminary data suggests that *Map* infection in cattle may be more prevalent in Portuguese cattle than initially expected, based on the previous surveys referred to above, employing mainly serological assays. Even with the absence of clinical signs, our data points out that the animals are shedding the agent in faeces, even with the possibility of being a passive shedding, perpetuating the cycle of infection.

The analysis of the polymorphisms in MIRU/VNTR *loci* has proven to be very useful for the discrimination of *Map* isolates [12; 13; 28] and the correspondent alleles have been found to be very stable after several subcultures *in vitro*, on different media, and after *in vivo* passage [29]. However, several distinct sets of MIRU/VNTR *loci* have been used to characterize *Map* isolates [12; 13; 30], which makes it difficult to compare between different studies. In this work we used the MIRU/VNTR *loci* set proposed by Thibault *et al*. (2007) [12] and Castellanos *et al*. (2010) [13], which were used before to genotype major collections of *Map* isolates from different countries, allowing the comparison of the correspondent allelic profiles with the profiles of the Portuguese isolates. The 12 Portuguese *Map* isolates shared the same MLVA allelic profile, suggesting that they belong to the same clonal lineage. This profile corresponds to the *Map* INMV2 type, according to [12]. The INMV2 type seems to be, together with INMV1, the most abundant in Europe [12]. For example, 35% and 61% of the French and Dutch bovine *Map* isolates, respectively, analysed by Thibault *et al*. (2007) [12] represented the INMV2 type. This type is widely disseminated, occurring in many other countries such as Germany, Czech Republic, Finland, Scotland, Greece, Spain and Canada [28; 31; 32]. Our preliminary data suggests that INMV2 *Map* strains are also abundant in Portugal and potentially circulate in the environment, by the shedding of the agent in the faeces of infected cattle.

This is the first study reporting the isolation and identification of *Map* from Portuguese asymptomatic cattle, along with its molecular characterization with a MLVA approach. Identification of shedding animals is extremely important for the prevention of the spread of *Map* infection. We highlight the need for systematic evaluation for the presence of shedding bovines in subclinically infected dairy herds and this can be accomplished by the use of PCR-based assays that can be applied at the herd or individual level, regardless of animal age or production stage. Testing of additional and

different cattle samples - for instance milk - from different geographical regions, are currently underway in order to have a clearer picture of the real situation of paratuberculosis in cattle in Portugal.

Acknowledgements

DVM Pedro Meireles is acknowledged for providing bovine fecal samples. Célia Leão is recipient of a PhD grant with reference SFRH/BD/62469/2009, from the Portuguese "Fundação para a Ciência e a Tecnologia". This work was partially funded by the project PTDC/CVT/111634/2009 from "Fundação para a Ciência e a Tecnologia".

Author's contribution

CL contributed to the designing of the study, experimental work of samples manipulation, growing and characterization of isolates, DNA extraction, VNTR PCRs, data analysis and writing the manuscript. AA contributed to the laboratory work and revised the manuscript. ISS, JI and AB contributed to the designing of the study and revised the manuscript. All authors have read and approved the final manuscript

3.6. References

1. **OIE. 2014**. Paratuberculosis (Johne's Disease). In OIE Terrestrial Manual (World Organisation for Animal Health), Chapter 2.1.11.
2. **Timms VJ, Gehringer MM, Mitchell HM, Daskalopoulos G, Neilan BA**. 2011. How accurately can we detect *Mycobacterium avium* subsp. *paratuberculosis* infection? J Microbiol Meth **85**:1-8.
3. **Salem M, Heydel C, El-Sayed A, Ahmed SA, Zschöck, Baljer G**. 2012. *Mycobacterium avium* subspecies *paratuberculosis*: an insidious problem for the ruminant industry. Trop Anim Health Prod. **45**:351-366.
4. **Englund S**. 2002. Molecular biology techniques as a tool for detection and characterization of *Mycobacterium avium* subsp. *paratuberculosis*. Ph.D. thesis. Swedish University of Agriculture Science, Uppsala, Sweden.
5. **Grant IR**. 2005. Zoonotic potential of *Mycobacterium avium* ssp. *paratuberculosis*: the current position. J Appl Microbiol **98**:1282-1293.
6. **Juste RA, Perez V**. 2011. Control of paratuberculosis in sheep and goats. Vet Clin Food Anim **27**:127–128.
7. **Atreya R, Bülte M, Gerlach G F, Goethe R, Hornef M H, Köhler H, Meens J, Möbius P, Roeb E, Weiss S**, on behalf of the ZooMAP Consortium. 2014. Facts, myths and hypotheses on zoonotic nature of Mycobacterium avium subspecies paratuberculosis. Int J Med Microbiol **304**:858-867.
8. **Bölske G, Herthnek D**. 2010. Diagnosis of Paratuberculosis by PCR. *In* Behr MA, Collins DM (eds) Paratuberculosis: Organism, Disease, Control. CAB International, Cambridge, MA, pp 267-283.
9. **Bird BB, Madison BM**. 2000. Use of fluorochrome staining for detecting acid-fast mycobacteria. *In* Current Laboaratory Practice Series. Department of Health and Human Services. Centers for Diseases Control and Prevention, Atlanta, pp 1-17.
10. **Sanderson JD, Moss MT, Tizard ML, Hermon-Taylor J.** 1992. *Mycobacterium paratuberculosis* DNA in Crohn's disease tissue. Gut **33**:890–896.
11. **Vansnick E, De Rijk P, Vercammen F, Geysen D, Rigouts L, Portaels F**. 2004. Newly developed primers for the detection of *Mycobacterium avium* subspecies *paratuberculosis*. Vet Microbiol **100**:197–204.
12. **Thibault VC, Grayon M, Boschiroli ML, Hubbans C, Overduin P, Stevenson K, Gutierrez MC, Supply P, Biet F**. 2007. New variable-number tandem-repeat markers for typing *Mycobacterium avium* subsp. *paratuberculosis* and *M. avium* strains: comparison with IS*900* and IS*1245* restriction fragment length polymorphism typing. J Clin Microbiol **45**:2404–2410.
13. **Castellanos E, Romero B, Rodriguez S de JL, Bezos J, Mateos A, Dominguez L, Aranaz A**. 2010. Molecular characterization of *Mycobacterium avium* subspecies *paratuberculosis* Types II and III isolates by a combination of MIRU-VNTR *loci*. Vet Microbiol **144:**118–126.
14. **Fecteau M-E, Whitlock R H**. 2010. Paratuberculosis in cattle. *In* Behr MA, Collins DM (eds) Paratuberculosis: Organism, Disease, Control. CAB International, Cambridge, MA, pp 144-156.
15. **Ferreira A, Mariano I, Caetano MC, Núncio P, Carrilho E, Sousa C, Lopes S, Almeida V, Penha-Gonçalves A**. 2002. Epidemiological study of paratuberculosis in ruminants in Alentejo, Portugal. *In*: Juste RA (Ed), Proc. VII Int Coll PTBC. Bilbao.
16. **Correia-Gomes C, Mendonça D, Niza-Ribeiro J**. 2010. Risk associations to milk ELISA result for paratuberculosis in dairy cows in northern Portugal using a multilevel regression model. Revue Méd Vét **161**:295-301.
17. **Ferreira A.** 1989. Patologia dos pequenos ruminantes. Panorama da paratuberculose (doença de Johne), em ovinos e caprinos na região de Trás-os-Montes e Alto Douro. Repositório de trabalhos do L.N.I.V. **21**:71–76 (in Portuguese).

18. **Amado A, Albuquerque T, Ferreira F**. 1994. Paratuberculosis. Epidemiological study in goats and sheep in the Vouzela area of Portugal. In: Chiodini RJ, Collins MT, Bassey EO (Eds), Proc IV Int Coll PTBC. Cambridge, pp34.
19. **Mendes S, Boinas F, Albuquerque T, Fernandes L, Afonso A, Amado A**. 2004. Epidemiological studies on paratuberculosis in small ruminants in Portugal. Epidémiol et santé anim. **45**:61-71.
20. **Coelho AC, Pinto ML, Silva S, Coelho AM, Rodrigues J, Juste RA**. 2007. Seroprevalence of ovine paratuberculosis infection in the Northeast of Portugal. Small Ruminant Res **71**:298-303.
21. **Vala H, Santos C, Esteves F, Albuquerque T, Afonso A, Botelho A, Seixas C, Amaral M, Amado A**. 2007. Paratuberculosis in sheep from Serra da Estrela Region, Portugal. Proceedings of the 9th International Colloquium on Paratuberculosis, Japan. Epidemiology and Control Strategies 250-253.
22. **Coelho AC, Pinto ML, Coelho AM, Rodrigues J, Juste R**. 2008. Estimation of the prevalence of *Mycobacterium avium* subsp. *paratuberculosis* by PCR in sheep blood. Small Ruminant Res **76**:201-206.
23. **Quintas H, Coelho AC, Valentim R, Vila AG, Prendes SM, Maurício R, Mendonça A**. 2012. Serological survey of *Map* infection in goats in the Northeast of Portugal. Proceedings of the XXVII World Buiatrics Congress, Lisbon, Portugal, pp. 315.
24. **Matos AC, Figueira L, Martins MH, Matos M, Andrade S, Alvares S, Mendes A, Sousa N, Coelho A, Pinto ML**. 2013a. Granulomatous lesions and *Mycobacterium avium* subsp. *paratuberculosis* in Portuguese wild boars (*Sus scofa*). Proceedings of the ESVP/ECVP 2012 **148**:1.
25. **Matos AC, Figueira L, Martins MH, Matos M, Pires MA, Alvares S, Mendes A, Sousa N, Coelho A, Pinto M**L. 2013b. Renal lesions in deer (*Cervus elaphus*): involvement of *Mycobacterium avium* subsp. *paratuberculosis*. Proceedings of the ESVP/ECVP 2012 **148**:1.
26. **Matos AC, Figueira L, Martins MH, Matos M, Alvares S, Pinto ML, Coelho AC**. 2013c. Disseminated *Mycobacterium avium* subsp. *paratuberculosis* infection in two wild Eurasian otters (*Lutra lutra L.*) from Portugal. J Zoo Wildl Med **44**:193-195.
27. **Miranda C, Matos M, Pires I, Ribeiro P, Álvares S, Vieira-Pinto M, Coelho AC**. 2011. *Mycobacterium avium* subsp. *paratuberculosis* infection in slaughtered domestic pigs for consumption detected by molecular methods. Food Res Int **44:**3276–3277.
28. **Stevenson K, Alvarez J, Bakker D, Biet F, de Juan L, Denham S, Dimareli Z, Dohmann K, Gerlach GF, Heron I, Kopecna M, May L, Pavlik I, Sharp JM, Thibault VC, Willemsen P, Zadoks RN, Greig A**. 2009. Occurrence of *Mycobacterium avium* subspecies *paratuberculosis* across host species and European countries with evidence for transmission between wildlife and domestic ruminants. BMC Microbiol **9**:212.
29. **Kasnitz N, Köhler H, Weigoldt M, Gerlach GF, Möbius P**. 2013. Stability of genotyping target sequences of *Mycobacterium avium* subsp. *paratuberculosis* upon cultivation on different media, in vitro- and in vivo passage, and natural infection. Vet Microbiol **167**:573-583.
30. **Castellanos E, de Juan L, Domínguez L, Aranaz A**. 2012. Progress in molecular typing of *Mycobacterium avium* subspecies *paratuberculosis*. Res Vet Sci **92**:169-179.
31. **Fritsch I, Luyven G, Köhler H, Lutz W, Möbius P**. 2012. Suspicion of *Mycobacterium avium* subsp. *paratuberculosis* transmission between cattle and wild-living red deer (*Cervus elaphus*) by multitarget genotyping. Appl Environ Microbiol **78**:1132-1139.
32. **Ahlstrom C, Barkema H W, Stevenson K, Zadoks R N, Biek R, Kao R, Trewby H, Haupstein D, Kelton D F, Fecteau G, Labrecque O, Keefe G P, McKenna S L B, Buck J De**. 2015. Limitations of variable number of tandem repeat typing identified through whole genome sequencing of Mycobacterium avium subsp. paratuberculosis on a national and herd level. BMC Genomics **16**:161.

Chapter IV

Effectiveness of nested IS*900*-targeted real time PCR to detect *Mycobacterium avium* subsp. *paratuberculosis* in faeces and milk

1 – "Effectiveness of nested IS900-targeted real time PCR to detect Mycobacterium avium subsp. paratuberculosis in faeces"
Célia Leão, Catarina Cruz, Ana Amaro, Carlos Pinto, Ilda Santos-Sanches, Joyce McLuckie, Craig Watkins, Karen Stevenson, Ana Botelho and João Inácio

2 – "Presence of Mycobacterium avium subs. paratuberculosis DNA in milk used to feed calves in Portugal"
Célia Leão, Ana Botelho, Elisabete Martins, Carla Aguiar, Inês Rebelo, Ricardo Bexiga

Manuscript number 1 and Manuscript number 2 are in preparation.

Effectiveness of nested IS*900*-targeted real time PCR to detect *Mycobacterium avium* subsp. *paratuberculosis* in faeces and milk samples

4.1. Abstract

Mycobacterium avium subsp. *paratuberculosis* (*Map*) is the causative agent of paratuberculosis. We optimized a nested IS*900*-targeted real time PCR assay that combines a first step of conventional PCR followed by a real time PCR, for improving the efficiency of *Map* detection in animal faeces and milk. The PCR assay was validated and tested using faecal samples from goats, cattle and sheep and milk samples from cattle. For faecal samples nested IS*900* PCR showed a diagnostic sensitivity and specificity of 96.6% and 51.1%, respectively, and a kappa coefficient of 0.60 when compared with the culture gold standard. Percentage of *Map* PCR positivity for faecal samples was calculated based on IS*900*-targeted real time PCR results, both with and without previous amplification step (nested and non-nested real time PCR). The nested approach showed an improved performance when compared with the non-nested PCR with a proportion of positives of 83% and 44%, respectively. For milk samples the percentage of *Map* PCR positivity, based on nested IS*900* real time PCR, was 26% representing 48.6% of the evaluated farms. This *ante mortem* test allowing the fast detection of *Map* in faeces and milk from live animals will be a great advantage in improving the efficiency of paratuberculosis monitoring programs.

Keywords

Mycobacterium avium subsp. *paratuberculosis*; Paratuberculosis; Johne's disease; IS*900*; *F57*; nested real time PCR

4.2. Introduction

Mycobacterium avium subsp. *paratuberculosis* (*Map*) is the causative agent of paratuberculosis, or Johne's disease, a chronic granulomatous enteritis infecting a wide range of animals, causing disease especially in ruminants, camelids, rabbits and hares [1; 2; 3; 4; 5]. *Map* infection has also become a highly contentious issue as a possible contributory factor of the human inflammatory bowel disease known as Crohn's disease [1; 6; 7]. Humans can be exposed to the agent mainly via the food chain where *Map* has been found in milk, milk products and meat [6; 8; 9]. Distinct phenotypic and genotypic *Map* strains have been isolated from different host species; Type C (Type II), associated with multiple host's infections, predominantly cattle; Type S (Type I and III), associated with primarily sheep and goats, and Bison Type associated with buffalo, cattle, goats, humans and other species infections [10].

During the past decade, paratuberculosis was considered as one of the most important diseases for worldwide livestock industries, due to its considerable economic impact triggered by a progressive and fatal weight loss of the animals and a diminution of milk production [6; 11; 12]. The transmission of *Map* usually occurs during the animals' first months of life, by faecal-oral route, with calves under 6 months of age being the most susceptible to infection due to their immature immune systems [13]. Clinical signs of disease usually appear after a long period of incubation, between two to five years [2; 6]. The definitive diagnosis of paratuberculosis is difficult and time consuming. Different and complementary diagnostic approaches can be used, including: (i) anatomo-histopathological examination, with microscopic observation of lesions and acid-fast bacilli in tissues; (ii) immunodiagnostics, with the analysis of the animal's immune response; (iii) bacteriological diagnostics, involving the culture and isolation of *Map* from biological samples such as faeces and tissues using specific media supplemented with mycobactin; and (iv) molecular diagnostics, based on the detection of *Map*'s specific nucleic acid sequences with the multi-copy IS*900* and the single-copy *F57* elements being the most used *Map*-specific genomic targets [1; 2; 14; 15; 16]. All these diagnostic approaches have advantages and disadvantages but the gold standard for *ante mortem* diagnostic, still remains the culture and identification of viable *Map* cells, which require several months to obtain due to the extremely fastidious growth of the agent.

In spite of the worldwide importance of paratuberculosis for the livestock industry, only a few nucleic acid testing assays, particularly the ones based on real time PCR technologies, have been validated and routinely used in veterinary laboratories for the *ante mortem* detection of *Map* in animal faeces [11; 17; 18; 19; 20; 21; 22; 23], milk [8; 9; 24; 25; 26] and blood [27]. Furthermore, an important limitation for the molecular detection of *Map* in faeces and milk samples is related to the inefficiency of mycobacterial DNA extraction procedures from those matrices [23; 24; 28; 29; 30; 31].

In a previous report from Sidoti and colleagues (2011) [32], an IS*900*-targeted hydrolysis probe, and respective flanking primers, was fully optimized and validated for the specific detection of *Map* in human biopsy specimens using real time PCR approaches. However, despite its high analytical sensitivity and specificity, this system was not fully satisfactory in our reference veterinary laboratory when used for the direct detection of *Map* in animal faeces and milk.

In this work we aim to improve the efficiency of *Map* detection in faecal and milk samples by optimizing a procedure for faecal and milk DNA extraction and by developing a nested IS*900*-targeted real time PCR assay that combines a first step of conventional PCR followed by a real time PCR. A different Map-specific target for real time PCR assay*, F57,*that has not been isolated from other organisms, unlike IS*900*, was also optimized for the identification of *Map* isolates to confirm identity of the isolates.

4.3. Materials and Methods

4.3.1. Bacterial strains

Nineteen reference, clinical and environmental strains of *Mycobacterium avium* Complex (MAC), non-MAC mycobacteria and non-mycobacterial species maintained at the Instituto Nacional de Investigação Agrária e Veterinária (INIAV, IP), Portugal, were used for the optimisation of the real time PCR-based assays, namely for the assessment of the analytical specificities and sensitivities of the respective primers and probes (Table 4.1).

Table 4.1. Bacterial strains used for the determination of the analytical specificity of the real time PCR assay and respective results.

Species	Reference strains/Isolates	IS*900*[1]	F57[1]
Acinetobacter baumanni	INIAV 845	-	-
Aeromonas hidrofila	INIAV 29172/12	-	-
Enterococcus	INIAV 27757/12	-	-
Escherichia coli	INIAV 17591/12	-	-
Klebsiella oxytoca	INIAV 27778/12	-	-
Klebsiella pneumonia	INIAV 26548/12	-	-
Pseudomonas aeruginosa	INIAV 838	-	-
Pseudomonas putida	INIAV 832	-	-
Salmonella	EURL 51	-	-
Staphilococcus intermedius	INIAV 831	-	-
Streptococcus bovis 1	INIAV 837	-	-
Mycobacterium avium subsp. *hominissuis*	INIAV 19	-	-
Mycobacterium avium subsp. *avium*	ATCC^{T} 25291	-	-
Mycobacterium avium subsp. *paratuberculosis*	INIAV 1568; ATCC^{T} 19698; INIAV 3	+	+
Mycobacterium bovis BCG	ATCC 27291	-	-
Mycobacterium tuberculosis	ATCC 25177	-	-
Mycobacterium scrofulaceum	INIAV 31389	-	-

ATCC – American Type Culture Collection, USA; INIAV – Instituto Nacional de Investigação Agrária e Veterinária, Lisbon, Portugal; [1]Presence (+) or absence (-) of IS*900* and F57 in the genome.

4.3.2. Faecal samples

Four sets of faecal samples were used in this work (Table 4.2). Set A was collected on the farm and Sets B and C were collected at the official abattoir, sent to the reference laboratory INIAV, IP, and stored at -20ºC until being processed for *Map* culture and/or DNA extraction. Samples for Set D were stored and processed for DNA extraction at the Moredun Research Institute, Scotland in collaboration with the University of Edinburgh, and DNA sent to INIAV, IP, for PCR analysis.

4.3.3. Milk samples

Ninety nine milk samples were collected from 37 dairy farms from 16 Portuguese counties (Table 4.2, Set E). Each milk sample, from 33 farms, was composed of a pool of waste milk used to feed calves. From the remaining four farms bulk tank milk was collected. Milk samples were collected on three different days on each farm, separated at least one week between collections, to increase the likelihood that the source of the milk was from different animals, and stored at -20ºC until being processed.

Table 4.2. Faecal and milk samples used in this study.

Set	Sample type	N. samples	Animal species	Geographic region	Sample collection	Paratuberculosis evidence
A	Faeces	17	Caprine	Azores, Portugal	Single farm	Positive clinical signs, 16 samples from seropositive animals and one pool from seronegative animals
B	Faeces	58	Bovine	Azores, Portugal	Official abattoir	Positive serology and histopathology
C	Faeces	40	Bovine	Azores, Portugal	Official abattoir	Some animals with clinical signs, serology and histopathological evidences
D	Faeces	66	Ovine	Scotland	Multiple farms	Suspicious animals from farms with history of paratuberculosis
E	Milk	99	Bovine	Portugal Mainland	Multiple farms	Some farms with history of paratuberculosis

4.3.4. Spiked faecal samples

Faecal samples were pooled from cattle that tested negative for paratuberculosis by traditional culture, immunological and PCR tests, and showed no histopathology consistent with paratuberculosis following slaughter. Six grams of pooled faeces was spiked with ten-fold dilutions of a suspension of *Map* cells, either with the *Map* K10 (Type C) or *Map* 235G (Type S) strains, in a range of 10^4 to 10^1 cells per gram of faeces, tested in triplicate. An additional 6 g of faeces without *Map* cells was used as a negative control. The number of *Map* cells was estimated by microscope count. The tubes were mixed well and stored at -20ºC until required. The genomic DNA of the faecal spiked samples was extracted and tested with the same procedures as the faecal samples.

4.3.5. Spiked milk samples

Milk spiked samples were split into nine tubes each with 10 ml of a bovine milk sample tested negative by standard procedures for paratuberculosis. The milk was spiked with ten-fold dilutions of a *Map* ATCC19698^T suspension in a range of 10^7 to 0 cells per ml of milk. Tubes were mixed well and stored at -20ºC until being tested. The genomic DNA of the milk spiked samples was extracted and tested with the same procedures as the milk samples.

4.3.6. Bacteriological culture of faeces

All faecal samples of sets A, B and C (Table 4.2) were tested for the presence of *Map* using culture assays according to the OIE (2014) [16], with minor modifications. Briefly, 20 mL of sterile distilled water were added to 1 g of faeces and stirred at room temperature for 30 minutes. After settling for 30 minutes, 5 mL of the uppermost suspension were transferred to a new tube containing 20 mL of 0.9% hexadecylpyridinium chloride (HPC), inverted several times and allowed to stand undisturbed at room temperature for 18 hours. After this period, the sediment was carefully transferred to a new tube, washed with 10 mL of sterile distilled water and centrifuged at 900 ×g at room temperature for 30 minutes. The pellet was resuspended in 500 µL of sterile distilled water and volumes of 100 µL were inoculated on Herrold's egg yolk medium (HEYM) slants with and without mycobactin J (bovine and caprine samples) and on Löwenstein–Jensen (LJ) medium with mycobactin J (caprine samples). The incubation was performed at 37 ºC for up to 6 months. Isolated colonies were confirmed to be acid-fast bacilli by auramine-rhodamine staining as described by Bird *et al.* (2000) [33].

