Development of a multiplex real-time PCR to distinguish between *Mycoplasma* species found in South African poultry

Pamela P Wambulawaye

Submitted in partial fulfilment in accordance with the requirements for the degree

Magister Scientiae (Veterinary Science)

Department of Production Animal Studies

Faculty of Veterinary Science

University of Pretoria

December 2022

Declaration

I, Pamela Wambulawaye, student number 04611579 hereby declare that this dissertation, “Development of a multiplex real-time PCR to distinguish between *Mycoplasma* species found in South African poultry” is submitted in accordance with the requirements for the Magister Scientiae (MSc) degree at University of Pretoria, is my own original work and has not previously been submitted to any other institution of higher learning. All sources cited or quoted in this research paper are indicated and acknowledged with a comprehensive list of references.

Pamela Wambulawaye

December 2022

Acknowledgements

I would like to extend my sincere gratitude to the following people:

- My supervisor, Prof Celia Abolnik, for her direction, patience, and guidance, for believing in me and wanting to see me grow academically, and without whom this degree would not have been possible.
- My co-supervisor, Dr Amanda Beylefeld, for her knowledge and assistance throughout the study.
- My funders, SARChI (Grant no. N00705/114162) and AgriSETA, for providing the financial support that enabled me to study.
- My colleagues, Drs Ayesha Hassim and Raksha Bhoora, for their assistance with laboratory work.
- My family for all the support and encouragement they have given me.
- My friends, Desirè Mazwi, Faith Nkosi, Thandeka Phiri and Rebone Mahlare, for the helping hand they were always ready and willing to lend.
- Finally, to my amazing fiancé Mark Bessinger, who gives me the strength to keep going every day and loves me unconditionally.

List of Abbreviations

°C	degree Celsius
µl	microlitre
µM	micromolar
AT	adenine-thymine
ATCC	American Type Culture Collection
CFUs	Colony forming units
CO_2	carbon dioxide
cPCR	conventional PCR
C_t	Cycle Threshold
DNA	deoxyribonucleic acid
DsDNA	double-stranded DNA
DVTD	Department of Veterinary Tropical Diseases
EDTA	Ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbent assay
FAA	FASTA file - protein
FASTA	FAST All
FNA	FASTA file - nucleotide
FSA	Contig sequences - nucleotide
Gb	gigabytes
GBK	Genbank file - sequences and annotations
GC	guanine-cytosine
GFF3	general features format (version 3) file
HI	haemagglutination inhibition
IGRs	intergenic regions
kb	kilobases
LOD	Limit of detection
LOQ	Limit of quantification
MG	*Mycoplasma gallisepticum*
MGB	Minor Groove Binder

ml	millilitre
mol%	Mole percent
MS	*Mycoplasma synoviae*
Mspp	*Mycoplasma* species
NCTC	National Collection of Type Cultures
PBS	Phosphate-buffered saline
PCR	polymerase chain reaction
PPLO	pleuropneumonia-like
qPCR	real-time PCR
RAST	Rapid Annotation using Subsystem Technology
RSA	rapid serum agglutination
SA	South Africa
SAPA	South African Poultry Association
SMMEs	Small, Medium and Micro Enterprises
T_a	Annealing temperature
TE	Tris-EDTA
T_m	Melting temperature