4.3.7. Bacteriological culture of milk

Milk samples (Table 4.2, set E) were prepared for culture according to Dimareli-Malli (2010) [34], with minor modifications. Briefly, 20 mL of milk were centrifuged at 2000 × g for 15 minutes at room temperature, the supernatant was discarded and the pellet was resuspended in 10 mL of 0.75% HPC and incubated at room temperature for 5 hours. The mixture was centrifuged at 2000 × g for 15 minutes at room temperature, the pellet was resuspended in 2 mL of sterile distilled water and 200 µL were inoculated on HEYM slants with and without Mycobactin J. The incubation was performed at 37 ºC for up to 6 months.

4.3.8. DNA extraction from faeces

All the caprine and bovine samples from the Azores (sample sets A, B and C), as well as the spiked samples, were processed for DNA extraction at INIAV, IP. The approach was based on the commercially available Invisorb® Spin Tissue Mini Kit (Stratec), but including previous additional steps to concentrate *Map* within the faecal samples and to mechanically disrupt the cells. For the concentration, 1 g of sample was stirred at room temperature for 30 minutes with 20 mL of sterile distilled water and allowed to settle for 30 minutes (the same homogenate was used also for culture – see 4.3.6). Five millilitres of the uppermost suspension were transferred to a new tube, centrifuged for 20 minutes at 3800 × g at room temperature and 4 mL of the supernatant was discarded. The pellet was resuspended in the remaining volume and 250 µL of the suspension was transferred to a sterile tube for mechanical disruption of the cells. Zirconium beads (1 mm) and 400 µl of the kit lysis buffer were added and the cells disrupted in a FastPrep FP120 Bio101 bead shaker (Savant Instruments Inc., Holbrook, NY) at 6.5 $msec^{-1}$ for 45 seconds, twice. Disrupted samples were cooled on ice for 15 minutes, followed by the addition of 50 µL of kit proteinase K solution and incubation overnight at 52 ºC. The remaining procedure was performed with the DNA extraction kit, according to the manufacturer’s instructions. The genomic DNA was eluted with 100 µL of elution buffer and stored at - 20ºC until being tested. The ovine faecal samples from Scotland (set D) were processed at the Moredun Research Institute, Scotland (in collaboration with the University of Edinburgh), with the PowerFecal® DNA isolation kit (MO BIO laboratoires, Inc.) according to the manufacturer’s instructions, and sent to INIAV, IP for analysis.

4.3.9. DNA extraction from milk

All the milk samples were submitted to a first treatment procedure, as described by Gao and colleagues (2007) [29] with minor modifications. Briefly, 10 mL of milk were incubated at 95 ºC during 10 minutes and cooled on ice for 10 minutes. Samples were centrifuged at 3100 × g for 30

minutes at 8 ºC and the whey was carefully removed. The pellet and the fat layer were resuspended in 15 mL of 0.75% HPC and incubated at room temperature for 30 minutes under agitation. After the incubation step, samples were centrifuged at 2000×g for 15 minutes at room temperature, the fat layer and liquid phase were decanted and the pellet was used for DNA extraction with the Invisorb® Spin Tissue Mini Kit (Stratec) with an additional mechanical disruption step.

The obtained pellet was resuspended in the kit lysis buffer, and transferred to a sterile tube for mechanical disruption of the cells with Zirconium beads (1 mm). Mechanical disruption was performed twice in a FastPrep FP120 Bio101 bead shaker (Savant Instruments Inc., Holbrook, NY) at 6.5 $msec^{-1}$ for 45 seconds. Disrupted samples were cooled on ice for 15 minutes, followed by the addition of 40 µL of proteinase K and incubation overnight at 52ºC. The remaining procedure was performed with the DNA extraction kit, according to the manufacturer's instructions. Genomic DNA was eluted with 100 µL of elution buffer and stored at - 20ºC until tested.

4.3.10. *TaqMan* probes and primers

DNA sequences from *Map* IS*900* and *F57* specific regions were retrieved from NCBI-GenBank and analysed for designing novel probe and primers. For the IS*900*, additional external forward (EXT-IS*900*-FW) and reverse (EXT-IS*900*-RV) primers were designed, flanking the primers and probe previously described by Sidoti and colleagues (2011) [32] (Figure 4.1 and Table 4.3). To set a nested IS*900*-targeted real time PCR assay the external primers were used in first amplification step by conventional PCR whose amplified products are used in the second amplification step by qPCR. A novel *TaqMan* probe and the respective flanking primers were also designed targeting the *F57* region (Table 4.3). All primers and probes were tested for their *in silico* specificity using the BLASTn from NCBI-GenBank (http://blast.ncbi.nlm.nih.gov). All probes and primers were synthesized by MWG Biotech (Germany).

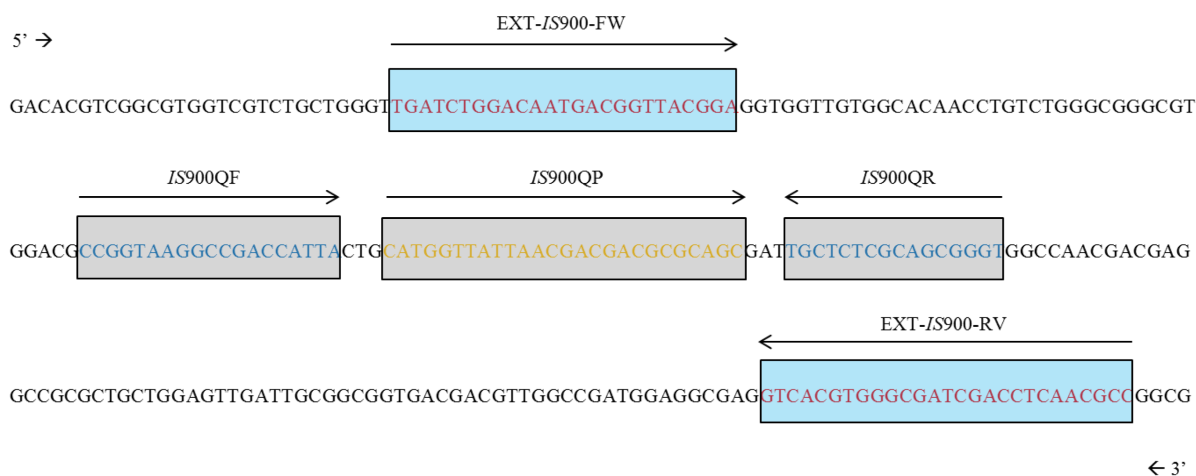

Figure 4.1. Schematic representation of complementary *Map* IS*900*-targeted primers and probe used in this study.

Boxed in blue are the newly designed external primers used for the first amplification step of the nested IS*900* real time PCR. Boxed in grey are the primers and probe described by Sidoti and colleagues (2011) [32]. Accession number of IS*900* sequence in Genebank is AF416985.1.

Table 4.3. Primers and probes used in this study.

Type of PCR	Primer/Probe	Sequence (5’-3’)	Target/References
Nested **IS*900*** qPCR - 1st step conventional PCR	EXT-IS*900*-FW	TGA TCT GGA CAA TGA CGG TTA CGG A	IS*900* element of *Map*/ this study
	EXT-IS*900*-RV	GGC GTT GAG GTC GAT CGC CCA CGT GAC	
Nested **IS*900*** qPCR - 2nd step qPCR and **IS*900*** qPCR	IS*900*QF	CCG GTA AGG CCG ACC ATT A	IS*900* element of *Map*/ [32]
	IS*900*QR	ACC CGC TGC GAG AGC A	
	IS*900*QP[1]	TET - CAT GGT TAT TAA CGA CGA CGC GCA GC - BHQ1	
F57 qPCR	F57_F	GCA GCT CCA GAT CGT CAT TC	*F57* element of *Map*/ this study
	F57_Rb	GTC CAG TTC GCT GTC ATC GA	
	TqF57b[2]	FAM - AGC ACG CAG GCA TTC CAA GTC C - BHQ1	
β-actin qPCR	F_Actin	GGC TCY ATY CTG GCC TC	*β-actin* gene of mammals/ [35]
	R_Actin	GCA YTT GCG GTG SAC RAT G	
	P_Actin[3]	Cy5.5 - TAC TCC TGC TTG CTG ATC CAC ATC - BHQ2	

[1] Probe labelled with tetrachlorofluorescein (TET) and Black Hole Quencher-1 (BHQ-1); [2]Probe labelled with 6-carboxyfluorescein (FAM) and Black Hole Quencher-1 (BHQ-1); [3]Probe labelled with Cyanine 5.5 fluorophore (Cy 5.5) and Black Hole Quencher-2 (BHQ-2).

4.3.11. Identification of *Map* isolates using *F57*-targeted PCR

Isolates were identified as *Map* by an *F57*-targeted TaqMan-based real time PCR assay. The analytical specificity of the assay was determined by testing extracted genomic DNA from MAC, non-MAC mycobacteria and non-mycobacteria isolates. The analytical sensitivity was determined by the construction of a standard curve based on the analysis of 10-fold serial dilutions of *Map* ATCC 19698^T DNA, extracted from the pure culture.

DNA was extracted from cultures by a heat treatment. For this purpose, one colony was transferred to a microtube containing 100 µL of TE buffer, mixed in a vortex and centrifuged at 2000 × g for 5 minutes at room temperature. The pellet was resuspended in 100 µL of Tris EDTA buffer pH 8 (10 mM Tris.Cl, 1 mM EDTA) and incubated for 45 minutes at 95 ºC. After bacterial lysis the suspension was centrifuged at 2000 × g for 1 minute at room temperature and the supernatant containing the extracted DNA was transferred to a new microtube and used directly as template for PCR reactions. For the confirmation of *Map* isolates, real time PCR reactions were carried out in a total volume of 20 µl containing 1× SSO Fast Super Mix (Bio-Rad), 0.4 µM of each *F57*-targeted primer and 0.15 µM of probe (Table 4.3) and 5 µl of the extracted DNA template. DNase free water (GIBCO) was used as negative control. Thermal cycling, fluorescent data collection, and data analysis were performed in a CFX96 (Bio-Rad) detection system real time PCR instrument with the following conditions: 1 cycle at 95 °C for 2 minutes, followed by 45 cycles at 95 °C for 5 seconds and 60 °C for 10 seconds.

4.3.12. Detection of *Map* in samples using IS*900*-targeted real time PCR

All set A and set B samples from Azores (Table 4.2), tested for the presence of *Map* using the gold standard of culture, were also screened with the IS*900* real time PCR, with and without a previous nested step, in order to estimate the diagnostic sensitivity and specificity of both molecular detection approaches compared to the gold standard. For the nested approach, the first step consisted of a conventional PCR using the external primers EXT-IS*900*-FW and EXT-IS*900*-RV (Figure 4.1). The reaction was carried out in a final volume of 25 µl containing 200 µM of each deoxynucleotide triphosphate (Applied Biosystems), 2.0 mM of $MgCl_2$ (Life Technologies), 0.4 µM of each primer (Table 4.3), 1 U of *Taq* DNA polymerase and 1× of the respective buffer (Life Technologies), and 5 µl of the extracted DNA solution. The amplification was performed in a MJmini™ Thermocycler (BioRad) with an initial step at 94 ºC for 3 minutes, followed by 40 cycles at 94 ºC for 45 seconds, 55

ºC for 30 seconds and 72 ºC for 90 seconds, ending with a step at 72 ºC for 10 minutes. DNA from *Map* (ATCC 19698^T) and DNase free water (GIBCO) were used as positive and negative controls of amplification, respectively. The second step consisted of a real-time PCR including the IS*900*-targeted probe/primers described by Sidoti and colleagues (2011) [32], and additional *β-actin* gene-targeted probe/primers as internal control [35] (Table 4.3). The real-time PCR reaction was carried out in a total volume of 20 µl containing 1× SSO Fast Super Mix (Bio-Rad), 0.4 µM of each primer (Table 4.3), 0.15 µM of each probe (Table 4.3) and 5 µl of the previously amplified PCR products. DNase free water (GIBCO) was used as negative control. Thermal cycling, fluorescent data collection, and data analysis were performed as above. Each sample was tested in triplicate and considered positive if at least one of the triplicates had a positive threshold cycle (Ct) value below 40.

Similarly to the above mentioned set A and set B samples, the additional 66 samples from Scotland (set D) were also tested both by the nested IS*900* real time PCR and by using only the second step of the assay (i.e., without the first nested step). The remaining 40 bovine faecal samples from Azores (set C), collected from animals with clinical suspicion of paratuberculosis, and the 99 bovine milk samples, were later screened only by the nested IS*900* PCR approach and culture.

The limit of detection (LOD) for the nested real time PCR was assessed using faecal and milk *Map* spiked samples. Each sample was tested in triplicate and the LOD corresponds to the highest dilution at which the assay could detect a positive result in at least one of the replicates.

Sensitivity, specificity and kappa coefficient were computed using the public available clinical research calculators VassarStats website (http://vassarstats.net).

4.4. Results

4.4.1. Analytical specificity and sensitivity of probes and primers

The probe and respective flanking primers targeting IS*900* were previously validated [32]. A novel set of primers and *TaqMan* probe targeting the single copy, *Map*-specific *F57* sequence was designed, showing 100% sensitivity and specificity by targeting only *Map* strains. None of the remaining MAC members, non-MAC mycobacteria or non-mycobacteria yielded any positive amplification (Table 4.1). The *F57*-targeted real time PCR LOD was one genome copy in the reaction mixture.

4.4.2. Culture and identification of *Map* from faecal and milk samples

Seventeen caprine (set A) and 58 bovine (set B) faecal samples from Azores were cultured (Table 4.4). From the 17 caprine samples, 16 were culture positive after two to four months of incubation. From these 16 culture positive samples, 15 were from seropositive animals and one sample was a pool of faeces from seronegative animals. From the 58 bovine samples, 43 showed growth of *Map* colonies after two to four months of incubation and 13 were considered to be culture negative after six months of incubation (Table 4.4). Culture results were not obtained for two samples probably because of the overgrowth of contaminants (Table 4.4).

Culture tests were also performed for the additional 40 bovine samples of set C but only four samples yielded a positive result after six months of incubation.

Ninety nine milk samples (Set E) from 37 dairy farms were cultured but none of the samples showed positive result after six months of incubation and 18 samples presented growth of contaminants (Table 4.4).

All cultures were confirmed to be acid-fast bacilli by auramine-rhodamine staining and were identified as *Map* by the *F57*-targeted real-time PCR assay.

Table 4.4. Results for the *Map* detection in faecal and milk samples by culture and real time PCR assays

Animal/Sample/Origin	**Numbers of samples**	**Culture**[1]	**IS*900* qPCR**[2]	**Nested IS*900* qPCR**[2]
Caprine/faeces/Azores – Portugal (set A)	1	-	-	+
	2	+	-	-
	11	+	-	+
	3	+	+	+
Bovine/faeces/ Azores – Portugal (set B)	5	-	-	-
	1	-	-	+
	3	-	+	-
	4	-	+	+
	2	+	-	+
	41	+	+	+
	2	*	+	+
Bovine/faeces/ Azores – Portugal (set C)	15	-	**	+
	21	-	**	-
	4	+	**	+
Ovine/faeces/Scotland (set D)	14	**	-	-
	43	**	-	+
	9	**	+	+
Bovine/milk /Portugal Continental (Set E)	6	*	**	+
	20	-	**	+
	73	-	**	-

[1]Presence (+) or absence (-) of colonies on specific medium confirmed as *Map* by *F57*-targeted real time PCR; [2]Presence (+) of absence (-) of IS*900*-specific fluorescence amplification curves (considered positive if at least one of three replicates had a Ct < 40);*No culture result due to contamination; **Not performed

4.4.3. Detection of *Map* by IS*900* real time PCR

The above mentioned 17 caprine (set A) and 58 bovine (set B) samples were further tested by IS*900* real time PCR (IS*900* qPCR Tables 4.3 and 4.4) and with a previous conventional amplification step (nested IS*900* qPCR, Table 4.3 and 4.4). From the caprine samples, 14 were positive for both culture and nested IS*900* PCR, including the pool of faeces from sera negative animals (Table 4.4). Two samples were nested IS*900* PCR negative but culture positive, and one sample was nested IS*900* PCR positive but culture negative. Only three caprine samples were positive in IS*900* real time PCR

(non-nested IS*900* qPCR). Regarding the 58 bovine samples of set B, 43 were culture and nested IS*900* PCR positive, and eight samples were culture and nested IS*900* PCR negative (Table 4.4). Of these eight, three were only positive in IS*900* qPCR, but negative in the nested IS*900* qPCR, which constitutes a strange result. Additionally, five samples showed discrepant results between the culture and the nested IS*900* PCR, being positive in PCR and negative in culture (Table 4.4) and two samples, positive in both PCR, were impossible to evaluate by culture due to contamination with other bacteria. Overall for these caprine and bovine samples, 87% were positive when using the nested IS*900* PCR approach, while only 71% of the samples were positive when using the non-nested IS*900* qPCR. Using culture as gold standard when testing these set A and set B samples, the diagnostic sensitivity and specificity of the non-nested IS*900* real time PCR approach was 74.6% and 50.0%, respectively, with a kappa coefficient of 0.20 (poor strength of agreement). The nested IS*900* PCR showed a diagnostic sensitivity and specificity of 96.6% and 57.1%, respectively, and a kappa coefficient of 0.60 (moderate strength of agreement) when compared with the gold standard of culture.

Spiked faecal samples were used to assess the LOD of the nested IS*900* real time PCR, comprising serial dilutions of *Map* strains of both Type C and Type S. For *Map* Type C it was possible to detect positive amplification results at an infection rate of 10^1 cells per gram of faeces, while for Type S the LOD was 10^2 cells per gram of faeces.

The additional 40 bovine samples from Azores (set C) were tested only by the nested IS*900* PCR assay, of which 19 (47.5%) were positive for *Map*. Similarly, from the additional 66 ovine samples from Scotland (set D), tested with the same nested PCR assay, 52 (78.8%) yielded a positive result for the presence of *Map*. For set D, the non-nested IS*900* real time PCR only detected 13.6% of *Map* positive samples (and all these positive samples were also positive for the nested approach).

From the set E, 26 milk samples (26%), representing 48.6% out of the 37 evaluated farms were positive for nested IS*900* real time PCR, including one of the four farms where bulk tank milk was collected. Twenty milk samples were positive for nested IS*900* real time PCR and negative for culture, while 73 were negative for both methods (Table 4.4).

Spiked milk samples were used to evaluate the LOD of the nested IS*900* real time PCR comprising serial dilutions of *Map* Type C resulting in detection of positive amplification at an infection rate of 10^2 cells per millilitre of milk.

In order to assess the presence of potential PCR inhibitors the co-amplification of the β-actin gene was used as an internal control of amplification in the real time PCR reaction. The amplification of this target was observed for all samples tested.

4.5. Discussion

The gold standard for the definitive *ante mortem* diagnosis of paratuberculosis is the isolation and identification of *Map* in specific medium. The principal advantage of culture is its specificity and possibility of quantification of *Map* expressed as colony forming units (CFU) per unit of sample (e.g. grams of faeces). This allows the classification of infected animals according to the excretion level of *Map* in faeces as: low (1 - 10 colonies), moderate (10 - 100 colonies) or high (> 100 colonies) shedders [15; 23]. Our faecal culture results showed that infection rates of set A and set B samples were 94% and 77%, respectively. Among culture positive samples, about 40% were considered to be from low *Map* shedders and the remaining from moderate or high shedders.

Map culture procedures require a variable incubation period due to the selected culture medium, number of bacteria present in the sample and to the *Map* strain Type, with Type C requiring about 2-3 months and Type S taking more than 6 months for colonies to first be observed [12; 15; 23]. Therefore, molecular diagnostic tools, particularly based on real time PCR technologies, have been proposed to shorten this *Map* time-to-detection.

Aiming to implement an efficient and rapid approach to detect *Map* in biological samples, we optimized a real time PCR assay using a previously published IS*900*-targeted hydrolysis (*TaqMan*) probe, and respective flanking primers. These assay and probe/primers were fully validated by Sidoti and colleagues (2011) [32] for use with human biopsy samples. However, faeces and milk are challenging biological matrices for the molecular detection with sample preparation and DNA extraction considered as critical steps where a high DNA quality is required [2; 8]. Faeces need an efficient extraction method due to the presence of PCR inhibitors such as phytic acid and polysaccharides, and to the large amounts of nucleic acids from other bacteria and host cells [23, 30]. Milk is considered a very difficult matrix due to the large quantity of fat and calcium ion [24]. Because of the characteristic thick waxy and lipid-rich *Map* cells, these bacteria are preferentially located in the fat fraction of milk, which is why many studies have been performed to investigate the best methodology for milk processing, including aspects like sample volume, sample concentration or inclusion of cell disruption and digestion steps [8; 9; 29].

We tested several DNA extraction approaches to process faecal and milk samples, including the use of different commercial DNA extraction kits and preliminary steps for *Map* cell concentration and lysis (data not shown). The most efficient extraction method, detecting a higher number of PCR positive samples, is described above in the materials and methods section and involves the use of the Invisorb® Spin Tissue Mini Kit with mechanical (bead beating) disruption, commonly considered as an essential step for DNA extraction from *Map* cells [23; 30]. However, for faecal samples we used more quantity of sample (1 g instead of 25 mg) to increase the number of *Map* cells and for both faecal

and milk samples the enzymatic lysis incubation time was performed in 12h in order to maximize the amount and quality of extracted DNA and improve detection by PCR.

The amplification results were not satisfactory when using the optimised DNA extraction procedure associated with the single IS*900*-targeted real time PCR approach to test both set A (caprine) and set B (bovine) faecal samples, when compared with the culture gold standard, only detecting 74.6% positive results among all culture positive samples. Therefore, a nested real time PCR approach was attempted by the inclusion of an initial conventional PCR step resulting in an increase in the diagnostic sensitivity to 96.6% (considering set A and set B faecal samples). Overall, considering the PCR positivity for all faecal samples tested by both the non-nested and nested real time PCR approaches (sets A, B and D), the proportion of positives was 44% and 83%, respectively, with a better performance for the nested approach, particularly for sets A and D.

Two *Map* faecal culture positive samples were negative by the nested IS*900* PCR, which lowered the diagnostic sensitivity. In these cases, the presence of inhibitors hampering the amplification reaction was ruled out due to the co-amplification of the control mammalian β-actin gene from the same samples. However, it is known *Map* cells may form aggregates and are not homogeneously distributed in the samples, and due to the procedures for culture and DNA extraction, we could potentially have more *Map* cells for culture than correspondent *Map* DNA to PCR once we only use 5 µL of the DNA suspension for *Map* detection, which may explain discrepant results between tests using separate faeces subsamples. There is also the possibility of unknown polymorphisms within the IS*900* target sequence preventing hybridization with the *Map*-specific probe. However, the IS*900* seems to be highly conserved among different strains, with a small number of nucleotide polymorphisms [36].

The detection of three faecal samples where amplification results were detected using the non-nested real time PCR approach but not the nested approach is more difficult to explain and deserves further studies (Table 4.4). As mentioned above, the complementary regions of the IS*900*-targeted external primers might present polymorphisms for a subset of *Map* strains, hampering the PCR amplification. Regardless of these discrepancies, the kappa measure of agreement between *Map* faecal culture and the nested IS*900* PCR was estimated to be 0.60, when comparing the results for faecal samples from sets A and B (animals with suspicion of paratuberculosis). Although the criteria for judging kappa statistic are not completely objective nor universally accepted, this value may allow us to infer a "moderate strength" of agreement between the two detection methods [37].

In total, 21 faecal and 93 milk samples were culture negative but yielded positive PCR results. This suggests nested IS*900* PCR to be more sensitive than culture for *Map* detection, since occurrence of PCR false positives was ruled out by use of effective negative controls. Furthermore storing conditions of freezing samples and the chemical processing before culture may adversely affect *Map* cells viability leading to their non-recovery from samples [38; 39], namely when using HPC decontamination as in the current work. However, PCR can still detect nucleic acids from these non-

viable or dormant *Map* cells. Also, animals may be shedding *Map* cells in numbers below the threshold of detection by culture, but still detected by PCR.

The difficulty of isolation of *Map* from milk samples despite the positive results on molecular detection obtained in our study is in accordance to Hanifian and colleagues (2013) [11] that reported ten times higher rates of positive results by real time PCR detection in comparison to culture.

The probability of a positive culture result in milk samples depends on the viability of the bacteria, the animal's infection load, the quality of the milk sample and the sample volume that is used in the analysis. It has been described that the volume that should be used for testing should be as high as possible, from 1 to 250 mL per sample [9]. In this work we used only 20 mL of milk for culture and 10 mL of milk for DNA extraction, due to limited availability of higher sample volumes. Other possible explanations for the culture negative results from milk samples are the composition of waste milk, a mixture of milk from animals with other infections (e.g. mastitis) and from animals under treatment; contamination rate of the cultures and the presence of antimicrobials could also restrict the growth of a fastidious agent like *Map* and the fact that *Map* is naturally concentrated at the fat layer of milk and the decontamination step with HPC could confine the recovery and viability of the agent.

Other previous works also describe that more samples scored positive for *Map* when using real time PCR approaches when compared with culture [11; 18; 19; 20; 21; 22; 40].

A previous report was published describing a nested IS*900*-targeted real time PCR approach using SYBR green chemistry to detect *Map* in animal faecal samples [18]. The authors found that the majority of the nested real time PCR products corresponded to unspecific amplification artefacts, not producing the fragment with the correct size. According to these authors, nested products should be confirmed for their correct length. In our nested real time PCR assay, which uses *TaqMan* chemistry, this issue is overcome by the use of a *Map* specific hydrolysis probe.

The enhancement of the *Map* detection rates using the nested IS*900* assay need to be balanced against the associated increased risk of cross contamination of samples. Therefore, we should emphasize the need for working in a diagnostic laboratory observing good laboratory practice for molecular PCR analysis, which includes working in separate clean rooms for DNA preparation and PCR analyses and the use of positive and negative controls [35].

The IS*900* target was selected for our PCR assay since it's a multi-copy sequence in the *Map* genome, which increases the chance of detecting the agent in the samples. As an example, we compared the performance of the nested IS*900* real time PCR with a similar nested assay targeting the single-copy *F57* element, using DNA extracts from set D ovine faecal samples (data not shown). As expected, of the 66 faecal samples, the IS*900* assay detected 79% positives, compared with only 21% positives for the *F57*-targeted assay. However, some reports describe the occurrence of IS*900*-like sequences in non-*Map* mycobacteria [41; 42]. This could raise some concerns about the specificity of IS*900*-targeted molecular assays for detecting *Map*. Nonetheless, as far as we know, the occurrence of these IS*900*-like elements among mycobacteria seems to be very rare [4], with only very few

sequences disclosed in public databases such as GenBank-NCBI. We believe the rare occurrence of these elements does not pose any significant specificity concerns when using IS*900*-targeted assays for the detection of *Map*.

The nested IS*900*-targeted PCR assay described above, associated with improved sample preparation to concentrate the agent and to reduce PCR inhibitors, allows the detection of *Map* from animal faeces and milk, with high sensitivity and specificity, reducing the time for confirmatory *Map* diagnosis from several months to a few hours. The assay is amenable to future automation possibilities regarding both the DNA extraction and amplification steps, particularly when used in reference veterinary laboratories. The availability of *ante mortem* tests allowing a fast and conclusive detection of *Map* in biological samples from live animals will be a great advantage in improving the efficiency of paratuberculosis monitoring programs and in decreasing the associated economic burden.

Acknowledgements

This study was funded by the project PTDC/CVT/111634/2009 from the *Fundação para a Ciência e a Tecnologia* (FCT), Portugal and by the Scottish Government Rural and Environment Science and Analytical Services Division in Scotland. Célia Leão is a recipient of PhD grant from FCT (SFRH/BD/62469/2009). We kindly acknowledge the work done by Jelena Jelena Nikolić, in Dr Andrew Free's research group at the School of Biological Sciences, University of Edinburgh for some of the DNA extractions from faecal samples in Group D. Lurdes Clemente (INIAV, I.P.) is acknowledged for providing bacterial strains.

Author's contribution

CL contributed to the designing of the study, experimental work of samples preparation, culture, growing and characterization of isolates, DNA extraction, real time PCR development, data analysis and writing the manuscripts. CC collaborated with collection and preparation of faecal samples, faecal culture, faecal DNA extraction, and faecal real time PCR testing. AA contributed for the laboratory work and revised the manuscript. CP contributed to the faecal samples collection and revised the manuscript. ISS contributed to the designing of the study and revised the manuscript. JM contributed with the spiked faecal samples. CW contributed with DNA from sheep faecal samples and revised the manuscript. KS contributed with the spiked faecal samples and revised the manuscript. AB contributed to the designing of both studies and revised both manuscripts. JI contributed to the designing of the faecal study and revised the manuscript. EM, CA, IR and RB contributed with collection of milk samples, all the material for the experiment, contributed to the designing of the study and revised the milk manuscript. All authors have read and approved the final manuscript.

4.6. References

1. **Timms VJ, Gehringer MM, Mitchell HM, Daskalopoulos G, Neilan B A**. 2011. How accurately can we detect *Mycobacterium avium* subsp. *paratuberculosis* infection? J Microbiol Meth **85**:1-8.
2. **Salem M, Heydel C, El-Sayed A, Ahmed SA, Zschöck M, Baljer G**. 2012. *Mycobacterium avium* subspecies *paratuberculosis*: an insidious problem for the ruminant industry. Trop. Anim Health Prod **45**:351-366.
3. **Carta T, Álvarez J, Pérez de la Lastra JM, Gortázar C**. 2013. Wildlife and paratuberculosis: A review. Res Vet Sci **94**:191-197.
4. **Rindi L, Garzelli C**. 2014. Genetic diversity and phylogeny of *Mycobacterium avium*. Infec Genet Evol **21**:375-383.
5. **Salgado M, Monti G, Sevilla I, Manning E**. 2014 Association between cattle herd *Mycobacterium avium* subsp. *paratuberculosis* (MAP) infection and infection of a hare population. Trop Anim Health Prod **46**:1313–1316.
6. **Atreya R, Bülte M, Gerlach GF, Goethe R, Hornef MH, Köhler H, Meens J, Möbius P, Roeb E, Weiss S**, on behalf of the ZooMAP Consortium. 2014. Facts, myths and hypotheses on zoonotic nature of *Mycobacterium avium* subspecies *paratuberculosis*. Int J Med Microbiol **304**:858-867.
7. **Park KT, Allen A J, Davis WC**. 2014. Development of a novel DNA extraction method for identification and quantification of *Mycobacterium avium* subsp. *paratuberculosis* from tissue samples by real-time PCR. J Microbiol Meth **99**:58-65.
8. **Donaghy JA, Johnston J, Rowe MT**. 2010. Detection of *Mycobacterium avium* ssp. *paratuberculosis* in cheese, milk powder and milk using *IS*900 and f57- based qPCR assays. J Appl Microbiol **110**:479-489.
9. **Slana I, Paolicchi F, Janstova B, Navratilova P, Pavlik I**. 2008. Detection methods for *Mycobacterium avium* subsp. *paratuberculosis* in milk and milk products: a review. Vet Med **53**:283-306.
10. **Stevenson K**. 2015. Genetic diversity of *Mycobacterium avium* subspecies *paratuberculosis* and the influence of strain type on infection and pathogenesis: a review. Vet Res **46**:64.
11. **Hanifian S, Khani S, Barzegari A, Shayegh J**. 2013. Quantitative real-time PCR and culture examination of *Mycobacterium avium* subsp. *paratuberculosis* at farm level. Vet Microbiol **162**:160-165.
12. **Sting R, Hrubenja M, Mandl J, Seemann G, Salditt A, Waibel S**. 2014. Detection of *Mycobacterium avium* subsp. *paratuberculosis* in faeces using different procedures of pre-treatment for real-time PCR in comparison to culture. Vet J **199**:138-142.
13. **Windsor PA, Whittington RJ**. 2009. Evidence for age susceptibility of cattle to Johne's disease. Vet J **184**:37–44.
14. **Vansnick E, De Rijk P, Vercammen F, Geysen D, Rigouts L, Portaels F**. 2004. Newly developed primers for the detection of *Mycobacterium avium* subspecies *paratuberculosis*. Vet Microbiol **100**:197–204.
15. **Gilardoni LR, Paolicchi FA, Mundo SL**. 2012. Bovine paratuberculosis: a review of the advantages and disadvantages of different diagnostic tests. Rev Argent Microbiol **44**:201-215.
16. **OIE**. 2014. Paratuberculosis (Johne's Disease). In OIE Terrestrial Manual (World Organisation for Animal Health), Chapter 2.1.11.
17. **Fang Y, Wu WH, Pepper JL, Larsen JL, Marras SA E, Nelson EA, Epperson WB, Christopher-Hennings J**. 2002. Comparison of real-time, quantitative PCR with molecular beacons to nested PCR and culture methods for detection of *Mycobacterium avium* subsp. *paratuberculosis* in bovine fecal samples. J Clin Microbiol **40**:287–291.
18. **Bögli-Stuber K, Kohler C, Seitert G, Glanemann B, Antognoli MC, Salman MD, Wittenbrink MM, Wittwer M, Wassenaar T, Jemmi T, Bissig-Choisat B**. 2005. Detection of *Mycobacterium avium* subspecies *paratuberculosis* in Swiss dairy cattle by real-time PCR and culture: a comparison of the two assays. J Appl Microbiol **99**:587–597.

19. **Irenge LM, Walravens K, Govaerts M, Godfroid J, Rosseels V, Huygen K, Gala J L**. 2009. Development and validation of a triplex real-time PCR for rapid detection and specific identification of *M. avium* subsp. *paratuberculosis* in faecal samples. Vet Microbiol **136**:166–172.
20. **Kawaji S, Taylor DL, Mori Y, Whittington RJ**. 2007. Detection of *Mycobacterium avium* subsp. *paratuberculosis* in ovine faeces by direct quantitative PCR has similar or greater sensitivity compared to radiometric culture. Vet Microbiol **125**:36–48.
21. **Kawaji S, Begg DJ, Plain KM, Whittington RJ**. 2010. A longitudinal study to evaluate the diagnostic potential of a direct faecal quantitative PCR test for Johne's disease in sheep. Vet Microbiol **148**:35–44.
22. **Kralik P, Slana I, Kralova A, Babak V, Whitlock RH, Pavlik I**. 2011. Development of a predictive model for detection of *Mycobacterium avium* subsp. *paratuberculosis* in faeces by quantitative real time PCR. Vet Microbiol **149**:133–138.
23. **Leite FL, Stokes KD, Robbe-Austerman S, Stabel J**. 2013. Comparison of fecal DNA extraction kits for the detection of *Mycobacterium avium* subsp. *paratubeculosis* by polymerase chain reaction. J Vet Diagn Invest **25**:27-34.
24. **Selim A, El-haig M, Galila ES**. 2013. Direct detection of *Mycobacterium avium* subsp. *paratuberculosis* in bovine milk by multiplex real-time PCR. Biotech Anim Husb **29**:513-525.
25. **Khare S, Ficht TA, Santos RL, Romano J, Ficht AR, Zhang S, Grant IR, Libal M, Hunter D, Adams LG**. 2004. Rapid and Sensitive Detection of *Mycobacterium avium* subsp. *paratuberculosis* in bovine milk and feces by a combination of immunomagnetic bead separation-conventional PCR and Real-Time PCR. J Clin Microbiol **42**: 1075–1081.
26. **Botsaris G, Slana I, Liapi M, Dodd C, Economides C, Rees C, Pavlik I**. 2010. Rapid detection methods for viable *Mycobacterium avium* subspecies *paratuberculosis* in milk and cheese. Int J Food Microbiol **141**:S87–S90.
27. **Khol JL, Kralik P, Slana I, Beran V, Aurich C, Baumgartner W, Pavlik I**. 2010. Consecutive excretion of *Mycobacterium avium* subspecies *paratuberculosis* in semen of a breeding bull compared to the distribution in feces, tissue and blood by IS*900* and *F57* quantitative real-time PCR and culture examinations. J Vet Med Sci **72**:1283-1288.
28. **Bull TJ, McMinn EJ, Sidi-Boumedine K, Skull A, Durkin D, Neild P, Rhodes G, Pickup R, Hermon-Taylor J**. 2003. Detection and verification of *Mycobacterium avium* subsp. *paratuberculosis* in fresh ileocolonic mucosal biopsy specimens from individuals with and without Crohn's disease. J Clin Microbiol **41**:2915–2923.
29. **Gao A, Mutharia L, Raymond M, Odumeru J**. 2007. Improved template DNA preparation procedure for detection of *Mycobacterium avium* subsp. *paratuberculosis* in milk by PCR. J Microbil Meth **69**:417-420.
30. **Plain KM, Marsh IB, Waldron AM, Galea F, Whittington AM, Saunders VF, Begg DJ, de Silva K, Purdie AC, Whittington RJ**. 2014. High-throughput direct fecal PCR assay for detection of *Mycobacterium avium* subsp. *paratuberculosis* in sheep and cattle. J Clin Microbiol **52**:745-757.
31. **Timms VJ, Mitchell HM, Neilan BA**. 2015. Optimisation of DNA extraction and validation of PCR assays to detect *Mycobacterium avium* subsp. *paratuberculosis*. J Microbiol Meth **112**:99–103.
32. **Sidoti F, Banche G, Astegiano S, Allizond V, Cuffini AM, Bergallo M**. 2011. Validation and standardization of IS900 and F57 real-time quantitative PCR assays for the specific detection and quantification of *Mycobacterium avium* subsp. *paratuberculosis*. Can J Microbiol **57**:347-354.
33. **Bird BB, Madison BM**. 2000. Use of fluorochrome staining for detecting acid-fast mycobacteria. In Current Laboaratory Practice Series. Department of Health and Human Services. Centers for Diseases Control and Prevention, Atlanta, pp 1-17.
34. **Dimareli-Malli Z**. 2010. Detection of *Mycobacterium avium* subsp. *paratuberculosis* in milk from clinically affected sheep and goats. Intern J Appl Res Vet Med **8**:44-50.
35. **Costa P, Ferreira A, Amaro A, Albuquerque T, Botelho A, Couto I, Cunha MV, Viveiros M, Inácio J**. 2013. Enhanced detection of tuberculous mycobacteria in animal tissues using a semi-nested probe-based real-time PCR. PLoS ONE **8**:e81337.

36. **Castellanos E, Aranaz A, de Juan L, Alvarez J, Rodríguez S, Romero B, Bezos J, Stevenson K, Mateos A, Domínguez L**. 2009. Single nucleotide polymorphisms in the *IS*900 sequence of *Mycobacterium avium* subsp. *paratuberculosis* are strain type specific. J Clin Microbiol. **47**:2260-2264.
37. **Landis JR, Koch GG**. 1977. The measurement of observer agreement for categorical data. Biometrics **33**:159–174.
38. **Reddacliff LA, Vadali A, Whittington RJ**. 2003. The effect of decontamination protocols on the numbers of sheep strain *Mycobacterium avium* subsp. *paratuberculosis* isolated from tissues and faeces. Vet Microbiol **95**:271–282.
39. **Whittington RJ**. 2009. Factors affecting isolation and identification of *Mycobacterium avium* subsp. *paratuberculosis* from fecal and tissue samples in a liquid culture system. J Clin Microbiol **47**:614–622.
40. **Kralik P, Pribylova-Dziedzinska R, Kralova A, Kovarcik K, Slana I**. 2014. Evidence of passive faecal shedding of *Mycobacterium avium* subsp. *paratuberculosis* in a Limousin cattle herd. Vet J **201**:91–94.
41. **Cousins DV, Whittington R, Marsh I, Masters A, Evans RJ, Kluver P**. 1999. Mycobacteria distinct from *Mycobacterium avium* subsp. *paratuberculosis* isolated from the faeces of ruminants possess *IS*900-like sequences detectable by *IS*900 polymerase chain reaction: implications for diagnosis. Mol Cell Probes **13**:431–442.
42. **Englund S, Bölske G, Johansson KE**. 2002. An *IS*900-like sequence found in a *Mycobacterium* sp. other than *Mycobacterium avium* subsp. *paratuberculosis*. FEMS Microbiol Lett **209**:267-271.

Chapter V

Novel SNP-based assay for genotyping *Mycobacterium avium* subsp. *paratuberculosis*

Celia Leão, Robert J Goldstone, Josephine Bryant, Joyce McLuckie, João Inácio, David GE Smith and Karen Stevenson

Manuscript accepted for publication with minor modifications in 2015 in Journal of Clinical Microbiology (doi: 10.1128/JCM.01958-15)

Novel SNP-based assay for genotyping *Mycobacterium avium* subsp. *paratuberculosis*

5.1. Abstract

Typing of *Mycobacterium avium* subspecies *paratuberculosis* strains presents a challenge since they are genetically monomorphic and traditional molecular techniques have limited discriminatory power. The recent advances and availability of whole genome sequencing has extended possibilities for the characterization of *Mycobacterium avium* subspecies *paratuberculosis* and it can provide a phylogenetic context to facilitate global epidemiology studies. In this study we developed a SNP assay based on polymerase chain reaction and restriction enzyme digestion or sequencing of the amplified product. The SNP analysis was performed using genome sequence data from 133 *Mycobacterium avium* subspecies *paratuberculosis* isolates with different genotypes from eight different host species and seventeen distinct geographic regions around the world. A total of 28402 SNPs were identified among all the isolates. The minimum number of SNPs required to distinguish between all the 133 genomes was 93 and between only the Type C isolates was 41. To reduce the number of SNPs and PCRs required we adopted an approach based on sequential detection of SNPs and a decision tree. By the analysis of 14 SNPs *Mycobacterium avium* subspecies *paratuberculosis* isolates can be characterized within 14 phylogenetic groups with a higher discriminatory power compared to MIRU-VNTR assay and other typing methods. Continuous updating of genome sequences are needed in order to better characterize new phylogenetic groups and SNP profiles. The novel SNP assay is a discriminative, simple, reproducible method and requires only basic laboratory equipment for the large-scale global typing of *Mycobacterium avium* subspecies *paratuberculosis* isolates.

Keywords

Paratuberculosis; Whole Genome Sequencing; SNPs; MIRU-VNTR; typing; phylogenetic relationship

5.2. Introduction

Mycobacterium avium subspecies *paratuberculosis* (*Map*) causes Johne's disease, a chronic infectious enteritis principally of ruminants. The disease occurs worldwide and is responsible for significant losses to the livestock industry. *Map* also has been detected in a subset of human patients with Crohn's disease [1] although the zoonotic role of the bacterium remains controversial.

Strain typing is a prerequisite for tracing the sources of infection and for studying the epidemiology, population structure and evolutionary relationships between isolates. It can also reveal the genetic diversity underlying important phenotypic characteristics such as host specificity, pathogenicity, antibiotic resistance and virulence. Typing of *Map* strains presents a challenge since *Map*, like *Mycobacterium tuberculosis*, is genetically monomorphic [2]. Genetic diversity among *Map* strains has been investigated using molecular techniques such as restriction fragment length polymorphism and *IS*900 analysis (*IS*900 RFLP) [3], pulsed-field gel electrophoresis (PFGE) [4], amplified fragment length polymorphism (AFLP) analysis [5], random amplified polymorphic DNA (RAPD) analysis [6], mycobacterial interspersed repetitive unit – variable number tandem repeat (MIRU-VNTR) analysis [7] and short-sequence repeat (SSR) analysis [8]. However, these techniques have limited discriminatory power when applied to *Map* and although this can be increased by combining complementary genotyping techniques, it is often insufficient for accurately determining relationships among isolates or global epidemiological studies [9; 10].

Whole genome sequencing (WGS) provides the ultimate resolution of isolates and, unlike the techniques above, it can provide a phylogenetic context to facilitate global epidemiology studies and affirm epidemiological connections [10; 11; 12]. Although WGS is becoming cheaper, it is still too expensive to be used for routine genotyping of *Map* isolates and requires robust data handling and analysis processes. Single nucleotide polymorphisms (SNPs) have been used successfully to type several genetically monomorphic pathogens, including *M. tuberculosis* [13], *Mycobacterium bovis* [14], *Salmonella enteritica* Typhi [15] and *Yersinia pestis* [16]. SNP assays have been used to discriminate between *Map* strain Types I, II and III [17; 18] and between strains derived from animal and human hosts [19]. However, these assays were based on a limited number of SNPs, many of which were not informative when applied to a wider wild-type population.

The purpose of this study was to develop a SNP assay that is discriminative, practicable and reproducible for the large-scale global typing of *Map* isolates. Hence we developed a SNP typing method based on polymerase chain reaction (PCR) and restriction enzyme digestion, which would minimize costs and require only basic laboratory equipment. Additionally, we adopted an approach based on sequential detection of SNPs and a decision tree to reduce the number of SNPs and PCRs required.

5.3. Materials and methods

5.3.1. Selection of genome sequences and *Map* strains

Genome sequences from 133 *Map* isolates generated in a previous study [10] were selected for SNP and phylogenetic analyses. This panel was chosen to maximise genetic diversity and reduce phylogenetic discovery bias [2] since it comprised isolates with different genotypes (determined by multiplex PFGE and MIRU-VNTR), which were selected from 17 different countries and isolated from eight different host species. The panel also included isolates representing the major strain types that have been described and previously reviewed [20; 21]. The sequence reference numbers are given in Figures 5.1 and 5.2. The field isolates for which genome sequence data was not available and that were used to validate the SNP assay are shown in Table 5.1.