TABLE OF CONTENTS

List of Figures

List of Tables

Summary

Avian mycoplasmosis is a serious and chronic bacterial disease caused by *Mycoplasma* species that can greatly impact the sustainability and profits of poultry production. The pathogens significant to poultry are *Mycoplasma gallisepticum* (MG) and *Mycoplasma synoviae* (MS), thus molecular techniques that are readily available focus mainly on these pathogens. Previously, six mycoplasma species were identified from South African poultry flocks, viz. *Mycoplasma gallisepticum*, *Mycoplasma synoviae*, *Mycoplasma gallinarum*, *Mycoplasma pullorum*, *Mycoplasma iners* and *Mycoplasma gallinaceum,* and minimum inhibitory concentration tests performed demonstrated evidence of multidrug resistance in some of the non-pathogenic mycoplasma species. The project is aimed to develop, validate and test a multiplex real-time PCR that could detect and distinguish between four of the *Mycoplasma* species in circulation, for which whole *Mycoplasma* genomes are available. A pan genome analysis identified genes in conserved regions for primer and probe design and synthesis; and a literature review conducted to compare published primer and probe sequences for mycoplasma detection and differentiation. Oligonucleotide primers and probes for the PCR detection and differentiation of *M. gallisepticum*, *M. synoviae*, *M. gallinaceum*, and *M. pullorum* were successfully designed, tested and PCR conditions optimised. A multiplex real-time PCR assay using these oligonucleotides was developed, optimised, and used to test field samples (n=203) collected from farms known to have persisting *Mycoplasma* infections, in conjunction with cultivation and identification. The multiplex real-time PCR assay detected MG in 62 % of the samples tested, MS in 83 %, *M. gallinaceum* in 15 % and *M. pullorum* in 32 %; and coinfections observed in 68 % of the samples. Culture and identification yielded only 9 *Mycoplasma* species: MG, *M. gallinaceum*, *M. pullorum* (n=2), *M. gallinarum*, *M. glycophilum*, and *M. iners* (n=3); all of which are fast growing *Mycoplasma* species, excluding MG. The assay can accurately and simultaneously detect and differentiate between the four *Mycoplasma* species listed. The results obtained give an indication that although there are proportionately more MG and MS species circulating in poultry populations, non-pathogenic *Mycoplasma* species are exceedingly present and appear mostly in coinfections with either MG, MS, or both.

CHAPTER 1: BACKGROUND

1.1 AVIAN MYCOPLASMOSIS

Mycoplasma species causing avian mycoplasmosis were first identified in 1926 in turkeys and then in 1936 in chickens (Nascimento et al., 2005, Umar et al., 2017). With more than 120 named species in the genus *Mycoplasma* (Swayne and Glisson, 2013), more than 20 are capable of infecting avian hosts (Umar et al., 2017). Most significant to poultry are *Mycoplasma gallisepticum* (MG) and *Mycoplasma synoviae* (MS), while *Mycoplasma meleagridis* and *M. iowae* are more significant in turkeys (Nascimento et al., 2005, Umar et al., 2017, Yadav et al., 2021). *Mycoplasma gallisepticum* causes chronic respiratory disease in chickens, and conjunctivitis and sinusitis in turkeys and game birds (Pattison et al., 2007, Swayne and Glisson, 2013), which usually runs a long course with high morbidity, but low mortality in the absence of complicating factors (Pattison et al., 2007). *Mycoplasma synoviae* causes subclinical upper respiratory infections in chickens and turkeys, but is known to become systemic resulting in infectious synovitis which manifests as bursitis, an exudative synovitis, and tenovaginitis (Swayne and Glisson, 2013).

Mycoplasma gallinaceum, *M. gallinarum*, *M. iners* and *M. pullorum*, and other pathogens, including *M. columbinsale*, *M. gallopavonis*, *M. lipofaciens*, *M. columborale*, *M. columbinum*, *M. imitans*, *M. anseris*, *M. glycophilum*, *M. cloacale*, *Acholeplasma laidlawii* and *Ureaplasma galorale* have very low or no pathogenicity and thus are of very little significance in the poultry industry (Nascimento et al., 2005, Swayne and Glisson, 2013). *Mycoplasma imitans* is of interest due to its close relationship to MG, sharing many phenotypic properties such as haemagglutination, haemadsorption, and biochemical reactions, resulting in possible misidentification with MG and serologic cross-reactions in field flock testing (Bradbury et al., 1993, Swayne and Glisson, 2013). In many attempts to detect pathogenic mycoplasmas, *M. gallinaceum* and *M. gallinarum* are usually isolated as contaminants (Swayne and Glisson, 2013).