Table 5.1. Additional *Map* field isolates used for validation of SNP assay

Isolate	Host	Geographic location	Multiplex PFGE profile	INMV profile	SNP profile
F043531	cattle	Northern Ireland (UK)	2.1	1	1
F012398	cattle	Cumbria (UK)	2.1	1	11
F005713	cattle	Wiltshire (UK)	2.1	1	11
C217551	cattle	Selkirkshire (UK)	2.1.	1	9
C216785/2	cattle	Mid Glamorgan (UK)	2.1	1	9
C219376	cattle	Shropshire (UK)	2.1	1	9
C221325	cattle	Gwynedd (UK)	2.1	1	9
C524656	cattle	Aberdeenshire (UK)	2.1	1	1
C326442	cattle	Ayr (UK)	2.1	1	9
C216962/6	cattle	Leicestershire (UK)	2.1	1	9
C1	goat	São Miguel (Azores, PT)	ND	ND	3
C2	goat	São Miguel (Azores, PT)	ND	ND	3
C4	goat	São Miguel (Azores, PT)	ND	ND	3
C7	goat	São Miguel (Azores, PT)	ND	ND	3
C9	goat	São Miguel (Azores, PT)	ND	ND	3
C11	goat	São Miguel (Azores, PT)	ND	ND	3
C13	goat	São Miguel (Azores, PT)	ND	ND	3
C14	goat	São Miguel (Azores, PT)	ND	ND	3
C16	goat	São Miguel (Azores, PT)	ND	ND	3
C4A4	goat	São Miguel (Azores, PT)	ND	ND	3
B1	cattle	Vila do Conde (PT)	ND	2	9
B3	cattle	Vila do Conde (PT)	ND	2	9
B13	cattle	Póvoa do Varzim (PT)	ND	2	11
B18	cattle	Póvoa do Varzim (PT)	ND	2	11
B21	cattle	Póvoa do Varzim (PT)	ND	2	11
B22	cattle	Póvoa do Varzim (PT)	ND	2	11

ND - not determined; UK – United Kingdom; PT - Portugal

5.3.2. Preparation of genomic DNA

Genomic DNA from the UK field isolates (Table 5.1) was extracted from plugs used to perform the PFGE analysis. Briefly, half of each plug was washed three times with 2 ml of Tris EDTA buffer pH 8 (10 mM Tris.Cl, 1 mM EDTA) for 10 minutes with shaking. After washing the plugs, 0.5-1 ml of sterile deionised water was added and the agarose was melted at 70ºC. The suspension was stored at -20ºC until required for PCR. DNA from the Portuguese field isolates (Table 5.1) was extracted from pure cultures grown in Middlebrook 7H9 medium supplemented with 10% OADC (Oleic Albumin Dextrose Catalase) and 2 mg/l of mycobactin as previously described [10] using the QIAamp DNA Mini Kit (Qiagen) according to manufacturer's instructions with minor modifications. Briefly, the bacterial suspension was centrifuged at 5000 × g for 10 minutes and the pellet was resuspended in 180 µl of ATL buffer. Zirconium beads (0.1 mm) were added to the mixture and mechanical disruption of the cells was performed twice with a FastPrep FP120 Bio101 bead shaker (Savant Instruments Inc., Holbrook, NY) at 6.5 $msec^{-1}$ for 45 seconds. The disrupted mixture was cooled on ice for 15 minutes and 20 µl of proteinase K was added. The remaining procedure was completed according to manufacturer's instructions. The DNA was eluted with 200 µl of AE buffer and stored at -20ºC until required for PCR.

5.3.3. SNP analysis and phylogenomics

The genome sequence of *Map* K10 [22; Accession number NC_002944.2] was used as the reference genome for the phylogenetic analyses. Raw genome sequence data for the *Map* isolates are available from the European Nucleotide Archive under accession PRJEB2204. The assembly of the reads into contigs was performed using Velvet assembler program (freely available at https://www.ebi.ac.uk/~zerbino/velvet/) and the alignment of the sequences and positioning of SNPs was executed utilising MUMmer package (freely available at http://mummer.sourceforge.net/) using *Map* K10 as reference sequence. SNPs were then extracted and concatenated and a phylogenetic tree was calculated using phyML. This phylogenetic tree was then imported into R to explore the 'paths' which exist between the genomes. These 'paths' are a description of the strains which group on the same branch as the structure of the phylogenetic tree descends (i.e. all strains are included on the 'root branch', which then bifurcates to include a subset of strains on one branch and the remaining strains on the other, and so on). Using these 'paths', SNPs which were shared by all strains on each branch were investigated. This data then allowed comparison of the SNPs present in reducing sized groups of genome sequences (groups become progressively smaller as the number of branches increases as the tree descends). This data permitted the detection of discriminate SNPs which are present in one group of sequences yet absent in others. A 'set cover' analysis was then performed on this dataset to calculate the minimum number of SNPs which were required to discriminate between the groups of

strains at each selected level of the tree. These SNPs were then selected and taken forward for validation as below. SNPs were named according to their base position in the revised *Map* K10 genome sequence.

5.3.4. SNP selection and primer design

In order to determine the minimum number of SNPs necessary to distinguish between all the 133 *Map* sequences, the minimum 'set cover' approach was adopted and a file was generated with all the SNP's combinations for the characterization of all the genomes. The selection of the SNPs was a computer random choice between all the possible SNP combinations.

Primers for PCR amplification of the selected SNPs were designed using the online software Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/) and the revised *Map* K10 genome sequence [23; Accession number AE016958.1] (Table 5.2). Primers were designed to be 18-20 mer with a melting temperature between 63 and 67ºC.

Table 5.2. PCR primers and restriction endonucleases used in the SNP assay for this study

SNP name	Primer name	Primer sequence	Annealing temperature (^{0}C)	Product size (bp)	Enzyme	Restriction results§
snp3842359*	MAP_F1	CACCTGGCCAAGTACTACCA	63	528	BsmBI	(A) Type C - 528 bp
	MAP_R1	GCGATGTCATGATGCTGCTG				(G/C) Type S - 312, 216 bp
snp343677	TypeS_F	AACACCAGGATCGCGTTCTT	65	511	AvaII	(G) Type I - 297, 152, 62 bp
	TypeS_R	CAATTAGCGGTCGAGTCGTC				(A) Type III - 293, 218 bp
snp50173	BisonF	GGACGATTACTCGGTTCCAG	63	469	ApoI	(T) Bison group - 226, 192, 51 bp
	BisonR	ACCCGTGTTCGGCTACCT				(G) Cattle group - 277, 192 bp
snp4111202	SNP4_F	GTCAGAAACATCCCGCCTTC	65	461	FatI	(G) Type C sub-group A – 284, 177 bp
	SNP4_R	GTATTGAGTGAGGCAAGCGG				(C) Type C sub-group B - 461 bp
snp3879247	SNP5_F	GTTGATCGACAGCGAGTGC	64	465	BlpI/DdeII	(C) – 227, 238 bp
	SNP5_R	GTGGTGTCCGAGGTGAACTT				(T) – 465 bp
snp2939977	SNP6_F	TATCTCCAAGGACGCATTCC	64	516	-	-
	SNP6_R	CTGCCATGTCCGTCCTTAAT				
snp1932058	SNP7_F	GGCTTGAAACTCCAACTGCT	63	452	-	-
	SNP7_R	CGTCGTACATCCTCGTGGT				
snp1327872	SNP8_F	GCGCTTGTTGTACAGGTTGA	64	528	AvaI/BsoBI	(G) – 292, 171, 65 bp
	SNP8_R	TACGACGAAGACCCCGACTA				(T) – 463, 65 bp

SNP name	Primer name	Primer sequence	Annealing temperature (^{0}C)	Product size (bp)	Enzyme	Restriction results§
snp3844632	SNP9_F	GATCGATGCGGAGCTCGT	64	457	FatI	(G) – 457 bp
	SNP9_R	TGACAGGAAGGTCCATAGCC				(C) – 241, 217 bp
snp1966028	SNP10_F	GTCGAGGGCTTCCAGGTT	67	427	SapI/EarlI	(A) – 427 bp
	SNP10_R	GTCTGAGGCCAGCGACAC				(C) – 246, 182 bp
snp305277†	SNP11_F	CCATCCCGAGTTCAACAAGT	64	461	BspMI/BfuAI	(G) – 310, 151 bp
	SNP11_R	ACTTGTCGGGGTTGTAGCTG				(A) – 461 bp
snp4339946	SNP12_F	AACCGCTCAAGGCGAAAG	64	464	BstEII	(T) – 292, 172 bp
	SNP12_R	TCCCTTATCTGCGAAGTGCT				(A) – 464 bp
snp2087274	SNP13_F	CAGACCGAGCACCTCCTG	65	453	HpyAV	(C) – 453 bp
	SNP13_R	CCGCGTTGAAGGATCTCAAG				(A) – 227, 226 bp
snp1686154	SNP14_F	GAATCCCCGGAACTGGTG	65	525	MscI	(G) – 525 bp
	SNP14_R	GCAGTCCAGATAACGGAACG				(A) – 284, 241 bp

*SNP can also distinguish between Type I and III (Type I has a "G" and Type III has a "C" at base position 3842359 detectable by sequencing the PCR product); †SNP can also distinguish between two phylogenetic subgroups of Bison group isolates (Figure 5.3); § in parenthesis is the expected base at the correspondent SNP position; - not applicable.

PCR reactions were carried out in 50 µl containing 200 µM of each deoxynucleotide triphosphate (Invitrogen), 0.5 µM of each primer (Table 5.2), 1 U of Phusion® High- Fidelity DNA polymerase, 1× Phusion GC buffer (New England Biolabs), and 4 µl of extracted DNA. Amplification was performed in a TC-PLUS Thermal cycler (Techne) with an initial step at 98 ºC for 3 minutes, followed by 35 cycles at 98 ºC for 30 seconds, 63-67 ºC for 30 seconds (annealing temperatures provided in Table 5.2) and 72 ºC for 40 seconds, ending with a step at 72 ºC for 10 minutes. The amplified products were electrophoretically analysed in a 1.5% (w/v) agarose gel stained with SYBR® Safe DNA Gel Stain (Life Technologies) in 0.5× Tris-Borate-EDTA (TBE) buffer. Gel electrophoresis images were acquired with an Alphaimager™ 2200 (Alpha Innotech). DNA ladder IV (Bioline) and *Map* K10 (positive control) were included on each gel.

Restriction endonuclease analysis of PCR products was performed according to the manufacturer's instructions using 10 µl of amplified product in a total reaction volume of 25 µl. All restriction endonucleases were purchased from New England Biolabs (Table 5.2). Restricted products were detected by electrophoresis on 1.5% (w/v) agarose gels as described above.

Confirmation of the presence of the SNPs was obtained by sequencing the PCR products. PCR product (40µl) was purified using QIAquick PCR purification Kit (Qiagen) according to manufacturer's instructions. Sequencing of PCR products was carried out by Eurofins Genomics (MWG-Biotech) using the same primers used to the amplification of the fragments. Confirmation of the presence or absence of the SNP in the expected position of the genome was achieved using Basic Local Alignment Search Tool (Blast - http://blast.ncbi.nlm.nih.gov/Blast.cgi). The phylogenetic profile for each isolate was obtained by combining the results for all SNPs.

5.3.5. Discriminatory power of SNP-based genotyping assay

The discriminatory power of the assay was calculated using the Hunter-Gaston Discriminatory Index (HGDI), according to Hunter and Gaston [24], using the formula:

$$HGDI = 1 - \left[\frac{1}{N(N-1)} \sum_{j=1}^{s} n_j (n_j - 1) \right]$$

where N is the total number of isolates, s the number of typing groups obtained and n_j the number of isolates belonging to the j-th typing group.

5.4. Results

The genome sequence data from 133 *Map* strains were compared to the reference *Map* strain K10 to identify SNPs. A total of 28402 SNPs were identified among all the isolates. A phylogenetic tree was generated based on the SNP analysis and distinct phylogenetic groups were identified (Figures 5.1 and 5.2), which conformed to the broadly recognised phylogenetic structure of *Map* [10].

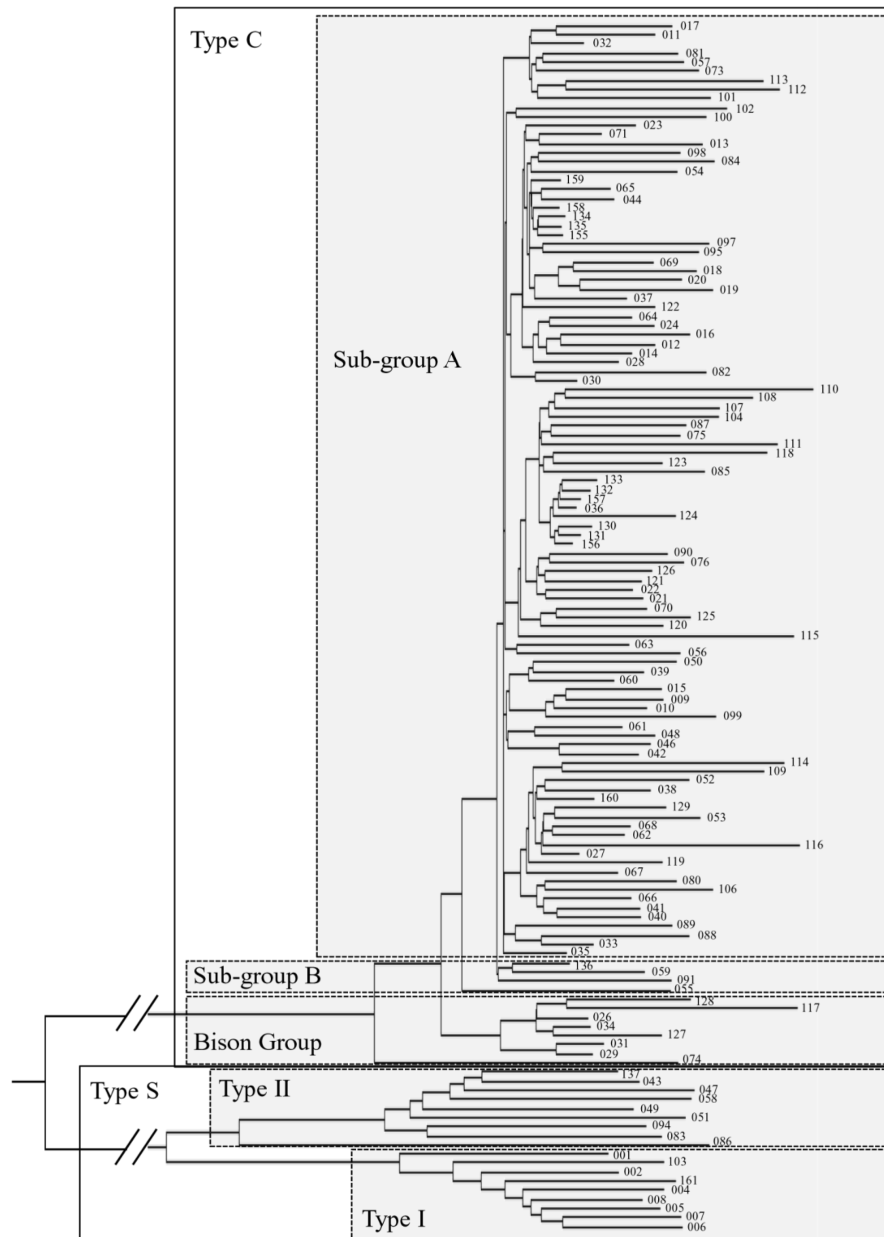

Figure 5.1. Whole genome SNP-based phylogenetic tree of 133 *Map* isolates included in this study.

Strain sequence reference MAPMRI numbers are indicated. Previously described lineages and sub-groups A and B described in this study are highlighted in grey.

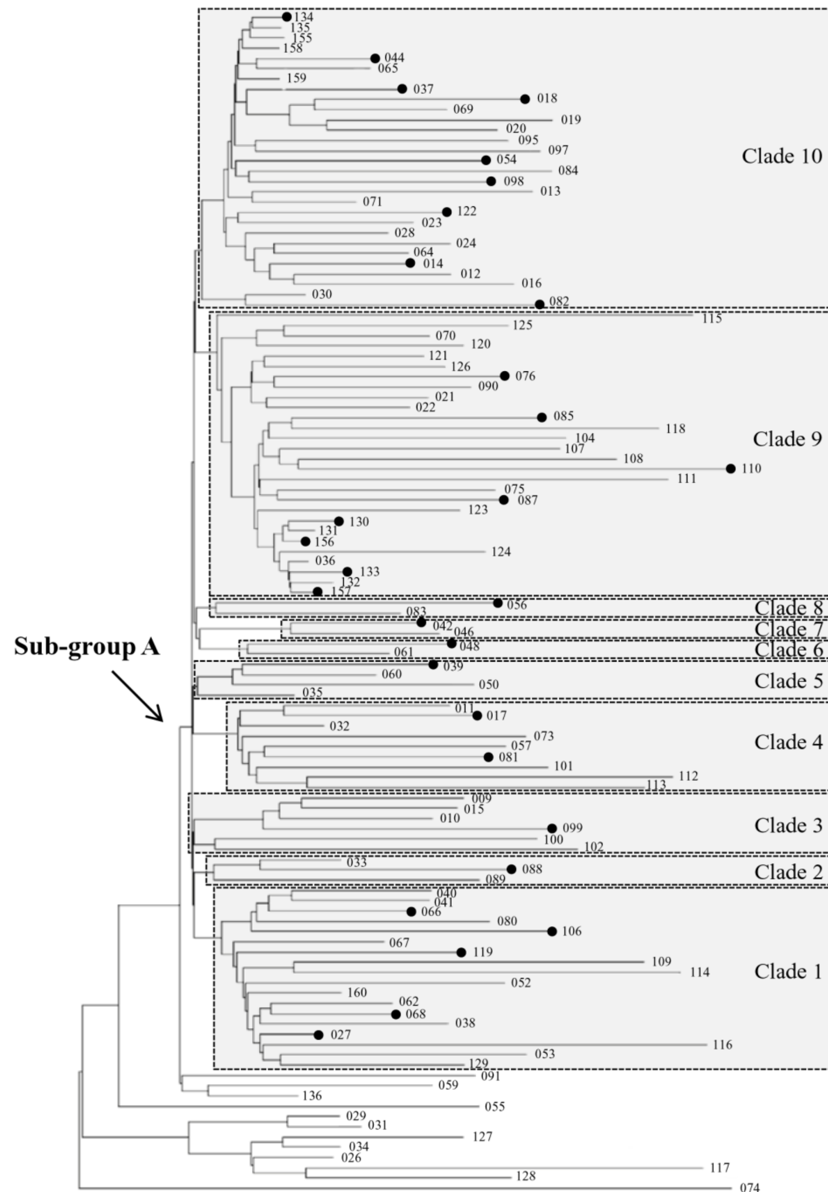

Figure 5.2. Whole genome SNP-based phylogenetic tree of 115 Type C *Map* isolates.

Strain sequence reference MAPMRI numbers are indicated. Ten clades within the phylogenetic sub-group A can be distinguished by PCR and sequencing following the analysis of ten SNPs (grey boxes). Black circles indicate the 30 strains used for the validation of the method.

By using an adaptation of the 'set cover' problem, the minimum number of SNPs required to discriminate between all the isolates was calculated to be 93. To refine the number of SNPs to a number manageable for routine laboratory procedures, we considered a strategy based on sequential detection of SNPs and a decision tree (Figure 5.3).

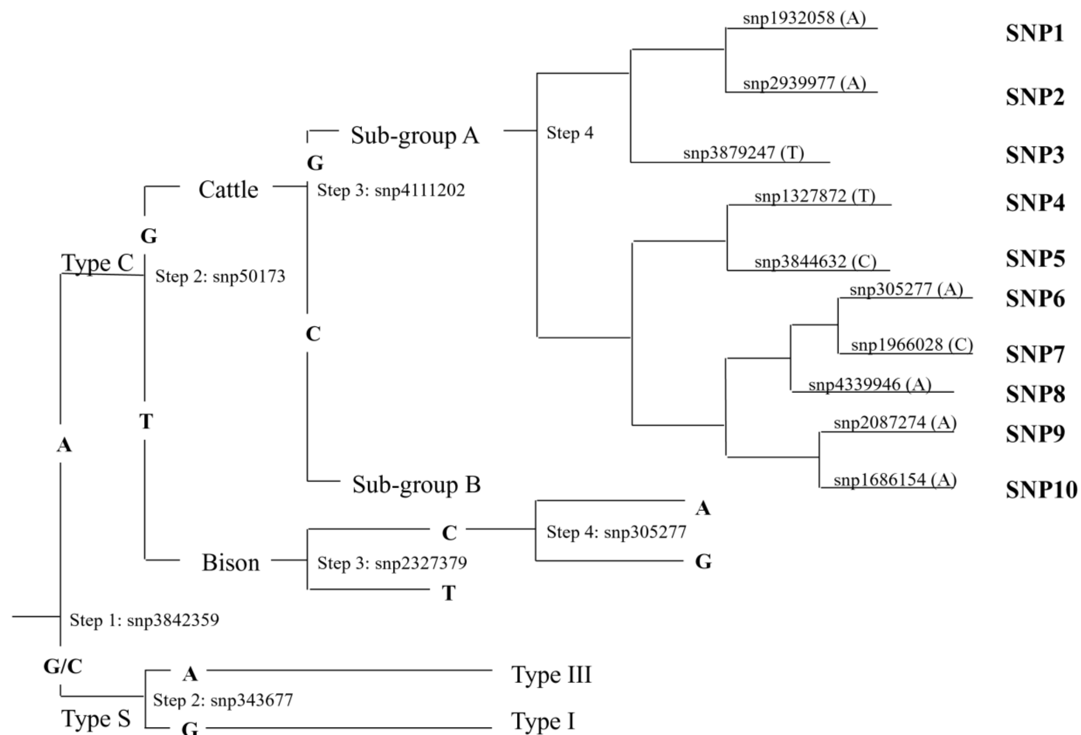

Figure 5.3. Work flow with a schematic representation of the decision tree with the sequential numbered steps, SNPs positions, expected bases and SNP profiles obtained based on the SNPs analysis.

The phylogenetic analysis distinguished two major strain groups corresponding to those previously designated Type C and Type S (Figure 5.1). We identified a SNP (snp3842359), which could be detected using BsmB1 (Table 5.2) that would discriminate between these two groups. Type C strains have an 'A' and Type S strains either a 'G' or 'C' at base position 3842359. This constituted the first step in the decision tree, the next step being determined according to whether the isolate was Type C or Type S (Figure 5.3).

For further analysis of Type S strains, we identified a SNP (snp343677), which could be detected using AvaII that discriminated the Type S sub-groups Type I and Type III (Table 5.2, Figures 5.1 and 5.3). Additionally, snp3842359 also could be used to distinguish Type I and Type III strains since Type I have a "G" and Type III have a "C" at base position 3842359 (Figure 5.3), which could be detected by PCR amplification and sequencing of the product.

The Type C group comprised the majority of the isolates and the high homogeneity within this group posed a challenge for identification of clade specifying SNPs. Firstly, SNP analysis was repeated with only the sequence data from the 115 Type C isolates (Figure 5.2) and the minimum number of SNPs required to distinguish between these 115 isolates was determined to be 41. We then

considered three principal sub-groups designated Bison [as reported previously 10] and A and B as shown in Figures 5.1 and 5.2. We identified a SNP (snp50173), which could distinguish the Bison group from both sub-groups A and B using ApoI (Table 5.2, Figures 5.1 and 5.3). This constituted step 2 in the decision tree (Figure 5.3).

Within the Bison group, Indian Bison type could be differentiated from US Bison type using snp2327379 and further differentiation of the Indian bison type was possible using snp305277 (Table 5.2, Figure 5.3). Due to the limited number of Bison type strains available extensive verification of these SNPs was not possible in this study.

To further discriminate between Type C isolates in sub-groups A and B, we identified a SNP (snp4111202), which could be detected using FatI (Table 5.2, Figures 5.1 and 5.3). Due to the small number of isolates in sub-group B, we did not, at this stage, seek additional SNPs to further discriminate the isolates within this group. This constituted step 3 in the decision tree (Figure 5.3).

Within the larger subgroup A, SNPs were identified that could discriminate ten groups (snp3879247, snp2939977, snp1932058, snp1327872, snp3844632, snp1966028, snp305277, snp4339946, snp2087274 and snp1686154, Tables 5.2 and 5.3, Figures 5.2 and 5.3). These SNPs were verified by sequencing of the PCR products and comparison of the sequences with the reference *Map* K10 strain. It was not possible to identify SNPs with specific restriction endonuclease sites for the clades differentiated using snp2939977 and snp1932058 but all other SNPs could be detected by restriction endonuclease analysis (Table 5.2).