1.2 LITERATURE REVIEW

1.2.1 Aetiological agent

Mycoplasmas are the smallest self-replicating prokaryotes known, that lack several capabilities normally expressed by other bacteria (Whitford et al., 1994). Mycoplasmas are known to have evolved from a common Gram-positive ancestor with low guanine-cytosine (GC) content (Citti and Blanchard, 2013). They have small genomes as a result of genome reduction, leading to the lack of a variety of metabolic capabilities such as the ability to synthesize a cell wall (Sherwood et al., 2011). The genome size of all mollicutes ranges from less than 600 kilobases (kb) with *M*. *genitalium* at 580 kb, to over 2200 kb with *Spiroplasma ixodetis* at 2220 kb (Razin et al., 1998); while the genome sizes in the genus *Mycoplasma* are between 580-1350 kb (Razin et al., 1998). Characteristically, mycoplasma genomes contain low GC content and apart from *M. pneumoniae* (at 40 mol%) and a few other exceptions, the GC content of all other mollicutes ranges between 24 and 35 mol% (Razin, 1985, Razin et al., 1983). Consequently, mycoplasmal intergenic regions are more adenine-thymine (AT) rich (at 90 mol%) than coding regions (Dybvig and Voelker, 1996).

The cells are built of only three organelles, namely the cell membrane, ribosomes, and a circular double-stranded DNA molecule (Bradbury, 2005, Razin and Hayflick, 2010). Mycoplasmas are pleomorphic and vary in shape from round or pear-shaped cells (0.3-0.8 µm) to flask-shaped cells with terminal tip structures as well as branched or helical filaments (Razin and Hayflick, 2010, Sherwood et al., 2011); thus maintenance of these different shapes suggests the presence of a cytoskeleton (Razin and Hayflick, 2010). As a consequence of having a limited number of genes, mycoplasmas essentially have an intimate association with mammalian cell surfaces which leads to the requirement of complex growth media for *in vitro* cultivation (Sherwood et al., 2011, Whitford et al., 1994). When grown on solid media, they form a “fried egg” colony shape. Replication of cells is not different from other prokaryotes, mycoplasmas reproduce by binary fission where replicated chromosomes are delivered to each cell length before division (Miyata and Seto, 1999, Razin and Hayflick, 2010).

The lack of a cell wall results in detergent and osmotic shock sensitivity, resistance to penicillin and other antibiotics affecting the cell wall (Razin and Hayflick, 2010). Mycoplasmas are able to pass through filters that block the passage of bacteria and

thus were thought to be viruses for years (Razin and Hayflick, 2010). The first cultivation of a mycoplasma, reported in 1898, was the bovine pleuropneumonia agent (Razin and Hayflick, 2010) after which, all similar agents were named pleuropneumonia-like (PPLO) organisms (Nascimento et al., 2005). Following their discovery, mycoplasmas were then confused with bacterial L-form (bacteria that have lost their cell wall partially or entirely) due to their close morphological resemblance (Razin and Hayflick, 2010). They belong to the class *Mollicutes* derived from the Latin '*mollis*' and '*cutis*' meaning 'soft' and 'skin' (Bradbury, 2005), and taxonomy is shown in Table 1.1.

Table 1.1: Taxonomy of all the Mycoplasmas in the class Mollicutes and some characteristics.

Order & Family	Genus	Genome size (kb)	Mol % G+C content	Distinctive features
Mycoplasmatales *Mycoplasmataceae*	*Mycoplasma*	580-1350	23-40	Growth at 37°C
	Ureaplasma	760-1170	27-30	Urea hydrolysis
Entomoplasmatales *Entomoplasmataceae* *Spiroplasmataceae*	*Entomoplasma*	790-1140	27-29	Growth at 30°C
	Mesoplasma	870-1100	27-30	0.04% tween 80 in serum-free media
	Spiroplasma	780-2220	25-30	Helical filaments
Acheloplasmatales *Acheloplasmataceae*	*Acheloplasma*	1500-1650	26-36	Growth at 30-37°C
Anaeroplasmatales *Anaeroplasmataceae*	*Anaeroplasma*	1500-1600	29-34	Strict anaerobes
	Asteroleplasma	1500	40	Oxygen sensitive anaerobes