5.4.1. SNP validation and genotyping

For the validation of the selected SNPs comprising the decision tree (Figure 5.3), DNA from isolates belonging to Type C (Bison group: MAPMRI029, MAPMRI031, MAPMRI127, MAPMRI034, MAPMRI117 and MAPMRI026; Sub-group A: MAPMRI110, MAPMRI120 and MAPMRI027; Sub-group B: MAPMRI059, MAPMRI136 and MAPMRI091) and Type S (Type I: MAPMRI007and MAPMRI001; Type III: MAPMRI051, MAPMRI045 and MAPMRI047) (Figure 5.2) were used for amplifying products containing SNPs (as indicated in Figure 5.3) and the PCR products were digested with the correspondent restriction enzymes (Table 5.2). PCR products were also purified and sequenced to confirm SNPs. To assess the validity of the ten SNPs for discriminating the clades within Sub-group A, 30 isolates from the original panel of 115 sequenced strains were re-tested. These were selected to be representative of the ten different phylogenetic clades as shown in Figure 5.2 and were subjected to analysis for all 14 SNPs (Figure 5.3), all of which were confirmed to be present. The *Map* isolates were grouped into SNP profiles as shown in Table 5.3 and Figure 5.3.

Table 5.3. SNP profiles of Type C *Map* isolates in phylogenetic sub-group A used in this study

Phylogenetic group	**SNP profile**	**No. isolates verified**	**Base at SNP position***													
			3842359	343677	50173	4111202	3879247	2939977	1932058	1327872	3844632	1966028	305277	4339946	2087274	1686154
Reference base (K10)			A	A	G	G	C	G	G	G	G	A	G	T	C	A
Clade 1	1	7	A	A	G	G	C	G	**A**	G	G	A	G	T	C	G
Clade 2	2	1	A	A	G	G	C	**A**	G	G	G	A	G	T	C	G
Clade 3	3	11	A	A	G	G	**T**	G	G	G	G	A	G	T	C	G
Clade 4	4	2	A	A	G	G	C	G	G	**T**	G	A	G	T	C	G
Clade 5	5	1	A	A	G	G	C	G	G	G	**C**	A	G	T	C	G
Clade 6	6	1	A	A	G	G	C	G	G	G	G	A	**A**	T	C	G
Clade 7	7	1	A	A	G	G	C	G	G	G	G	**C**	G	T	C	G
Clade 8	8	1	A	A	G	G	C	G	G	G	G	A	G	**A**	C	G
Clade 9	9	16	A	A	G	G	C	G	G	G	G	A	G	T	**A**	G
Clade 10	10	9	A	A	G	G	C	G	G	G	G	A	G	T	C	**A**
new clade	11	6	A	A	G	G	C	G	G	G	G	A	G	T	C	**G**
	Total no.	56														

* SNP position in the revised *Map* K10 sequence (Accession number: AE016958.1).
Defining SNP base marked in bold type

A further 26 *Map* isolates (Table 5.1) not previously sequenced or typed using this SNP assay were genotyped using the 14 SNPs. These isolates belonged to four phylogenetic groups within the Sub-group A. Significantly, ten isolates from different geographic regions of United Kingdom with the same MIRU-VNTR and PFGE profile were classified into three different SNP profiles, one of which was not identified in the original phylogenetic analysis and therefore represented a new SNP profile (SNP11) (Table 5.3). Two UK isolates were identified to belong to the SNP1 profile, six isolates to SNP9 and two isolates to profile SNP11. The 16 Portuguese isolates were distributed among three phylogenetic groups: all the isolates from Azores were present in the same group identified by profile SNP3; two isolates from the same region in the north of Portugal were identified in the same phylogenetic group as six isolates from the United Kingdom (profile SNP2); and, the remaining four Portuguese isolates from the same region were found to belong to the new phylogenetic group together with the two DNAs from the United Kingdom (profile SNP11) (Table 5.1).

5.4.2. Discriminatory power of SNP-based genotyping assay

The discriminatory power of the SNP-based assay was 0.8390 for the 56 isolates that were used for the validation of the assay. In order to compare the discriminatory power of the SNP assay with MIRU-VNTR analysis, we used the typing results for 46 isolates, which had been typed using both methods. The HGDI was calculated to be 0.8135 for the SNP assay and 0.6386 for MIRU VNTR.

5.5. Discussion

Several methods have been used to characterize *Map* strains but they have some limitations. Techniques based on the analysis of total genomic DNA such as RFLP and PFGE require culture of the isolates to prepare moderate amounts of high quality DNA and are therefore slow, can be technically demanding, labor intensive, hard to standardize and expensive. Furthermore, RFLP and PFGE can clearly distinguish between Type C and S but do not give sufficient discrimination within these strain types for detailed epidemiological studies. Techniques such as AFLP and RAPD employ PCR to detect smaller genomic DNA fragments but are less utilised for epidemiological studies due to difficulties in standardisation, reproducibility and limited discriminative power. Other typing methods based on repetitive sequences such as SSR and MIRU-VNTR are popular due to their ease of use and rapidity but are again limited with respect to their ability to discriminate within the two major strain types and the typing results may not reflect the evolutionary relationships between isolates [10; 12; 25; 26].

In this study a novel typing assay based on SNP analysis by PCR and restriction or sequencing of the amplified products was developed. This technique is easy to perform, is applicable to a small quantity of genomic DNA and is based on standard PCR and restriction endonuclease analysis. It was possible to refine the number of SNPs to a number manageable for routine laboratory procedures by adopting an approach based on sequential detection of SNPs via a decision tree. The SNP assay was highly discriminative, possessing a higher discriminatory power than MIRU-VNTR when applied to 46 *Map* isolates.

SNP-based typing assays are particularly useful for monomorphic pathogens that exhibit limited genetic diversity. Furthermore, they have the advantage that they can be used to determine phylogenetic relationships, unlike techniques based on mobile or repetitive DNA elements, which interrogate a relatively small proportion of the mycobacterial genome and can exhibit homoplasy. However, SNP discovery is subject to phylogenetic discovery bias [2], a phenomenon well described for *M. tuberculosis* [27] and *Bacillus anthracis* [28], and is most likely to be encountered where information is missing on strains geographically-restricted or belonging to rare phylogenetic groupings. For this reason, we utilised a large collection of global isolates, which had been previously genotyped by classical molecular tools (PFGE and MIRU-VNTR) to maximise genetic diversity and include representatives of all previously reported strain types. The SNPs identified in this study should provide the necessary means to unambiguously classify *Map* strains within this global framework. Even so, the composition of any panel of SNPs needs to be reviewed or augmented once additional groups of strains that were not included in the initial analysis are discovered. This has been illustrated in this study with the discovery of a new phylogenetic group represented by profile SNP 11 comprising six isolates when an additional uncharacterised 26 isolates were screened using the SNP assay. WGS needs to be performed and the sequence comparisons and SNP analysis repeated to determine the positions of these isolates within the phylogenetic tree and determine any additional SNPs that could be used to define the group. In this study, we concentrated on finding SNPs to differentiate within Type C strains in sub-group A. The SNP assay needs to be expanded to differentiate between strains within Type S, Bison type and Type C subgroup B, but WGS data from more strains belonging to these phylogenetic groups are required for SNP discovery to provide a phylogenetically robust framework for strain differentiation combined with sufficient discriminatory power for detailed genetic studies.

SNPs have been described in previous studies that differentiate between the two major phylogenetic groups, Type S and Type C, and between *Map* strain Types I and III. A PCR-REA assay described by Whittington et al. [29] based on a SNP at base position 223 in the *IS*1311 insertion sequence has been used extensively for discriminating between Type S and Type C strains. However, in a recent study [10] the *IS*1311-REA incorrectly identified strain MAPMRI074 as a Type S strain when WGS and SNP analysis clearly confirmed it to be Type C, suggesting that the C to T allelic

variation at base pair position 223 in IS*1311* occurred after the initial divergence of Type C from Type S strains. The SNP identified in this study, snp3842359, and the corresponding restriction endonuclease BsmBI for PCR-REA SNP detection could provide a more reliable alternative assay for differentiating Type S and C strains.

IS*1311* SNP analysis has also been used to distinguish Bison type strains from non-Bison Type C and Type S strains [30]. In Bison-type strains, all copies of IS*1311* have a "T" at base pair position 223, whereas the non-Bison Type C strains have one or more copies with a "C" or "T" at the same position. Copy number with respect to this allele is not always easy to assess and can be very variable [10]. Snp50173 and the corresponding restriction endonuclease ApoI for PCR-REA SNP detection could provide an easier, alternative assay for discriminating Bison-type strains from other Type C strains.

A study published by Castellanos and colleagues [17] developed a PCR-REA to detect a SNP present on the *gyrB* gene at base position 1626 that allowed discrimination of Type III from Type I and II strains. Additional SNPs in the *gyrA* gene were identified that could differentiate Types I and III from Type II. In our study it is possible to use only a single SNP to discriminate Type I, Type II and Type III based on the amplification and sequencing of a fragment containing the snp3842359 where in the same position of the genome Type I strains have a "G", Type II have an "A" and Type III have a "C". This is an improvement compared with the system previously reported.

We developed a novel SNP-based genotyping assay based on the analysis of 14 SNPs that can be used to characterize *Map* isolates within 14 phylogenetic groups with a higher discriminatory power compared to MIRU-VNTR assay and other typing methods. We adopted an approach based on sequential detection of SNPs and a decision tree based on PCR-restriction enzyme digestion to reduce the number of SNPs and required PCRs. This novel assay can overcome some issues regarding the genotyping of isolates characterized as Type I, Type III and Bison type. Continuous updating of genome sequences are needed in order to better characterize new phylogenetic groups and SNP profiles. The novel SNP assay is a discriminative, simple, reproducible method and requires only basic laboratory equipment for the large-scale global typing of *Map* isolates.

Acknowledgements

This work was supported by the Scottish Government Rural and Environment Science and Analytical Services Division. Célia Leão is a recipient of a PhD grant from "Fundação para a Ciência e a Tecnologia" SFRH/BD/62469/2009.

Author's contribution

CL contributed to the bioinformatics analysis, designing PCRs systems, experimental work of DNA extraction, enzymatic restrictions, DNA purification, data analysis and writing the manuscript. RG contributed with bioinformatic support and analysis, data analysis and revised the manuscript. JB contributed with sequencing data and revised the manuscript. JM contributed with isolates DNA. JI contributed to the designing of the study and revised the manuscript. DS and KS contributed to the designing of the study, data analysis and revised the manuscript. All authors have read and approved the final manuscript.

5.6. References

1. **Naser SA, Sagramsingh SR, Naser AS, Thanigachalam S**. 2014. *Mycobacterium avium* subspecies *paratuberculosis* causes Crohn's disease in some inflammatory bowel disease patients. World J Gastroenterol **20**:7403-7415.
2. **Achtman M.** 2008. Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. Ann Rev Microbiol **62**:53-70.
3. **Pavlik I, Horvathova A, Dvorska L, Bartl J, Svastova P, du Maine R, Rychlik I.** 1999. Standardisation of restriction fragment length polymorphism analysis for *Mycobacterium avium* subspecies *paratuberculosis*. J Microbiol Meth **38**:155-167.
4. **Stevenson K, Hughes VM, de Juan L, Inglis NF, Wright F, Sharp JM**. 2002. Molecular characterization of pigmented and non-pigmented isolates of *Mycobacterium avium* subspecies *paratuberculosis*. J Clin Microbiol **40**:1798-1804.
5. **Motiwala AS, Strother M, Amonsin A, Byrum B, Naser SA, Stabel JR, Shulaw WP, Bannantine JP, Kapur V, Sreevatsan S.** 2003. Molecular epidemiology of *Mycobacterium avium* subsp. *paratuberculosis*: Evidence for limited strain diversity, strain sharing, and identification of unique targets for diagnosis. J Clin Microbiol **41**:2015-2016.
6. **Pillai SR, Jayarao BM, Gummo JD, Hue Jr EC, Tiwari D, Stabel JR, Whitlock RH.** 2001. Identification and sub-typing of *Mycobacterium avium* subsp. *avium* by randomly amplified polymorphic DNA. Vet Microbiol **79**:275-284.
7. **Thibault VC, Grayon M, Boschiroli ML, Hubbans C, Overduin P, Stevenson K, Gutierrez MC, Supply P, Biet F.** 2007. New variable-number tandem-repeats markers for typing *Mycobacterium avium* subsp. *paratuberculosis* and *M. avium* strains: comparison with *IS*900 and *IS*1245 restriction fragment length polymorphism typing. J Clin Microbiol **45**:2404-2410.
8. **Sevilla I, Li L, Amonsin A, Garrido J, Geijo MV, Kapur V, Juste RA**. 2008. Comparative analysis of *Mycobacterium avium* subsp. *paratuberculosis* isolates from cattle, sheep and goats by short sequence repeat and pulsed-field gel electrophoresis typing. BMC Microbiol **8**:204.
9. **Stevenson K, Àlvarez J, Bakker D, Biet F, de Juan L, Denham S, Dimareli Z, Dohmann K, Gerlach G-F, Heron I, Kopecna M, May L, Pavlik I, Sharp JM, Thibault VC, Willemsen P, Zadoks R, Greig A.** 2009. Occurrence of *Mycobacterium avium* subspecies *paratuberculosis* across host species and European countries with evidence for transmission between wildlife and domestic ruminants. BMC Microbiol **9**:212.
10. **Bryant JM, Thibault VC, Smith DGE, McLuckie J, Heron I, Sevilla IA, Biet F, Harris SR, Maskell DJ, Bentley SD, Parkhill J, Stevenson K**. 2015. Phylogenomic exploration of the relationships between strains of Mycobacterium avium subspecies paratuberculosis. BMC Genomics In press.
11. **Bryant JM, Schürch AC, van Deutekom H, Harris SR, de Beer JL, de Jager V, Kremer K, van Hijum SA, Siezen RJ, Borgdorff M, Bentley SD, Parkhill J, van Soolingen D.** 2013. Inferring patient to patient transmission of *Mycobacterium tuberculosis* from whole genome sequencing data. BMC Infect Dis **13**:100.
12. **Ahlstrom C, Barkema HW, Stevenson K, Zadoks RN, Biek R, Kao R, Trewby H, Haupstein D, Kelton DF, Fecteau G, Labrecque O, Keefe GP, McKenna SLB, Buck JD**. 2015. Limitations of variable number of tandem repeat typing identified through whole genome sequencing of *Mycobacterium avium* subsp. *paratuberculosis* on a national and herd level. BMC Genomics **16**:161.
13. **Stucki D, Malla B, Hostettler S, Huna T, Feldmann J, Yeboah-Manu D, Borrell S, Fenner L, Comas I, Coscollà M, Gagneux S.** 2012. Two new rapid SNP-Typing methods for classifying *Mycobacterium tuberculosis* Complex into the main phylogenetic lineages. PLOS ONE **7**:e41253.

14. **Joshi D, Harris NB, Waters R, Thacker T, Mathema B, Krieswirth B, Sreevatsan S.** 2012. Single Nucleotide Polymorphisms in the *Mycobacterium bovis* genome resolve phylogenetic relationships. J Clin Microbiol **50**:3853-3861.
15. **Octavia S, Lan R.** 2007. Single-Nucleotide-Polymorphism typing and genetic relationships of *Salmonella enterica* Serovar Typhi Isolates. J Clin Microbiol **45**:3795–3801.
16. **Morelli G, Song Y, Mazzoni CJ, Eppinger M, Roumagnac P, Wagner DM, Feldkamp M, Kusecek B, Vogler AJ, Li Y, Cui Y, Thomson NR, Jombart T, Leblois R, Lichtner P, Rahalison L, Petersen JM, Balloux F, Keim P, Wirth T, Ravel J, Yang R, Carniel E, Achtman M.** 2010. *Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity. Nat Genet **42**:1140-1143.
17. **Castellanos E, Aranaz A, Romero B, de Juan L, Àlvarez J, Bezos J, Rodríguez S, Stevenson K, Mateos A, Domínguez L.** 2007. Polymorphisms in *gyrA* and *gyrB* genes among *Mycobacterium avium* subsp. *paratuberculosis* Type I, II, and III isolates. J Clin Microbiol **45**:3439-3442.
18. **Gastaldelli M, Stefani E, Lettini AA, Pozzato N.** 2011. Multiplexed typing of *Mycobacterium avium* subsp. *paratuberculosis* Type I, II and III by Luminex xMAP suspension array. J Clin Microbiol **49**:389-391.
19. **Wynne JW, Bellera C, Boyda V, Francisc B, Gwoźdźd J, Carajiasd M, Heinea HG, Wagnere J, Kirkwoode CD, Michalskia WP.** 2014. SNP genotyping of animal and human derived isolates of *Mycobacterium avium* subsp. *paratuberculosis*. Vet Microbiol **172**:479-485.
20. **Stevenson K. 2010**. Comparative Differences between Strains of *Mycobacterium avium* subsp. *paratuberculosis* p126-137. *In* Behr MA, Collins DM (eds), Paratuberculosis Organism, Disease, Control. CAB International, Oxfordshire, UK.
21. **Stevenson K. 2015**. Genetic diversity of *Mycobacterium avium* subsp. *paratuberculosis* and the influence of strain type on infection and pathogenesis: A review. Vet Res **46**:64.
22. **Li L, Bannantine JP, Zhang Q, Amonsin A, May BJ, Alt D, Banerji N, Kanjilal S, Kapur V**. 2005. The complete genome sequence of *Mycobacterium avium* subspecies *paratuberculosis*. Proc Natl Acad Sci USA **102:**12344-12349.
23. **Wynne JW, Seeman T, Bulach D, Coutts SA, Talaat AM, Michalski WP**. 2010. Re-sequencing the *Mycobacterium avium* subspecies *paratuberculosis* K10 genome: improved annotation and revised genome sequence. J Bacteriol **192**:6319-6320.
24. **Hunter PR, Gaston MA.** 1988. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. J Clin Microbiol **26**:2465–2466.
25. **Castellanos E, de Juan L, Domínguez L, Aranaz A**. 2012. Progress in molecular typing of *Mycobacterium avium* subspecies *paratuberculosis*. Res Vet Sci **92**:169-179.
26. **Sevilla I, Garrido JM, Geijo M, Juste RA.** 2007. Pulsed-field gel electrophoresis profile homogeneity of *Mycobacterium avium* subsp. *paratuberculosis* isolates from cattle and heterogeneity of those from sheep and goats. BMC Microbiol **7**:18.
27. **Gagneux S, Small PM.** 2007. Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. Lancet Infect Dis **7**:328-337.
28. **Pearson T, Busch JD, Ravel J, Read TD, Rhoton SD, U'Ren JM, Simonson TS, Kachur SM, Leadem RR, Cardon ML, Van Ert MN, Huynh LY, Fraser CM, Keim P.** 2004. Phylogenetic discovery bias in *Bacillus anthracis* using single-nucleotide polymorphisms from whole-genome sequencing. Proc Natl Acad Sci USA **101**:13536-13541.
29. **Whittington RJ, Marsh IB, Choy E, Cousins D.** 1998. Polymorphisms in *IS*1311, an insertion sequence common to *Mycobacterium avium* and *M. avium* subsp. *paratuberculosis*, can be used to distinguish between and within these species. Mol Cell Probe **12**:349-358.
30. **Whittington RJ, Marsh IB, Whitlock RH.** 2001. Typing of *IS*1311 polymorphisms confirms that bison (*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 **15**:139-145.

Chapter VI

Molecular characterization of a rare pigmented *Mycobacterium avium* subsp. *paratuberculosis* Type C strain

"Genome wide comparison between *Mycobacterium avium* subsp. *paratuberculosis* non-pigmented and pigmented type C isolates"

Barbosa P, **Leão C**, Amaro A, Pinto C, Chaves I, Usié A, Botelho A, Stevenson K, Inácio J, Ramos AM

Manuscript with whole genome sequencing analysis in preparation

Molecular characterization of a rare pigmented *Mycobacterium avium* subsp. *paratuberculosis* Type C strain

6.1. Abstract

Mycobacterium avium subspecies *paratuberculosis* (*Map*) is the causative agent of paratuberculosis, affecting a wide range of animals, especially ruminants, with a considerable economic impact. *Map* strains can be classified in two major groups, Type C (or Type II) and Type S (or Type I/III). Type C are the most frequently isolated strains from a wide variety of hosts including cattle and humans, while Type S strains may have a preference for sheep and goats. Yellow pigmented strains from both Type I and Type III have been isolated from small ruminants from a few restricted geographic regions in the United Kingdom, Spain and Faroe Islands. Here we report the characterisation of a rare pigmented *Map* strain isolated from a goat faecal sample. Growth characteristics and F*57* real time PCR indicated that the pigmented isolate was a *Map* strain suspected to belong to Type C. By SNPs analysis it was characterized as Type C strain and SNP profile 3. Whole genome sequencing analysis is being performed and by the analysis of specific LSPs using Basic Local Alignment Search Tool it was possible to identify the presence of $LSP^{A}20$ and deletion 2 and the absence of $LSP^{A}4$-II, MAV-14, $LSP^{A}18$ and GPL cluster in the pigmented strain genome, characteristic results for Type C strains. This is the first report of a pigmented *Map* strain isolated in Portugal. Furthermore, it constitutes the first evidence of the existence of pigmented Type C strains.

Keywords

Map; paratuberculosis; pigmented strain; Type C; LSPs; SNPs

6.2. Introduction

Mycobacterium avium subspecies *paratuberculosis* (*Map*) is the causative agent of paratuberculosis, or Johne's disease. *Map* affects a wide range of animals especially ruminants, causing an intestinal granulomatous infection leading to considerable economic impact, due to weight loss of animals and diminution of milk production [1; 2]. *Map* strains can be classified in two major groups, Type C (or Type II) and Type S (or Type I/III) strains. Type C strains are commonly isolated from a wide variety of hosts including cattle and humans and are the most predominant strain type in cattle, while Type S strains are assumed to have host species preference for sheep and goats [3; 4; 5]. Type S strains can also be classified into two different groups, designated Type I and Type III, according to distinct pulsed-field gel electrophoresis (PFGE) genotyping profiles [6; 7]. Other molecular methodologies have been used for strains characterization namely IS*900*-RFLP, MIRU-VNTR, presence or absence of LSPs, gyr genes analysis, SNPs detection and more recently whole genome studies [5; 8].

Despite the genetic homology of >98%, *Map* Type C and Type S can be distinguished by growth characteristics in culture medium and genotyping assays. Type C strains can be easily isolated when compared with Type S, requiring up to 6 months of incubation, depending on the culture medium, while Type S strains can take up to one year or more for the growth of visible colonies. Yellow pigmented strains have been isolated from small ruminants from a few restricted geographic regions in the United Kingdom, Spain and Faroe Islands [3; 7]. The majority of pigmented isolates genotyped to date belong to Type I but pigmented Type III strain has been identified [4; 5; 8; 9]. Also, non-pigmented Type I strains have been isolated so pigmentation is not an exclusive characteristic of Type I strains [K Stevenson personal communication]. There is only one report of a pigmented strain isolated after 8-10 weeks of incubation, from a cow from Scotland, UK, in 1954 [10]. Unfortunately, the strain is not available for genotyping. Another case from Scotland of a cow with intestines showing orange pigmentation suggestive of infection by pigmented *Map* has been observed [L Petrie, personal communication].

Genotypically, *Map* strains can be distinguished based on specific regions of the genome and presence or absence of genomic markers such as Large Sequence Polymorphisms (LSPs) and Single Nucleotide Polymorphisms (SNPs) and by traditional typing tools. Whole genome sequencing (WGS) affords more in depth knowledge about the genetic diversity of organisms enabling elucidation of phylogenetic relationships. Recently, a WGS project was performed with 141 *Map* strains, that revealed the genetic diversity and phylogenomic relationships between isolates from distinct geographic regions around the world and from a variety of hosts [11]. With this analysis it was

possible to confirm the classification of *Map* into major strain types and to develop a new typing assay based on SNPs identification (see Chapter V) [11].

The aim of this work was to characterize and elucidate by molecular methods the strain type of a pigmented *Map* strain isolated from a goat faecal sample from Portugal, suspected to belong to Type C.

6.3. Material and Methods

6.3.1. Isolation and identification of pigmented colonies

A few small yellow pigmented colonies obtained from the culture of a goat faecal sample, were analysed. Pigmented colonies mixed with non-pigmented (Figure 6.1) were obtained on slants of both Herrold's egg yolk (HEYM) and Löwenstein–Jensen (LJ) media with mycobactin J (Chapter IV, set A samples), after 8-10 weeks incubation at 37ºC.

Suspected colonies were stained by auramine-rhodamine staining [12], and sub-cultured on solid culture media (HEYM and LJ with mycobactin J) for up to 6 months.

Isolated colonies were used to inoculate liquid culture medium Middlebrook 7H9 supplemented with 10% OADC and mycobactin J and incubated for up to 3 months at 37ºC.

6.3.2. DNA extraction

Two types of DNA extraction methods were performed. For heat treatment DNA extraction one colony from solid medium sub-culture was transferred to a microtube containing 100 µL of Tris EDTA buffer pH 8 (10 mM Tris.Cl, 1 mM EDTA) (TE buffer), mixed using a vortex mixer and centrifuged at 2000 × g for 5 minutes at room temperature. The pellet was resuspended in 100 µL of TE buffer pH 8 and incubated for 45 minutes at 95 ºC. The lysed bacterial suspension was centrifuged at 2000 × g for 1 minute at room temperature and the supernatant, containing the extracted DNA, was transferred to a new microtube and used directly as template in PCR reactions for confirmation of identity.

Genomic DNA from liquid sub-culture was extracted as previously described on Chapter V, section 5.3.2, using the QIAamp DNA Mini Kit (Qiagen) according to manufacturer's instructions with minor modifications. Briefly, the bacterial suspension was centrifuged at 5000 × g for 10 min and the pellet was resuspended in 180 µl of kit ATL buffer. Zirconium beads (0.1 mm) were added to the mixture and mechanical disruption of the cells was performed twice with a FastPrep FP120

Bio101 bead shaker (Savant Instruments Inc., Holbrook, NY) at 6.5 $msec^{-1}$ for 45 seconds. The disrupted mixture was cooled on ice for 15 min and 20 µl of proteinase K was added. The remaining procedure was completed according to manufacturer's instructions. The DNA was eluted with 200 µl of AE buffer and stored at -20ºC until required for PCR.

6.3.3. Identification and confirmation of pigmented isolates

For the identification and confirmation of isolates obtained from solid and liquid sub-cultures, F57-targeted real time PCR reactions were performed as previously described on Chapter IV, section 4.3.11. Briefly, a reaction mixture with a total volume of 20 µl containing 1× SSO Fast Super Mix (Bio-Rad), 0.4 µM of each *F57*-targeted primer and 0.15 µM of probe (Table 4.3, Chapter IV) and 5 µl of the extracted DNA template was prepared. DNase free water (GIBCO) was used as negative control. Thermal cycling, fluorescent data collection, and data analysis were performed in a CFX96 (Bio-Rad) detection system real time PCR instrument with the following conditions: 1 cycle at 95 °C for 2 minutes, followed by 45 cycles at 95 °C for 5 seconds and 60 °C for 10 seconds.

6.3.4. Molecular characterization

A novel SNP-based assay for genotyping of *Map,* based on PCR amplification coupled with restriction endonuclease analysis, was performed as described in Chapter V in order to clarify the strain type of the pigmented isolate. Amplicon obtained from the amplification of a fragment containing the snp3842359 and BsmBI restriction analysis were used to distinguish between Type C and Type S strains; snp343677 and AvaII were used to distinguish between Type I and Type III within the S Type; snp50173 and ApoI were used to differentiate isolates from Bison group from cattle group (phylogenetic sub-groups A and B); snp4111202 and Fat I were used to differentiate between phylogenetic sub-groups A and B. For the differentiation of the 10 sub-groups within sub-group A, ten SNPs were tested and the amplified products were sequenced: snp3879247; snp2939977; snp1932058; snp1327872; snp3844632; snp1966028; snp305277; snp4339946; snp2087274; snp1686154.

Sequencing of PCR products was carried out by StabVida, Lda. with the same primers used in the amplification of the fragments. The sequencing result was analysed using Basic Local Alignment Search Tool (Blast - http://blast.ncbi.nlm.nih.gov/Blast.cgi). The phylogenetic profile for each isolate was obtained by combining the results for all SNPs, as described at Chapter V.

Whole Genome Sequencing was performed for further complete genome analysis. An "in silico" analysis to search for LSPs specific of each strain type in the genome data was performed

using Basic Local Alignment Search Tool (Blast - http://blast.ncbi.nlm.nih.gov/Blast.cgi). Specific LSPs searched in this study present in Type C strains and absent from Type S genome included $LSP^{A}20$ and deletion 2 and LSPs present in Type S strains and absent from Type C are $LSP^{A}4$-II, MAV-14, $LSP^{A}18$ and GPL [13].

6.4. Results

6.4.1. Isolation, sub-culture and confirmation of identity of pigmented isolate

Yellow pigmented and non-pigmented colonies (Figure 6.1.A) were isolated from a goat faecal sample on HEYM and LJ with mycobactin J after 8-10 weeks of incubation.

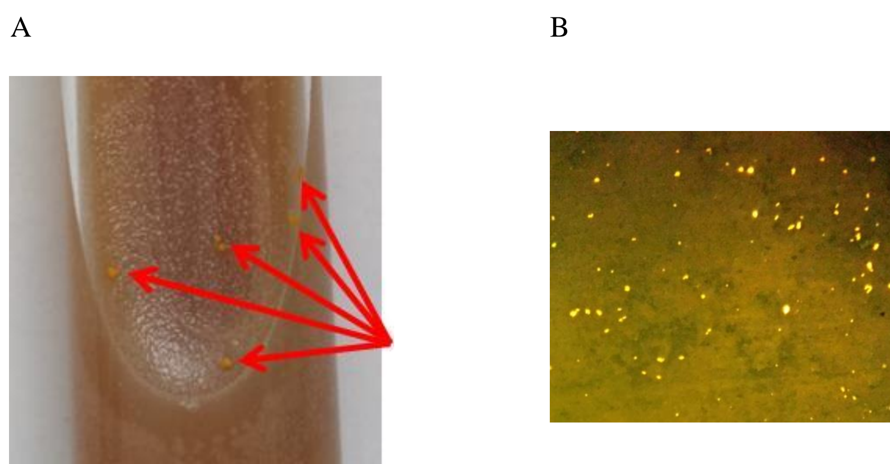

Figure 6.1. A - Image of the yellow pigmented (indicated by red arrows) mixed with non-pigmented *Map* colonies obtained from faecal culture after 8-10 weeks of incubation at 37ºC on HEYM with mycobactin J. **B** - Microscopical observation of the pigmented colonies stained by auramin-rhodamin procedure showing fluorescence characteristic of acid-fast bacilli.

The pigmented colonies were confirmed to be acid-fast bacilli by positive auramin-rhodamin staining (Figure 6.1.B). The same colonies were plated on solid medium to obtain pure cultures. A single colony from a pure culture was sub-cultured in liquid media and a yellow pigmented suspension was obtained after 2 months of incubation (Figure 6.2).

Figure 6.2. Image of the yellow pigmented *Map* strain suspension obtained from sub-culture in Middlebrook 7H9 supplemented with 10% OADC and mycobactin J after 2 months of incubation at 37ºC.

DNA was extracted from the pigmented sub-cultures on solid and liquid media and F*57* real time PCR was performed with positive amplification result (Figure 6.3).

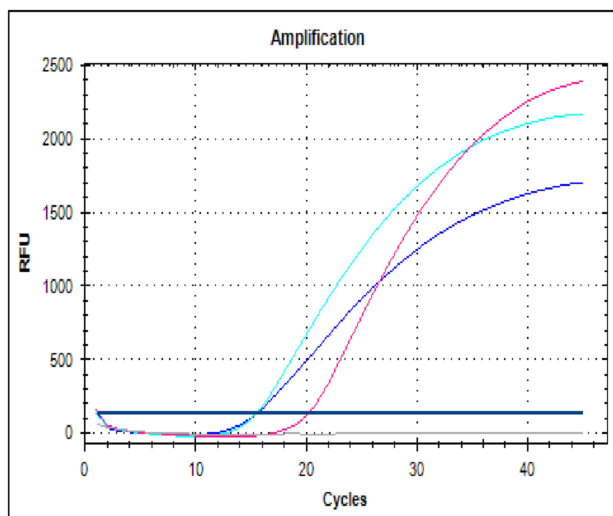

Figure 6.3. F*57* real time PCR result showing amplification for the F*57* specific *Map* marker.

Pink line - Positive control ATCC 19698^{T}; Dark blue line - Pigmented colony from the first isolation from faecal sample; Light blue line - Pigmented colony from sub-culture in Middlebrook 7H9; Grey line - Negative control.

6.4.2. Molecular characterization

A SNPs assay was performed to characterize the strain type of the pigmented isolate. The first steps of the assay were based on PCR amplification and analysis of the amplicon by restriction

endonuclease analysis according to a decision tree (see Chapter V). The obtained result is described in Table 6.1.

Table 6.1. Obtained results for the first steps of the decision tree of the SNP assay based on amplification and restriction endonuclease analysis.

SNP	Amplicon (bp)	Restriction Enzyme	Restriction products (bp)	Result
snp3842359	528	BsmBI	528	Type C
snp343677	-	-	-	-
snp50173	469	ApoI	277, 192	Cattle group
snp4111202	461	FatI	284, 177	group A

bp – base pairs; - not tested.

According to the decision tree, the pigmented isolate is a Type C strain, group A.

Ten additional SNPs were tested by amplification and sequencing of the amplified product in order to determine the SNP profile of the isolate. The pigmented strain was characterized as belonging to SNP profile 3, according to the characterization criteria indicated in Chapter V, section 5.4.1, table 5.3. The sequencing results of the ten SNPs are presented in Table 6.2.

Table 6.2. Results obtained by the analysis of the remaining 10 SNPs, by amplification and sequencing of the amplified product.

SNP	Amplicon (bp)	Sequencing result[1]
snp3879247	465	T
snp2939977	516	G
snp1932058	452	G
snp1327872	528	G
snp3844632	457	G
snp1966028	427	A
snp305277	461	G
snp4339946	464	T
snp2087274	453	C
snp1686154	525	G

[1] base at the SNP position in the genome obtained by sequencing of the amplified product.

WGS was performed in order to further characterize the pigmented isolate. This analysis is still on-going and future perspectives of the analysis are given in Chapter VII, section 7.2 of this dissertation.

With the genome data obtained by WGS, the presence and absence of LSPs were performed using Blast. LSP^A20 (Figure 6.4) and deletion 2 were found to be present with 100% and 99% of homology, respectively, with pigmented strain genome data. LSP^A18 (Figure 6.5), LSP^A4-II, MAV-14 and GPL were found to be absent from the pigmented strain genome data.

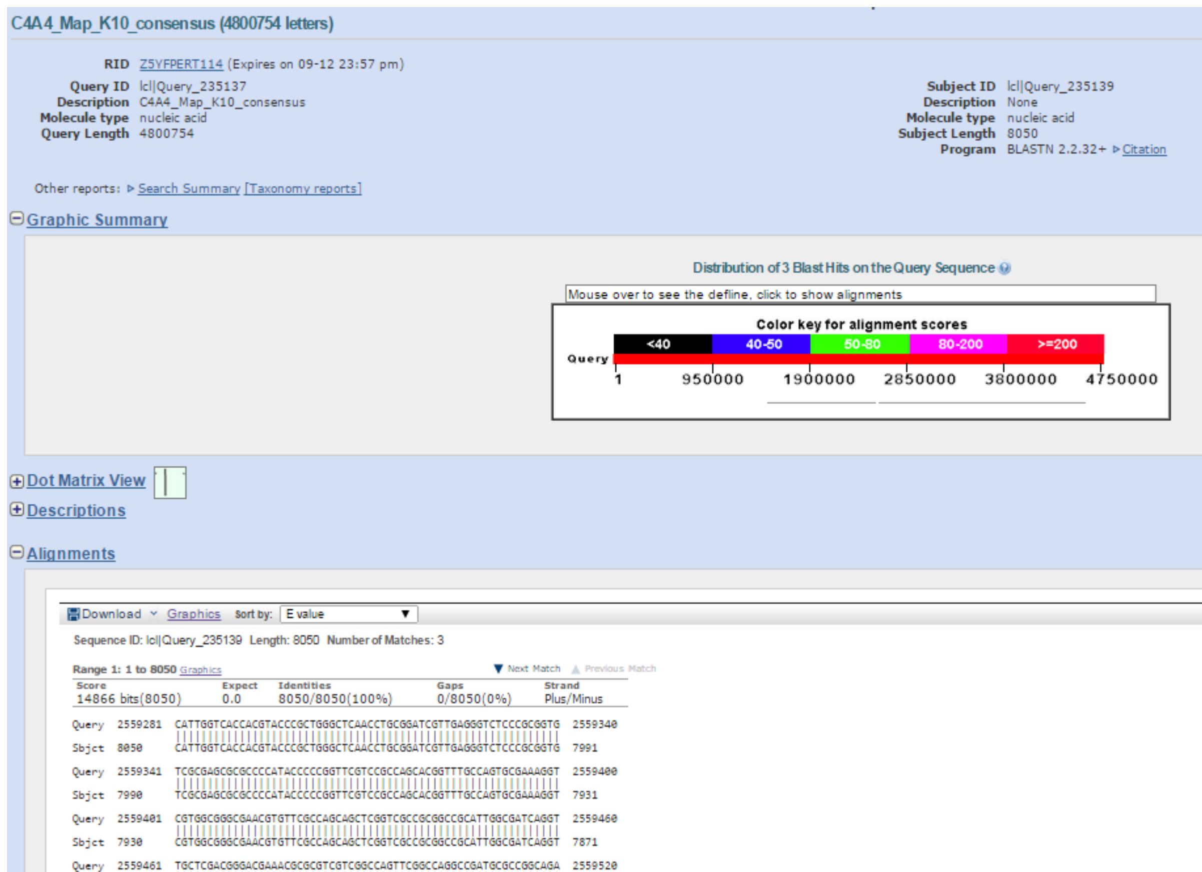

Figure 6.4. Image of the Blast analysis presenting 100% of homology of LSP^{A}20 (with 8050 bp), found to be present in the pigmented strain genome data.

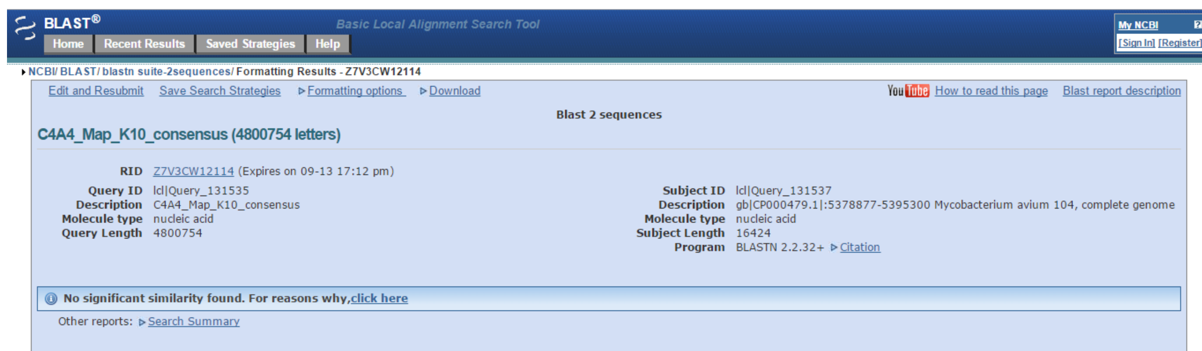

Figure 6.5. Image of the Blast analysis showing no similarity between LSP^{A}18 (with 16424 bp) with the pigmented strain genome data (indicating to be absent from the genome).

6.5. Discussion

Mycobacterium avium subspecies *paratuberculosis* (*Map*) strains were initially characterized by RFLP analysis and IS*900* DNA hybridisation [14] and were designated as Type C and Type S based on the species from which they were originally isolated, cattle and sheep, respectively. However, *Map* Type C strains can be isolated from a wide variety of hosts including cattle, humans, sheep and goats and Type S strains can also be found in cattle.

Subsequently, *Map* isolates were further characterized by PFGE and the two major strain groups were designated Type I (synonymous with Type S) and Type II (synonymous with Type C) to reduce the possible confusion between host origin and strain type [7]. A few years later, de Juan and colleagues (2005) [6], based on PFGE analysis, proposed the designation of Type III isolates considered as intermediate between Type I and Type II strains. For a decade it was assumed that the yellow pigmented *Map* isolates were exclusive to Type I. Then, in a study conducted by Biet and colleagues (2012) [8] a Type III pigmented ovine isolate was reported. Since then, non-pigmented Type I strains have been identified [K Stevenson, personal communication]. Pigment production does not therefore appear to be restricted to a particular strain type. However, there are no confirmed reports of Type C pigmented strains. In 1954 there was a study reporting a pigmented strain isolated in 8-10 weeks from a cow [10], from Scotland, but the strain is not available for genotyping. Another case of a cow from Scotland presenting intestines with orange pigmentation suggestive of infection by pigmented *Map* also has beed observed [L Petrie personal communication]. To date, no other pigmented *Map* strain has been associated with Type C strains.

In this study, the strain type of a yellow pigmented *Map* isolate from a goat was elucidated to be a rare Type C strain by SNPs assays and sequencing analysis.

From a goat faecal sample a few yellow pigmented colonies, mixed with white colonies, were obtained in HEYM with mycobactin J and on LJ with mycobactin J after 8-10 weeks of incubation at 37ºC. Based on the relatively short time to obtain culture, in comparison with Type S strains, it was immediately suspected that the yellow isolate might be a Type C strain. Auramin rhodamin staining and F57 real time PCR confirmed the identity of the yellow colonies as acid-fast bacilli and *Map*.

Using the SNPs assay developed in Chapter V, the pigmented isolate was found to belong to Type C, group A phylogenetic group and SNP profile 3.

Apparently a mixed infection of two different phenotypes of the same pathogen species *Map*, have occurred in this goat. Dozens of non-pigmented white colonies mixed with only two or three yellow pigmented colonies, isolated from faeces, classified the animal as a high shedder in agreement

with the observed clinical signs of paratuberculosis (weight loss and diarrhoea) and ELISA test positive.

Yellow pigmentation of colonies is associated with carotenoid pigmentation, natural pigments produced by a variety of bacteria, algae, fungi, animals and plants [15; 16]. Carotenoids are hydrophobic molecules found in a wide variety of microorganisms including *Streptomyces griseus* [17]; *Mycobacterium smegmatis* [18]; *Mycobacterium aurum* [19]; *Mycobacterium phlei* [20]; *Mycobacterium kansasii* [21]; *Mycobacterium tuberculosis* [22], and others, acting as antioxidative molecules, antimicrobial agents and associated with stress response [16]. Pigmentation has been linked to a higher virulence of pathogens by conferring resistance to stress conditions, like low pH and oxygen levels, and consequently more adaptability and survival capabilities in adverse environmental conditions [22]. Genetic sigma factors, such as sigF, present in *Mycobacterium tuberculosis* and other mycobacterial genomes, are responsible for stress responses and are associated with carotenoid production [17; 23].

WGS was performed in order to compare the genome of the pigmented strain with reference strains representative of Type C (*Map* K10) and Type S (*Map* S397) and clarify the strain type, but this analysis is still on-going. Using the sequencing data, specific LSPs characteristic of Type C strains were found to be present ($LSP^{A}20$ and deletion 2), while specific Type S LSPs were absent from the genome data of the pigmented isolate.

The virulence of the pigmented strain could not be assessed in this study as it was impossible to know if both strains infecting the goat were responsible for the animal's clinical situation or if one was more virulent than the other. Due to time constraints and funding, it was not possible to undertake a full genetic analysis of the genes that may be involved in pigmentation in *Map*. Future studies, proposed in Chapter VII, section 7.2, could compare pigmentation genes with other *Map* isolates and with other similar genes from other microorganisms to try to elucidate why some strains produce pigments while others do not, as well as virulence genes. However, a recent review was published [5] reporting that, based on WGS studies to date, no SNPs, insertions, deletions or the presence or absence of any single gene that could be exclusively associated with pigmented strains have been identified, suggesting that the underlying genetic basis for changes in this phenotype is likely to be complex and multi-factorial. For this reason, so far it is not clear why some *Map* strains produce yellow pigments and if there is any relationship with virulence level of the isolates contributing to the severity of the disease.

The growth characteristics and real time PCR amplification of F*57* fragment, the genetic homology between the pigmented isolate with *Map* K10 Type C genome with the identification of Type C specific LSPs, gave us enough confidence to support the initial suspicion that we isolated a

rare pigmented Type C *Map* strain, from a goat faecal sample from Portugal. More studies are being performed with WGS analysis.

This is a very significant discovery in the *Map* field as it is the first report of a pigmented *Map* strain isolated in Portugal and also the first report of a Type C pigmented strain. More studies are needed in order to elucidate pigmentation feature of some *Map* isolates and its possible relationship with virulence and pathogenesis.

Acknowledgements

Célia Leão is recipient of a PhD grant with reference SFRH/BD/62469/2009, from the Portuguese "Fundação para a Ciência e a Tecnologia". StabVida, Lda. and CEBAL are acknowledged for contributing to the whole genome sequencing and bioinformatics analysis, respectively.

Author's contribution

PB is contributing with the bioinformatics work and is writing the manuscript of WGS analysis. CL contributed to the designing of the study, experimental work of isolation and characterization of the strain, data analysis and writing this chapter. AA collaborated for the experimental work and data analysis. CP provided the faecal samples. AB, KS, JI, AMR contributed to the designing of the study. All authors have read and approved the final results.