(Razin et al., 1998, Tully et al., 1993, Whitford et al., 1994)

1.2.2 Epidemiology

1.2.2.1 Hosts - Avian mycoplasmosis

Mycoplasma gallinaceum infections occur primarily in chickens (*Gallus gallus domesticus*) and turkeys (*Meleagridis spp*.) - gallinaceous birds, but have also been isolated from natural infections of geese (*Anser anser*) and ducks (*Anas platyrhynchos*) that came in contact with infected chickens (Pattison et al., 2007, Swayne and Glisson, 2013). Infection can occur at any age but young and stressed birds will more likely develop clinical symptoms (Pattison et al., 2007). *Mycoplasma gallinaceum* has also been isolated from pheasants (*Phasianus colchicus*), guinea fowl (*Numida meleagris*), partridges (*Perdicinae*), peafowl (*Pavo cristatus*), quail (*Coturnix coturnix*), racing pigeons (*Columba livia domestica*), a yellow-naped

Amazon parrot (*Amazona auropalliata*), greater flamingos (*Phoenicopterus roseus*) and, as detected by polymerase chain reaction (PCR), four mature rooks (*Corvus frugilegus*) (Pattison et al., 2007, Swayne and Glisson, 2013).

The role of MG in free-flying birds is still unclear but it has been found in choughs (*Pyrrhocorax pyrrhocorax*) in Scotland, tree sparrows (*Passer montanus*) in Japan and house sparrows (*Passer domesticus*) in India (Pattison et al., 2007). In North America, an epidemic in finches (*Fringillidae*) affected other songbirds as well and while there is little knowledge on the influence of host gender or breed, female finches are more likely to survive MG infection than males (Pattison et al., 2007).

Chickens and turkeys are the main hosts of MS, *M. gallinarum* and *M. lipofaciens* (Pattison et al., 2007). Natural infections of MS have also occurred in geese, pheasants, Japanese quail (*Coturnix japonica*), pigeons, house sparrows in Spain, guinea fowl, red-legged partridges (*Alectoris rufa*) and ducks (Nascimento et al., 2005, Pattison et al., 2007, Swayne and Glisson, 2013). Ducks, geese, budgerigars (*Melopsittacus undulatus*) and pheasants are susceptible to MS when inoculated experimentally (Swayne and Glisson, 2013). Poveda et al*.,* (1990) reported isolating *M. gallinaceum*, *M. gallinarum*, *M. iners* and *M. pullorum* from battery hens with no clinical symptoms, whilst MG and MS showed low incidence (Poveda et al., 1990).

Pigeons are common hosts of *M. columbinsale*, *M. columbinum* and *M. columborale*, and chickens, pheasants and partridges are common hosts of *M. gallinaceum*, *M. glycophilum*, *M. iners* and *M. pullorum* (Benčina et al., 1987, Pattison et al., 2007). *M. gallopavonis* is found in turkeys, *M. anseri* in geese and *M. cloacale* in both turkeys and geese as well as ducks and chickens (Benčina et al., 1987, Pattison et al., 2007). MG, *M. iners*, *M. gallinarum* and *M. gallinaceum* have been isolated from two peregrine falcons (*Falco peregrinus*) that had infraorbital sinusitis and respiratory disorders, and *M. gallinaceum* from the tracheal exudate of a hoopoe (*Upupa epops*) (Poveda et al., 1990).

1.2.2.2 Occurrence - then and now

Mycoplasmas were first cultivated and reported in France in 1898 as the causative agent of bovine pleuropneumonia, now known as *Mycoplasma mycoides.* All similar agents isolated in subsequent years were named the pleuropneumonia-like

organism(s) (Fallon, 1967, Saraya, 2016). The next PPLO, isolated twenty-five years later, was *Mycoplasma agalactiae*, the causative agent of infectious agalactia (mastitis and eye and joint lesions) in sheep and goats; and since then PPLOs were isolated from many animals species including humans (Fallon, 1967, Kumar et al., 2014). It was not until 1961 that the causative agent of primary atypical pneumonia in humans was recognised as a mycoplasma (Marmion and Goodburn, 1961), leading to its isolation in 1962 (Chanock et al., 1962) and designation as *Mycoplasma pneumoniae* in 1964 (Couch et al., 1964). In 1972, a group of helical, motile plant-inhabiting microorganisms were reported in association with corn stunt disease (Davis et al., 1972), termed spiroplasma derived from the Greek ‘speira’ and ‘plasma’ meaning ‘coil’ and ‘shape or form’ (Davis and Worley, 1973); and believed to exist only in nature as insect and plant pathogens (Chen and Liao, 1975, Whitcomb and Williamson, 1975). Then in 1977, it was shown that spiroplasma were not only capable of provoking disease, but were able to persist for prolonged periods in the tissue of host vertebrates (Tully et al., 1977). *Mycoplasma* species causing avian mycoplasmosis were first identified in 1926 in turkeys and then in 1936 in chickens (Nascimento et al., 2005, Umar et al., 2017). Currently, avian mycoplasmosis is exceedingly prevalent and widely distributed in layer, breeder, commercial and even backyard farms worldwide (Felice et al., 2020, Yadav et al., 2021). *Mycoplasma* species are found in Asia, Parts of Africa, Europe, North- and South America, the Middle East, and Oceania (Armour et al., 2013, Behboudi, 2022, Nascimento et al., 2005, Yadav et al., 2021). Chaidez-Ibarra et al., (2022) conducted a systematic review and meta-analysis to estimate the global occurrence of MG and MS, and deduced that the pooled global occurrence was 27.0 % and 38.4 % for MG and MS respectively (Chaidez-Ibarra et al., 2022). In South Africa, *Mycoplasma* species can be found in all nine provinces (Armour et al., 2013, Beylefeld et al., 2018, Moretti, 2012), and MG and MS are the most economically significant.

1.2.3 Clinical signs and lesions

Infections generally occur without clinical signs and the development of disease depends on influencing factors or coinfection with other pathogens (Pattison et al., 2007, Swayne and Glisson, 2013). In MG infections, clinical manifestations are respiratory rales and breathing through a partly open beak, sneezing, conjunctivitis,

coughing, snicks, nasal discharge, swelling of either one or both infraorbital sinuses, and reduced feed consumption and loss of weight (Clark, 2019, Evans et al., 2012, Nascimento et al., 2005, Pattison et al., 2007). *Mycoplasma synoviae* infections have similar clinical signs to MG but in a milder form, apart from pale combs and heads, retarded growth and rapid weight loss, marked depression, swollen joints and foot pads, and lameness. As the disease progresses breast blisters, a shrunken and discoloured comb, and ruffled feathers (Clark, 2019, Nascimento et al., 2005, Pattison et al., 2007, Swayne and Glisson, 2013). Lesions in infected birds include catarrhal inflammation observed on the trachea, bronchi and sinus, thick and opaque air sacs; and histological alterations include a lymphoid follicular reaction that may also affect connective tissue, mucosal glandular hyperplasia and mononuclear cell infiltration (Nascimento et al., 2005).

1.2.4 Transmission and spread

Transmission of mycoplasmas may occur horizontally (direct and indirect contact) or vertically (transovarian) (Clark, 2019, Levisohn and Kleven, 2000). Horizontal transmission is when infections occur through inhalation of aerosols or droplets into the respiratory tract as a result of exhalation, sneezing or coughing up of the organisms (Nascimento et al., 2005, Pattison et al., 2007). This route of spread has been perpetuated by large numbers of flocks kept in close contact and high stock density (Bradbury, 2005). Mycoplasma can survive for a few days outside the host, however, under certain environmental conditions or if well protected by tissue debris or exudates, the pathogen may survive for longer (Jordan, 1985, Polak-Vogelzang, 1977). Fomites such as contaminated feathers, airborne dust, water, feed, contact personnel, suboptimal biosecurity and other avian hosts also play a role in transmission (Nascimento et al., 2005, Swayne and Glisson, 2013).