6.6. References

1. **Motiwala AS, Strother M, Amonsin A, Byrum B, Naser SA, Stabel JR, Shulaw WP, Bannantine JP, Kapur V, Sreevatsan S**. 2003. Molecular epidemiology of *Mycobacterium avium* subsp. *paratuberculosis*: evidence for limited strain diversity, strain sharing, and identification of unique targets for diagnosis. J Clin Microbiol **41**:2015–2026.
2. **Timms VJ, Gehringer MM, Mitchell HM, Daskalopoulos G, Neilan BA**. 2011. How accurately can we detect *Mycobacterium avium* subsp. *paratuberculosis* infection? J Microbiol Meth **85**: 1-8.
3. **Salem M, Heydel C, El-Sayed A, Ahmed SA, Zschöck M, Baljer G**. 2012. *Mycobacterium avium* subspecies *paratuberculosis*: an insidious problem for the ruminant industry. Trop Anim Health Prod **45**:351-366.
4. **Stevenson K**. 2010. Comparative differences between strains of *Mycobacterium avium* subsp. *paratuberculosis*. *In* Behr MA, Collins DM (eds) Paratuberculosis: Organism, Disease, Control. CAB International, Cambridge, MA, pp 126-137.
5. **Stevenson K**. 2015. Genetic diversity of *Mycobacterium avium* subspecies *paratuberculosis* and the influence of strain type on infection and pathogenesis: a review. Vet Res **46**:64.
6. **de Juan L, Mateos A, Domínguez L, Sharp JM, Stevenson K**. 2005. Genetic diversity of *Mycobacterium avium* subspecies *paratuberculosis* isolates from goats detected by pulsed-field gel electrophoresis. Vet Microbiol **106**:249–257.
7. **Stevenson K, Hughes VM, de Juan L, Inglis NF, Wright F, Sharp JM**. 2002. Molecular characterization of pigmented and non-pigmented isolates of *Mycobacterium avium* subspecies *paratuberculosis*. J Clin Microbiol **40**:1798–1804.
8. **Biet F, Sevilla IA, Cochard T, Lefrançois LH, Garrido JM, Heron I, Juste RA, McLuckie J, Thibault V C, Supply P, Collins D, Behr M A, Stevenson K**. 2012. Inter and intra-subtype genotypic differences that differentiate *Mycobacterium avium* subsp. *paratuberculosis* strains. BMC Microbiol **12**:264.
9. **Sevilla I, Garrido JM, Geijo M, Juste RA**. 2007. Pulsed-field gel electrophoresis profile homogeneity of *Mycobacterium avium* subsp. *paratuberculosis* isolates from cattle and heterogeneity of those from sheep and goats. BMC Microbiol **7**:18.
10. **Watt JAA**. 1954. Johne's disease in a bovine associated with the pigmented strain of *Mycobacterium johnei*. Vet Rec **66**:387.
11. **Bryant JM, Thibault VC, Smith DGE, McLuckie J, Heron I, Sevilla IA, Biet F, Harris SR, Maskell DJ, Bentley SD, Parkhill J, Stevenson K**. 2015. Phylogenomic exploration of the relationships between strains of *Mycobacterium avium* subspecies *paratuberculosis.* BMC Genomics In press.
12. **Bird BB, Madison BM**. 2000. Use of fluorochrome staining for detecting acid-fast mycobacteria. In Current Laboaratory Practice Series. Department of Health and Human Services. Centers for Diseases Control and Prevention, Atlanta, pp 1-17.
13. **Alexander D C, Turenne C Y, Behr M A**. 2009. Insertion and deletion events that define the pathogen *Mycobacterium avium* subsp. *paratuberculosis*. J Bacteriol **191**:1018–1025.
14. **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.
15. **Mohammadi M, Burbank L, Roper MC**. 2012. Biological role of pigment production for the bacterial phytopathogen *Pantoea stewartii* subsp. *stewartii*. Appl Env Microbiol **78**: 6859-6865.
16. **Kirti K, Amita S, Priti S, Kumar AM, Jyoti S**. 2014. Colourful world of microbes: Carotenoids and their applications. Adv Biol 837891.
17. **Krügel H, Krubasik P, Weber K, Saluz HP, Sandmann G**. 1999. Functional analysis of genes from *Streptomyces griseus* involved in the synthesis of isorenieratene, a carotenoid with aromatic end groups, revealed a novel type of carotenoid desaturase. Biochim Biophys Acta **1439**:57-64.

18. **Hümpel A, Gebhard S, Cook GM, Berney M**. 2010. The SigF regulon in *Mycobacterium smegmatis* reveals roles in adaptation to stationary phase, heat, and oxidative stress. J Bacteriol **192:**2491–2502.
19. **Viveiros M, Krubasik P, Sandmann G, Houssaini-Iraqui M**. 2000. Structural and functional analysis of the gene cluster encoding carotenoid biosynthesis in Mycobacterium aurum A+. FEMS Microbiol Lett **187**:95-101.
20. **Goodwin TW, Jamikorn M**. 1955. Studies in Carotenogenesis. 17. The carotenoids produced by different strains of *Mycobacterium phlei*. Biochem J **62**:269-281.
21. **David HL**. 1974. Carotenoid Pigments of *Mycobacterium kansasii*. Appl Microbiol **28**:696-699.
22. **Saviola B**. 2014. Pigments and Pathogenesis. J Mycobac Dis **4**:5.
23. **Gebhard S, Hümpel A, McLellan AD, Cook GM**. 2009. The alternative sigma factor SigF of *Mycobacterium smegmatis* is required for survival of heat shock, acidic pH and oxidative stress. Microbiol **154**:2786–2795.

Chapter VII

Final Conclusions and Future perspectives

"I think and think for months and years.
Ninety-nine times, the conclusion is false.
The hundredth time I am right"

Albert Einstein

7.1. Conclusions

Mycobacterium avium Complex (MAC) are known as non-tuberculous, or atypical, mycobacteria and comprises microorganisms that affect a wide range of wildlife and livestock animals, including humans. These mycobacteria have a worldwide distribution and are widespread in the environment [1; 2; 3]. MAC includes nine species of slow-growing mycobacteria being *Mycobacterium intracellulare* and *Mycobacterium avium* (*M. avium*) subspecies the most commonly isolated from human and animal samples. *M. avium* is classified in four distinct subspecies: *Mycobacterium avium* subspecies *avium* (*Maa*), *Mycobacterium avium* subspecies *silvaticum* (*Mas*), *Mycobacterium avium* subspecies *hominissuis* (*Mah*) and *Mycobacterium avium* subspecies *paratuberculosis* (*Map*) [2; 3; 4] with *Mah* and *Map* as the most relevant as pathogens and consequently more studied among the four subspecies.

Mah has triggered a high impact on global public health concern, being considered as one of the most important member from MAC, regarding to its' opportunistic zoonotic potential in disseminated human infections namely in immunocompromised individuals [3]. Acquired Immunodeficiency Syndrome (AIDS) patients are the most affected humans, but *Mah* can also cause pulmonary infections and cervical lymphadenitis usually in children [1; 5; 6; 7]. *Mah* also affects pigs and had been isolated from other sources like birds, cattle, sheep, soil and water supplies, hypothetically one of the main infection routes for humans. However, the vehicle of infection between animals and humans still remain not clearly understood [3; 5; 8]. Epidemiological studies based on typing assays have been performed to assess the genetic characteristics of human, animal and environmental isolates in order to better understand the relationship between strains and elucidate sources of infection. A good typing method should be easy and rapid to perform, reproducible, robust, cheap, with a high level of discrimination and stability of the used DNA target and easy to standardize and to compare results among laboratories [9].

Genotyping of *Mah* has been achieved mainly by IS*1245*-RFLP, PFGE and MIRU-VNTR assays [6; 10; 11; 12; 13]. RFLP and PFGE have the disadvantages of requiring large amounts of pure culture and highly purified DNA, are technically demanding, expensive and hard to standardize [9; 12; 13; 14]. Those downsides lead to the need of development and standardization of more rapid, reliable, cheap and reproducible tools with high discriminatory powers. VNTR analysis has proven to be reproducible, rapid and simple to perform, with a good discriminatory power and with the possibility of digitalization of results allowing the comparison between laboratories. More recently, MATR-VNTR studies have emerged for the characterization of Japanese isolates from human, animal and environmental isolates based on the amplification of specific VNTR *loci* from the *M. avium*

genome [3; 5; 15; 16]. This technique has been reported to present a similar or even higher discriminatory power compared to IS*1245*-RFLP typing and MIRU-VNTR *loci* analysis [15].

The real prevalence of *Mah* in Portugal is not accurately known with only one study from 2009 reporting the characterization by IS*1245*-RFLP of *Mah* isolates from an outbreak occurred in Portugal from 2004 to 2008 [11]. In *Chapter II* of this dissertation, we analysed human and pig isolates, obtained during this outbreak by Multiple-Locus Variable number tandem repeat Analysis (MLVA). We tested a combination of 15 MIRU-VNTR and 5 MATR-VNTR *loci* to elucidate the genetic diversity of *Mah* isolates from distinct geographic regions of Portugal and to unravel epidemiological links between human and porcine isolates. Based on the highest VNTR allelic diversity indexes we selected the six more discriminative *loci* (MATR-3, MATR-6, MATR-7, MATR-8, MATR-11 and MATR-15) to further characterize isolates. Comparing with other reported studies using MATR-VNTR, we concluded that apparently the discriminatory power of MATR-VNTR *loci* changes between *Mah* populations isolated from distinct geographical regions. Therefore, an eventual simplification of the VNTR typing approach, by choosing a reduced number of the most discriminative *loci* for epidemiological studies in different geographic regions, must take into consideration the average diversity patterns of local *Mah* populations. The MLVA showed a high genetic diversity among *Mah* clinical isolates of human and porcine origins from Portugal. We also concluded that there were no correlation of *Mah* isolates with respective geographic origin, host and type of biological sample. Based on the majority of the molecular epidemiological studies of *Mah* isolates from European countries we can presume that common environmental sources are the most probable origin of infections for both pigs and humans and that bedding materials or feed, and/or international import/export markets of the animals contribute to the source of infection of pigs at global level [2; 12; 13; 17; 18; 19]. With this study we reported that MLVA assay is useful for global evaluation of the genetic diversity of *Mah* isolates with distinct genotypes randomly distributed across Portugal. This was the first report of the characterization of *Mah* isolates from Portugal using MATR-VNTR *loci*. We also projected the utility of MLVA to other countries and the implementation of a *Mah* worldwide VNTR profiles open-access database, allowing global epidemiological studies of this pathogen.

The second more significant and most studied MAC member is *Map*, the etiological agent of paratuberculosis, a chronic granulomatous enteritis affecting a wide range of animals. This mycobacteria is very challenging for researchers in veterinary field and animal producers in what concerns: strains characterization, due to the fastidious growth and special requirements for isolation; high economic impact worldwide, caused by a debilitation condition's of the infected animals, making this agent as one of the most important in animal health; and the possible zoonotic link to humans' bowel disease, known as Crohn's disease.

Paratuberculosis is not listed in European veterinary legislation list of mandatory reporting diseases and the animals do not need to be in quarantine when producers buy new animals. However, it is well known that stress conditions can significantly influence the intestinal microbial flora and consequently lead to the progression of the disease. Recently, it was reported that animals tested negative for paratuberculosis before moving to new farms became positive sometime after, probably triggered by the associated surrounding stress [20].

Map is transmitted mainly by faecal-oral route and can be shed in the environment by infected faeces, contributing to the spread of the disease. This agent has a large incubation period of time without development of clinical signs in infected animals, with the majority of them acting like asymptomatic carriers, but shedding intermittently the agent in faeces and milk spreading the disease to other animals [21; 22; 23]. It is suspected that for each farm with 1 to 5% of infected cows, 50% of the herd are asymptomatic shedders [22].

Diagnostic of paratuberculosis is difficult and time consuming due to the fastidious growth of the agent and the characteristics and stage of the disease. Once culture of *Map* is considered the "gold standard" diagnostic it can take several months to a year to generate a final diagnostic of the disease. For this reason, validation and standardization of faster and reliable alternative methodologies are needed in order to better characterize infected animals and to establish and improve control programmes for the disease. Molecular assays based on the detection of specific nucleic acids from bacterial genome are attractive and highly sensitive approaches and have been increasingly used helping to reduce the diagnostic time of fastidious microorganisms. However, as culture remains the "gold standard" of *Map* diagnostic, so far, the molecular testing needs to be complemented with culture results. *Map* specific fragments like IS*900* and *F57* are the most studied targets for PCR detection of *Map* from biological samples. Standard PCR system most widely used for detecting IS*900* was designed by Sanderson and colleagues (1992) [24] with primers P90/P91 and nowadays is still being used for a variable source of animal and human samples [25; 26; 27].

As paratuberculosis is not a mandatory reporting disease the real worldwide prevalence is not clearly understood and it is assumed that is a neglected and under-diagnosed disease in many countries. In Portugal only a few studies reporting the prevalence of paratuberculosis in cattle and

small ruminants are available [28; 29; 30; 31; 32; 33; 34]. In *Chapter III* we reported for the first time the isolation, identification and characterization of *Map* from asymptomatic cattle from a restricted region of the north of Portugal. For this study, we used standard PCR and culture of *Map* from faecal samples. Even with a limited number of samples in a pilot study, our preliminary data suggests that *Map* infection in cattle may be more prevalent in Portugal than initially expected. Despite the absence of clinical signs, our data points out that animals are shedding the agent in faeces, with the possibility of being a passive intermittent shedding, perpetuating the cycle of infection. We also typed for the first time strains isolated from asymptomatic cattle from Portugal by MLST with the evidence of a clonal infection by INMV2 strains, according to the classification of Thibault and colleagues (2007) [35], one of the most common INMV type widespread in Europe.

PCR discovery has revolutionized the scientific research and diagnostic fields by nucleic acid detection and characterization. Despite the faster technology comparing to traditional diagnostic methodologies, standard procedures of PCR presents some limitations. Standard PCR is considered an “end-point” assay, requiring electrophoretic separation of the amplified products on an agarose gel stained with ethidium bromide or similar fluorescent dyes, making it a time consuming assay, expensive all over the time and presenting risk of cross-contamination of samples by manipulating open tubes. This assay also presents a difficult quantification of DNA, based on the fluorescence of the obtained bands, and low sensitivity, due to the required amount of stained DNA to be detected by fluorescence in the agarose gel, leading to possible false-negatives results [36; 37; 38]. Nested PCR can overcome the issue of the low amount of DNA template in samples by the use of two PCR amplification steps targeting the same DNA region, increasing the number of amplicon copies in the end of reaction, but requiring additional good laboratory practice standard conditions to avoid cross-contamination [39; 40]. Real time PCR was developed in order to monitor the reaction in real time by the observation of increasing fluorescent in each amplification cycle, excluding the “end-point” detection and consequent cross-contamination risk. This approach also allows the detection of less quantity of DNA in the reaction and the reduction of detection time by the shorter duration of each amplification cycle. In conclusion, real time PCR presents good advantages for diagnostics research field: is more sensitive, fast, robust, allows the quantification of nucleic acids and the reaction is performed in a close tube minimizing cross-contamination risks [36; 37; 39; 41].

One particular concern from our study reported in *Chapter III* was the observation of a non-specific multi band pattern in agarose gel after the amplification by standard PCR of the *Map* specific IS*900* target. This situation was specially observed when the direct detection of *Map* DNA was performed in complex biological matrices, like faeces, regarding diverse DNA targets present that can originate non-specific amplification fragments with different sizes. For this reason, more specific and sensitive assays are needed to better characterize the presence of *Map* in biological samples, including optimized DNA extraction methodologies. DNA extraction is considered a critical step for molecular

diagnostics tools, where highly purified DNA is required contributing to the specificity and detection limits of the molecular assay. PCR is being increasingly used as new “gold standard” in some scientific areas specially virology [41], however, for *Map* detection PCR was considered less sensitive than culture of the agent [42] requiring more improvements.

Therefore, in *Chapter IV* we optimized new DNA extraction methodologies for the detection of *Map* directly from faecal and milk samples. Faeces and milk are known to be challenging biological matrices for the molecular detection due to the presence of inhibitors [22]. Faecal PCR inhibitors included phytic acid and polysaccharides, and large amounts of nucleic acids from other bacteria and host cells [43; 44], while milk have large quantity of fat and calcium ions [45]. Due to *Map* cell wall composition, these bacteria are preferentially located in the fat fraction of milk [46; 47]. Commercial DNA extraction kits and preliminary steps for *Map* cell concentration and lysis from samples were tested with faecal *Map* culture positive samples and experimentally infected milk. Different amounts of faeces and milk were analysed with and without mechanical disruption steps and different enzymatic lysis incubation times. The most efficient extraction method, detecting a higher number of PCR positive samples was Invisorb® Spin Tissue Mini Kit with mechanical (bead beating) disruption, commonly considered as an essential step for DNA extraction from *Map* cells [43; 44]. Twelve hours of enzymatic incubation lysis step was adopted and, in case of faeces, the extraction was performed in 1g instead of 25 mg recommended by the manufacturer of the extraction kit, while milk samples were subjected to an initial sample preparation procedure described by Gao and colleagues (2007) [46] in order to recover the maximum quantity of *Map* cells from the fat layer.

We also developed a new IS*900* nested real time PCR, with the combination of two amplification steps, the first step by conventional PCR being the amplified product used as template for the second amplification step by real time PCR. With this approach the percentage of positives of the assay increased from 44% to 83% in faecal samples tested by IS*900* real time PCR and IS*900* nested real time PCR, respectively. It was not possible to calculate this parameter for milk samples due to the lack of culture positive results. With our IS*900* nested real time PCR the limit of detection (LOD) obtained from spiked with *Map* Type C faecal samples was 10 cells per gram of faeces while for *Map* Type S the LOD was 100 cells per gram of faeces. The LOD milk spiked with *Map* Type C was 100 cells per millilitre of milk. We also described the development of a novel real time PCR for the detection of a specific *F57* fragment with 100% of specificity and a very high sensitivity with a LOD of one genome copy in the reaction mixture. With this study we aimed to highlight the presence of *Map* in faecal and milk samples from different animals from distinct geographic regions from Portugal and the development of novel TaqMan-based real time PCR assays, with high sensitivity and specificity, associated with improved sample preparation to concentrate the agent and to reduce PCR inhibitors, shortening the time for confirmatory *Map ante mortem* diagnosis. The availability of *ante mortem* tests allowing a fast and conclusive biological samples testing from live animals offer a great

advantage in improving the efficiency of monitoring and control programs and decreasing the associated economic burden. Actually, the current European control programs are still mainly voluntary acts conducted by producers [20].

According to growth characteristics and host preference *Map* isolates were initially classified in two different groups designated Type C (C from cattle) and Type S (S from sheep). Collins and colleagues (1990) [48] observed distinct PCR-REA patterns obtained from those two major groups isolated from different animals and proved the existence of genetic differences between Type C and Type S. However, Type C strains can infect a wide range of animals and also humans while Type S strains are isolated from a restrictive variety of hosts suggesting to have preference by sheep and goats. Different typing assays started being used as useful tools for strains characterization and in 2002, Stevenson and colleagues [49] proposed a new designation for *Map* strains to avoid confusions with host association (Type I for Type S strains and Type II for Type C strains), based on PFGE and RFLP analysis. Other strains Type designations were proposed as Type III, suggested as an intermediate group of strains between Types I and II but more related with Type I [50] and Bison Type, genetically related with Type II [51]. The characterization of Bison strains have been performed by IS*1311*-REA but had presented some issues. MIRU-VNTR and presence or absence of specific LSPs are the most widely used typing assays for *Map* strains differentiation.

Genetic studies with the analysis of whole genome sequences can overcome concerns with strains characterization and contributes to the improvement of knowledge about phylogenetic relationships between strains providing a better understanding of *Map* genetics, evolution, virulence and possibly helping in the elucidation of *Map* zoonotic potential [52]. Whole Genome Sequencing (WGS) by next-generation sequencing (NGS) technologies have become widely available with a more reduced cost providing the opportunity of a large-scale DNA analysis for researchers. NGS offers a more complete and reliable genome analysis than traditional sequencing and typing assays [52; 53; 54; 55].

For a decade only one complete genome sequence from *Map* from strain Type C (*Map* K10) [56] was available in the public domain restringing the evolution of genetic studies by absence of more reference genomes sequences and specially genomes representative of all distinct strain types. In the last few years, other *Map* WGS studies have emerged namely: a study from Hsu and colleagues (2011) [54] reporting the sequencing of six genomes from strains isolated from cattle, oryx, goat, human, red deer and environmental samples; Wynne and colleagues (2011) [52] reported the sequencing of ten and three genomes from isolates of human and animal origin, respectively; Bannantine and colleagues (2012) [57] assembled the complete genome sequence of an ovine Type S strain by sequencing and optical mapping; Ghosh and colleagues (2012) [58] described the characterization and the whole genome analysis of two isolates from camel samples; and Singh and colleagues (2013) [59] performed the annotation of the complete genome sequence of an Indian Bison

Type strain. Recently, two WGS projects comprising the sequencing of 124 Canadian isolates and 144 *Map* isolates from distinct hosts from different geographic regions all over the world were performed to better understand the genetic of *Map* isolates and the utility of MIRU-VNTR and IS*1311*-REA was evaluated [60; 61]. Ahlstrom and colleagues (2015) [60] reported the WGS analysis of 124 Canadian isolates with the identification of over 3000 SNPs between all isolates divided into eight phylogenetic sub-groups with the presence of a dominant subtype. The utility of MIRU-VNTR typing assay was evaluated suggesting that this technique may lead to incorrect epidemiological conclusions based on the observation of different phylogenetic relationships between strains presenting the same MIRU-VNTR profile. This conclusion was corroborated in the second study conducted by Bryant and colleagues [61] with the observation of phylogenetic rearrangements between strains belonging to the same traditional typing profile more closely related with other strains with distinct profiles. No phylogenetic relationship between strains from the same typing profile, host origin or geographic region was observed. More than 40000 SNPs were identified between all the 144 sequenced isolates genomes and the existence of two major groups was confirmed, Type C and Type S. Type I and Type III are phylogenetic sublineages of Type S and Bison Type is a sublineage of Type C. It was also highlighted the lack of robustness of the traditional IS*1311* PCR-REA for the distinction of strains being necessary to develop new improved tools for this characterization [61].

In this context, in Ch*apter V* we developed a novel SNP-based assay for the characterization of *Map* isolates based on the analysis of the WGS data of 133 *Map* strains from the 144 genomes previously studied by Bryant and colleagues [61]. A total of 28402 SNPs were identified among all isolates being 93 and 41 the minimum required to distinguish between all genomes and between only the Type C isolates, respectively. The novel assay was developed based on sequential detection of 14 SNPs and a decision tree distinguishing 14 phylogenetic groups with a higher discriminatory power compared to MIRU-VNTR assay and other typing methods. This novel assay can overcome some issues regarding the genotyping of isolates characterized as Type I, Type III and Bison type. We described one SNP where in the same position of the genome Type I strains have a “G”, Type II have an “A” and Type III have a “C”. We also identified another SNP and the corresponding restriction endonuclease ApoI for PCR-REA SNP detection that could provide an easier, alternative assay for discriminating Bison-type strains from other Type C strains. These are improvements comparing with other previously reported typing systems. Our novel SNP assay is a discriminative, simple and reproducible method, applicable to a small amount of genomic DNA and based on standard PCR and restriction endonuclease analysis, requiring only basic laboratory equipment for a large-scale global typing of *Map* isolates. However, continuous updating of genome sequences is needed in order to better characterize new phylogenetic groups and SNP profiles.

During the work described in this dissertation, a total of 74 *Map* isolates were obtained from faecal samples: 12 from asymptomatic cattle; and 62 from goats (N=15) and cattle (N=47), with

suspicious disease. One of the goat samples presented a mixture of several non-pigmented colonies along with a few yellow pigmented ones. The reduced 8-10 weeks incubation time and the identification of these yellow isolates as *Map* constituted a focus of interest for their further characterization. Pigmented *Map* strains have been characterized as Type S belonging to both Type I and Type III [61], isolated from small ruminants, with only one case reported in 1954 about the isolation from a cow of a pigmented strain, after 8-10 weeks of incubation, suggestive to be a Type C strain [62]. Since then, no further evidences of the existence of pigmented Type C strains were described. For this reason, in *Chapter VI* we performed the study and molecular characterization of the pigmented colonies in order to clarify the classification of that isolate. The first steps of the study were the identification of the yellow pigmented colonies by auramine-rhodamin staining, with observation of acid-fast bacilli, and by *F57*-targeted real time PCR system described in *Chapter IV,* section 4.3.11, with confirmation as *Map*, the isolate was further characterized with the novel SNPs-based assay described in *Chapter V* revealing that it belongs to Type C, group A, SNP profile 3, according to our phylogenetic study. In order to deeper characterize this pigmented strain a WGS analysis was performed. The genome of a non-pigmented Type C *Map* strain, isolated from asymptomatic cattle described in *Chapter III*, was also sequenced for genome comparison. With the WGS data, $LSP^{A}20$ and deletion 2 were identified in both sequenced strains genomes. The annotation and mapping of the genome data from our isolates with both reference Type C (*Map* K10) and Type S (*Map* S397) strains showed a higher genetic similarity with *Map* K10. This fact confirms that our pigmented and non-pigmented *Map* strains belong to Type C. We reported for the first time a pigmented *Map* strain isolated in Portugal and a first evidence of the existence of pigmented Type C strains, since 1954, this time isolated from a goat.

Summing up in this work, the following conclusions were reached:

- Portuguese *Mah* isolates of human and porcine origins characterized by MLVA, revealed a high heterogenic *Mah* population with no genotype correlation with host, type of sample, or geographic origin.
- Presence of *Map* in faeces from asymptomatic cattle, evaluated in a small scale study by traditional molecular assays and culture, constitutes an awareness for the existence of paratuberculosis in unsuspected farms.
- For faecal and milk samples more sensitive and specific IS*900* nested real time PCR, associated with optimized extraction DNA methodologies, should be used in the future since it revealed a higher robustness and sensitivity, compared with standard procedures, enabling a reliable *ante mortem* direct diagnostic of paratuberculosis. A novel *F57* real time PCR can be applied to pure culture isolates, obtained in standard culture, constituting a more specific and sensitive alternative to the identification of *Map*.

- The novel typing assay described for *Map* strains characterization based on SNPs present in the whole genome is a valuable contribution to the improvement of simple, rapid and cheap typing tools.
- The isolation and characterization of a rare pigmented *Map* Type C strain, isolated from a goat, revealed and confirmed that pigmented strains can also belong to Type C and not only Type S group.

All together the work described in this dissertation, constitutes an improvement in paratuberculosis diagnostic tools, by direct molecular detection of *Map* in live animals; stakeholders awareness for the existence and widespread of unsuspected *Map* animal infections in Portugal; a contributibution for refine and alternative molecular tools and markers to unveil the epidemiological traits of *Mah* and *Map* infections enabling implementation of monitoring and control measures.

7.2. Future perspectives

With the findings described in this work we aimed to contribute to a better understanding of epidemiological traits of MAC infections, namely caused by *Mah* and *Map* strains, affecting the respiratory and the intestinal tract of both humans and animals. In this perspective, we intend to leave some viewpoints and contributions for future work highlighting the importance of research studies in animals and humans health areas. The real prevalence of MAC infections in Portugal is not well known and for that reasons more and large scale studies are needed to characterize the real situation in our country. It is only possible to implement disease control measures if the real situation is clearly understood. Therefore, we want to advise for the need of *Mah* studies using a larger number of isolates from human, animal and environmental samples in order to extend and perform a more representative study of *Mah* infections and evaluate genetic relationships between isolates from a wide range of sources to elucidate transmission routes and epidemiology of the disease.

Another relevant topic is what concerns to *Map*. We described the detection of *Map* from faeces from asymptomatic bovine from a restricted region of the north of Portugal and from goats and cattle from Azores, Portugal. We also evaluated milk samples from 16 Portuguese counties and from Azores, Portugal, with positive *Map* DNA detection. Our results opened new doors to the evidence of the presence of *Map* in Portuguese farms and that even with the absence of clinical signs Portuguese cattle is shedding the agent perpetuating the spread of the disease. These data has sensitized us for the fact that the disease is underdiagnosed in Portugal in an unknown scale contributing to an unidentified economic impact for producers and for the country, apart from the possible contribution for increasing the risk of human disease. For this purpose, we consider very relevant to boost the research in *Map* infections by testing a higher number and variety of livestock, as well as wildlife animals spread in our country. We also consider that the awareness of producers and nation for this disease and correspondent impact in economy is a crucial step to the implementation of preventing measures and control programs for paratuberculosis, as already exist for other infections.

Another interesting point is contributing for a better and rapid diagnostic of *Map* from live animals. Nowadays, there is no standardized molecular method for *Map* detection with distinct researcher groups and routine laboratories using their own in-house methodologies. In the Portuguese reference laboratory for animal diseases *Map* diagnostic is currently performed with more traditional assays based on ELISA, CF, AGID, AccuProbe® Culture Identification Tests (Gen-Probe), INNO-LIPA® Mycobacteria V2 (Innogenetics), Ziehl-Nielsen and culture. In our work we improved the DNA extraction from faeces and milk samples and real time PCR approaches with high sensitivity and sensibility which may contribute significantly to the improvement of routine diagnostic.

It was not possible in our current work, to optimize novel real time PCR systems for the distinction of different *Map* strains Type (I, II and III) and another issue that was not possible to perform was the optimization of multiplex real time PCR for the identification of *Mycobacterium* genus, MAC and *M. avium* subspecies, however, primers and TaqMan probes were designed and tested for this purpose and may constitute the aim of future work.

Regarding WGS projects, in the first study we have contributed to the availability of *Map* complete genomes and phylogenetic studies. However, more efforts are needed to better understand the genetic of *Map* populations. As future work we suggest the use of the novel SNP-based assay in a larger set of *Map* isolates. The improvement of the assay, with more WGS data and the inclusion of more SNPs, is needed to better characterize novel phylogenetic groups and attribution of new SNPs profiles designation enabling the creation of a *Map* SNPs profile open-access database. With the second study we confirmed the strain Type of a pigmented *Map* isolate based on genome comparison with reference *Map* strains. It was not possible in this work to maximise the analysis of the obtained WGS data from our isolates. However, genetic studies are still being performed to search for virulence and pigmentation genes to better characterize and explain the difference between pigmented and non-pigmented strains. Hereafter, it would be interesting to extensively explore genetic particularities between strains by comparing our sequencing data with other 280 *Map* sequenced genomes obtained in the last few years with the evolution of NGS studies. WGS studies provide massive quantities of data allowing more and novel opportunities of extensive genetic future works of populations and epidemiological links of diseases.

7.3. References

1. **Biet F, Boschiroli ML, Thorel MF, Guilloteau LA**. 2005. Zoonotic aspects of *Mycobacterium bovis* and *Mycobacterium avium-intracellulare* complex (MAC). Vet Res **36**:411–436.
2. **Turenne CY, Wallace R Jr, Behr MA**. 2007. *Mycobacterium avium* in the postgenomic era. Clin Microbiol Rev **20**:205–229.
3. **Rindi L, Garzelli C**. 2014. Genetic diversity and phylogeny of *Mycobacterium avium.* Infect Genet Evol **21**:375–383.
4. **Shin SJ, LeeB S, Koh W, Manning EJB, Anklam KA, Sreevatsan S, Lambrecht RS, Collins MT**. 2010. Efficient Differentiation of *Mycobacterium avium* Complex Species and Subspecies by Use of Five-Target Multiplex PCR. J Clin Microbiol **48**:4057–4062.
5. **Iwamoto T, Nakajima C, Nishiuchi Y, Kato T, Yoshida S, Nakanishi N, Tamaru A, Tamura Y, Suzuki Y, Nasu M**. 2011. Genetic diversity of *Mycobacterium avium* subsp. *hominissuis* strains isolated from humans, pigs, and human living environment. Infect Genet Evol **12**:846-852.
6. **Kaevska M, Slana I, Kralik P, Reischl U, Orosova J, Holcikova A, Pavlik I**. 2011. "*Mycobacterium avium* subsp. *hominissuis*" in neck lymph nodes of children and their environment examined by culture and triplex quantitative real-time PCR. J Clin Microbiol **49**:167–172.
7. **Agdestein A, Johansen TB, Kolbjørnsen O, Jørgensen A, Djønne B, Olsen I**. 2012. A comparative study of *Mycobacterium avium* subsp. *avium* and *Mycobacterium avium* subsp. *hominissuis* in experimentally infected pigs. BMC Vet Res **8**:11.
8. **Cayrou C, Turenne C, Behr MA, Drancourt M**. 2010. Genotyping of *Mycobacterium avium* complex organisms using multispacer sequence typing. Microbiol **156**:687–694.
9. **Jagielski T, Ingen J van, Rastogi N, Dziadek J, Mazur P K, Bielecki J**. 2014. Current methods in the molecular typing of *Mycobacterium tuberculosis* and other mycobacteria. Biomed Res Int Article ID 645802.
10. **Mobius P, Lentzsch P, Moser I, Naumann L, Martin G, Kohler H**. 2006. Comparative macrorestriction and RFLP analysis of *Mycobacterium avium* subsp. *avium* and *Mycobacterium avium* subsp. *hominissuis* isolates from man, pig, and cattle. Vet Microbiol **117**:284–291.
11. **Domingos D, Amado A, Botelho A**. 2009. IS1245 RFLP analysis of strains of *Mycobacterium avium* subspecies *hominissuis* isolated from pigs with tuberculosis lymphadenitis in Portugal. Vet Rec **164**:116-120.
12. **Pate M, Kušar D, Zǒlnir-Dovč M, Ocepek M**. 2010. MIRU–VNTR typing of *Mycobacterium avium* in animals and humans: Heterogeneity of *Mycobacterium avium* subsp. *hominissuis* versus homogeneity of *Mycobacterium avium* subsp. *avium* strains. Res Vet Sci **91**:376-81.
13. **Radomski N, Thibault VC, Karoui C, de Cruz K, Cochard T, Gutiérrez C, Supply P, Biet F, Boschiroli ML**. 2010. Determination of genotypic diversity of *Mycobacterium avium* subspecies from human and animal origins by mycobacterial interspersed repetitive-unit-variable-number tandem-repeat and *IS*1311 restriction fragment length polymorphism typing methods. J Clin Microbiol **48**:1026-1034.
14. **Sevilla I, Garrido JM, Geijo M, Juste RA**. 2007. Pulsed-field gel electrophoresis profile homogeneity of *Mycobacterium avium* subsp. *paratuberculosis* isolates from cattle and heterogeneity of those from sheep and goats. BMC Microbiol **7**:18.
15. **Inagaki T, Nishimori K, Yagi T, Ichikawa K, Moriyama M, Nakagawa T, Shibayama T, Uchiya K, Nikai T, Ogawa K**. 2009. Comparison of a variable number tandem-repeat (VNTR) method for typing *Mycobacterium avium* with mycobacterial interspersed repetitive-unit-VNTR and IS1245 restriction fragment length polymorphism typing. J Clin Microbiol **47**:2156–2164.

16. **Tatano Y, Sano C, Yasumoto K, Shimizu T, Sato K, Nishimori K, Matsumoto T, Yano S, Takeyama H, Tomioka H**. 2012. Correlation between variable-number tandem-repeat-based genotypes and drug susceptibility in *Mycobacterium avium* isolates. Eur J Clin Microbiol Infect Dis **31**:445-454.
17. **Tirkkonen T, Pakarinen J, Rintala E, Ali-Vehmas T, Marttila H, Peltoniemi OAT, Mäkinen J**. 2010. Comparison of variable-number tandem-repeat markers typing and *IS*1245 restriction fragment length polymorphism fingerprinting of *Mycobacterium avium* subsp. *hominissuis* from human and porcine origins. Acta Vet Scand **52**:21.
18. **Álvarez J, Castellanos E, Romero B, Aranaz A, Bezos J, Rodríguez S, Mateos A, Domínguez L, de Juan L**. 2011. Epidemiological investigation of a *Mycobacterium avium* subsp. *hominissuis* outbreak in swine. Epidemiol Infect **139**:143–148.
19. **Johansen TB, Agdestein A, Lium B, Jørgensen A, Djønne B.** 2014. *Mycobacterium avium* subsp. *hominissuis* infection in Swine associated with peat used for bedding. Biomed Res Int **2014**:189649.
20. **Pribylova-Dziedzinska R, Slana I, Lamka J, Pavlik I**. 2014. Influence of stress connected with moving to a new farm on potentially MAP-infected Mouflons. ISRN Microbiol **2014**: 450130.
21. **Sweeney RW, Whitlock RH, Rosenberger AE**. 1992. *Mycobacterium paratuberculosis* cultured from milk and supramammary lymph nodes of infected asymptomatic Cows. J Clin Microbiol **30**:166-171.
22. **Salem M, Heydel C, El-Sayed A, Ahmed S A, Zschöck M, Baljer G**. 2012. *Mycobacterium avium* subspecies *paratuberculosis*: an insidious problem for the ruminant industry. Trop Anim Health Prod **45**:351-366.
23. **Atreya R, Bülte M, Gerlach G F, Goethe R, Hornef M H, Köhler H, Meens J, Möbius P, Roeb E, Weiss S**, on behalf of the ZooMAP Consortium. 2014. Facts, myths and hypotheses on zoonotic nature of *Mycobacterium avium* subspecies *paratuberculosis*. Int J Med Microbiol **304**:858-867.
24. **Sanderson JD, Moss MT, Tizard ML, Hermon-Taylor J**. 1992. *Mycobacterium paratuberculosis* DNA in Crohn's disease tissue. Gut **33**:890–896.
25. **Whittington RJ, Taragel CA, Ottaway S, Marsh I, Seaman J, Fridriksdottir V**. 2001a. 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.
26. **Khare S, Ficht TA, Santos RL, Romano J, Ficht AR, Zhang S, Grant IR, Libal M, Hunter D, Adams LG**. 2004. Rapid and sensitive detection of *Mycobacterium avium* subsp. *paratuberculosis* in bovine milk and feces by a combination of immunomagnetic bead separation-conventional PCR and real-time PCR. J Clin Microbiol **42**:1075–1081.
27. **Naser SA, Sagramsingh SR, Naser AS, Thanigachalam S**. 2014. *Mycobacterium avium* subspecies *paratuberculosis* causes Crohn's disease in some inflammatory bowel disease patients. World J Gastroenterol **20**:7403-7415.
28. **Mendes S, Boinas F, Albuquerque T, Fernandes L, Afonso A, Amado A**. 2004. Epidemiological studies on paratuberculosis in small ruminants in Portugal. Epidémiol et santé anim **45**:61-71.
29. **Coelho AC, Pinto ML, Silva S, Coelho AM, Rodrigues J, Juste RA**. 2007. Seroprevalence of ovine paratuberculosis infection in the Northeast of Portugal. Small Ruminant Res **71**:298-303.
30. **Coelho AC, Pinto ML, Coelho AM, Rodrigues J, Juste R**. 2008. Estimation of the prevalence of *Mycobacterium avium* subsp. *paratuberculosis* by PCR in sheep blood. Small Ruminant Res **76**:201-206.
31. **Coelho AC, Pinto ML, Coelho AM, Rodrigues J**. 2009. Comparação de duas técnicas de isolamento do *Mycobacterium avium* subsp. *paratuberculosis* em amostras de fezes de ovinos com suspeita clínica de paratuberculose. Pesq Vet Bras **29**:415-420.
32. **Coelho AC, Pinto ML, Coelho AM, Aires A, Rodrigues J**. 2010. A seroepidemiological survey of *Mycobacterium avium* subsp. *paratuberculosis* in sheep from North of Portugal. Pesq Vet Bras **30**:903-908.

33. **Vala H, Santos C, Esteves F, Albuquerque T, Afonso A, Botelho A, Seixas C, Amaral M, Amado A**. 2007. Paratuberculosis in sheep from Serra da Estrela Region, Portugal. Proceedings of the 9th International Colloquium on Paratuberculosis, Japan. Epidemiology and Control Strategies 250-253.
34. **Martinho A, Campos N, Rodrigues P, Magro F**. 2010. *Mycobacterium avium* subsp *paratuberculosis* from blood of Crohn's disease patients in a Portuguese Hospital. RFCS **7**:74-79.
35. **Thibault VC, Grayon M, Boschiroli ML, Hubbans C, Overduin P, Stevenson K, Gutierrez M C, Supply P, Biet F**. 2007. New variable-number tandem-repeats markers for typing *Mycobacterium avium* subsp. *paratuberculosis* and *M. avium* strains: comparison with IS*900* and IS*1245* restriction fragment length polymorphism typing. J Clin Microbiol **45**:2404-2410.
36. **Heid CA, Stevens J, Livak KJ, Williams PM**. 1996. Real time quantitative PCR. Genome Res **196**:986-994.
37. **Valasek MA, Repa JJ**. 2005. The power of real-time PCR. Adv Physiol Educ **29**:151–159.
38. **Parashar D, Chauhan DS, Sharma VD, Katoch VM**. 2006. Applications of real-time PCR technology to mycobacterial research. Indian J Med Res **124**:385–398.
39. **Neonakis IK, Gitti Z, Krambovitis E, Spandidos DA**. 2008. Molecular diagnostic tools in mycobacteriology. J Microbiol Meth **75**:1-11.
40. **Costa P, Ferreira A, Amaro A, Albuquerque T, Botelho A, Couto I, Cunha MV, Viveiros M, Inácio J**. 2013. Enhanced detection of tuberculous mycobacteria in animal tissues using a semi-nested probe-based real-time PCR. PLoS ONE **8**: e81337.
41. **Mackay IM, Arden KE, Nitsche A**. 2002. Real-time PCR in virology. Nucleic Acids Res **30**:1292–1305.
42. **Timms VJ, Gehringer MM, Mitchell HM, Daskalopoulos G, Neilan BA**. 2011. How accurately can we detect *Mycobacterium avium* subsp. *paratuberculosis* infection? J Microbiol Meth **85**: 1-8.
43. **Leite FL, Stokes KD, Robbe-Austerman S, Stabel J**. 2013. Comparison of fecal DNA extraction kits for the detection of *Mycobacterium avium* subsp. *paratubeculosis* by polymerase chain reaction. J Vet Diagn Invest **25**:27-34.
44. **Plain KM, Marsh IB, Waldron AM, Galea F, Whittington AM, Saunders VF, Begg DJ, de Silva K, Purdie AC, Whittington RJ**. 2014. High-throughput direct fecal PCR assay for detection of *Mycobacterium avium* subsp. *paratuberculosis* in sheep and cattle. J Clin Microbiol **52**:745-757.
45. **Selim A, El-haig M, Galila ES**. 2013. Direct detection of *Mycobacterium avium* subsp. *paratuberculosis* in bovine milk by multiplex real-time PCR. Biotech Anim Husb **29**:513-525.
46. **Gao A, Mutharia L, Raymond M, Odumeru J**. 2007. Improved template DNA preparation procedure for detection of *Mycobacterium avium* subsp. *paratuberculosis* in milk by PCR. J Microbil Meth **69**:417-420.
47. **Slana I, Paolicchi F, Janstova B, Navratilova P, Pavlik I**. 2008. Detection methods for *Mycobacterium avium* subsp. *paratuberculosis* in milk and milk products: a review. Vet Med **53**:283-306.
48. **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.
49. **Stevenson K, Hughes VM, de Juan L, Inglis NF, Wright F, Sharp JM**. 2002. Molecular characterization of pigmented and non-pigmented isolates of *Mycobacterium avium* subspecies *paratuberculosis*. J Clin Microbiol **40**:1798–1804.
50. **de Juan L, Mateos A, Domínguez L, Sharp JM, Stevenson K**. 2005. Genetic diversity of *Mycobacterium avium* subspecies *paratuberculosis* isolates from goats detected by pulsed-field gel electrophoresis. Vet Microbiol **106**:249–257.
51. **Whittington RJ, Marsh IB, Whitlock RH**. 2001b. Typing of IS*1311* polymorphisms confirms that bison (*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 **15**:139–145.

52. **Wynne JW, Bull TJ, Seemann T, Bulach DM, Wagner J, Kirkwood CD, Michalski WP**. 2001. Exploring the zoonotic potential of *Mycobacterium avium* subspecies *paratuberculosis* through comparative genomics. PLoS One **6**:e22171.
53. **Shendure J, Ji H**. 2008. Next-generation DNA sequencing. Nat Biotechnol **26**:1135–1145.
54. **Hsu CY, Wu CW, Talaat AM**. 2011. Genome-wide variation among *Mycobacterium avium* subspecies *paratuberculosis* isolates: a better understanding of Johne's disease transmission dynamics. Front Microbiol **2**:236.
55. **Nielsen R, Paul JS, Albrechtsen A, Song YS**. 2011. Genotype and SNP calling from next-generation sequencing data. Nat Rev Genet **12**:443-451.
56. **Li L, Bannantine JP, Zhang Q, Amonsin A, May BJ, Alt D, Banerji N, Kanjilal S, Kapur V**. 2005. The complete genome sequence of *Mycobacterium avium* subspecies *paratuberculosis*. Proc Natl Acad Sci USA **102**:12344–12349.
57. **Bannantine JP, Wu C-W, Hsu C, Zhou S, Schwartz DC, Bayles DO, Paustian ML, Alt DP, Sreevatsan S, Kapur V, Talaat AM**. 2012. Genome sequencing of ovine isolates of *Mycobacterium avium* subspecies *paratuberculosis* offers insights into host association. BMC Genomics **13**:89.
58. **Ghosh P, Hsu C, Alyamani EJ, Shehata MM, Al-Dubaib MA, Al-Naeem A, Hashad M, Mahmoud OM, Alharbi KBJ, Al-Busadah K, Al-Swailem AM, Talaat AM**. 2012. Genome-wide analysis of the emerging infection with *Mycobacterium avium* subspecies *paratuberculosis* in the Arabian camels (*Camelus dromedarius*). PLoS One **7**:e31947.
59. **Singh SV, Kumar N, Singh SN, Bhattacharya T, Sohal JS, Singh PK, Singh AV, Singh B, Chaubey KK, Gupta S, Sharma N, Kumar S, Raghava GPS**. 2013. Genome sequence of the "Indian Bison Type" biotype of *Mycobacterium avium* subsp. *paratuberculosis* strain S5. Genome Announc **1**:e00005-13.
60. **Ahlstrom C, Barkema H W, Stevenson K, Zadoks R N, Biek R, Kao R, Trewby H, Haupstein D, Kelton D F, Fecteau G, Labrecque O, Keefe G P, McKenna S L B, Buck J De**. 2015. Limitations of variable number of tandem repeat typing identified through whole genome sequencing of *Mycobacterium avium* subsp. *paratuberculosis* on a national and herd level. BMC Genomics **16**:161 1-9.
61. **Bryant JM, Thibault VC, Smith DGE, McLuckie J, Heron I, Sevilla IA, Biet F, Harris SR, Maskell DJ, Bentley SD, Parkhill J, Stevenson K**. 2015. Phylogenomic exploration of the relationships between strains of *Mycobacterium avium* subspecies *paratuberculosis.* BMC Genomics In press.
62. **Biet F, Sevilla IA, Cochard T, Lefrançois LH, Garrido JM, Heron I, Juste RA, McLuckie J, Thibault VC, Supply P, Collins D, Behr MA, Stevenson K**. 2012. Inter and intra-subtype genotypic differences that differentiate *Mycobacterium avium* subsp. *paratuberculosis* strains. BMC Microbiol **12**:264.
63. **Watt JAA**. 1954. Johne's disease in a bovine associated with the pigmented strain of *Mycobacterium johnei*. Vet Rec **66**:387.