Vertical transmission is when the pathogens are transferred to the next generation through hatching eggs laid by naturally infected birds (Bradbury, 2005, Swayne and Glisson, 2013). Chicks infected with MG at hatch are a major source of infection in that transmission to eggs at low levels could cause widespread infections between the offspring (Levisohn and Kleven, 2000). *Mycoplasma gallisepticum* survives well in the yolk or allantoic fluid, marking the potential significance of indirect spread in hatcheries via egg debris (Clark, 2019, Pattison et al., 2007). Birds infected with MS become

persistently infected for life and remain carriers (Umar et al., 2017). Artificial insemination of pooled semen from turkey stags is another form of transmission which enables horizontal and probably vertical spread, where a single infected bird could lead to a wide distribution of contaminated semen to female flocks (Bradbury, 2005). Exception for egg transmission, natural infections occur through the inhalation of aerosols into the respiratory tract or through the conjunctiva (Pattison et al., 2007, Swayne and Glisson, 2013).

1.2.5 Pathogenesis

For mycoplasmas to survive in the host, evade the immune system, disseminate and successfully cause disease, some pathogenicity tools/mechanisms are employed (Nascimento et al., 2005). These include an attachment to host cell surfaces, penetrating and surviving in non-phagocytic host cells, apoptosis mediation, gliding motility, mitotic effect for B and/or T lymphocytes, molecular mimicry leading to tolerance, expression of antigenic cell surface variation and production of by-products such as hydrogen peroxide (Bradbury, 2005, Nascimento et al., 2005, Umar et al., 2017). Isolates and strains vary widely in virulence and tropism depending on their phenotypic and genotypic characteristics, propagation method and the number of passages, challenge route and dosage; and this property is readily lost through laboratory passages (Pattison et al., 2007, Swayne and Glisson, 2013).

1.2.6 Influencing factors

Severe airsacculitis (air sac disease) occurs as a result of MG and MS infection that has been complicated by *Escherichia coli* and Newcastle disease or infectious bronchitis (respiratory viral infections) (Swayne and Glisson, 2013). Other pathogens that act synergistically with MG in chickens are the infectious bursal disease virus, influenza A virus, reo- and adenoviruses, *E. coli* (in both chickens and turkeys), *Haemophilus paragallinarum* and infectious laryngotracheitis virus (Pattison et al., 2007). Factors other than pathogens include social stress due to intensive management, increased environmental ammonia from poor ventilation, immunosuppressive agents, high levels of dust and poor nutrition (Pattison et al., 2007).

1.2.7 Economic significance

Mycoplasma gallisepticum is the most economically significant pathogen of poultry (Swayne and Glisson, 2013) that causes the most losses in the poultry industry (Nascimento et al., 2005). An infection causes a decrease in egg production and quality, poor hatchability, reduced feed conversion efficiency, increase in mortality, downgrading of broilers and turkeys at slaughter and carcass condemnations (Clark, 2019, Nascimento et al., 2005, Pattison et al., 2007, Swayne and Glisson, 2013). In layers, infections during lay can cause losses of between 10-20% for periods of up to one month, and a loss of 5-20 eggs per bird in chronic infection without clinical signs (Pattison et al., 2007, Yadav et al., 2021). In breeders, just the suspicion of infection can result in egg and progeny export restrictions, and actual infection in the slaughter of valuable flocks (Pattison et al., 2007, Yadav et al., 2021). Additional costs include treatment, laboratory diagnosis/surveillance, increased biosecurity and vaccination (Pattison et al., 2007, Swayne and Glisson, 2013).

1.2.8 Diagnosis

Mycoplasma infections cannot be diagnosed from clinical symptoms and lesions alone and as such laboratory identification methods are essential. Diagnosis employs three approaches: culture for isolation and identification of the organisms (the gold standard), serology for specific antibody detection and molecular detection of pathogen DNA (Yadav et al., 2021).

1.2.8.1 Culture

Mycoplasmas can be cultured from swabs taken directly from the exudate of nasal sinuses, air sacs, joints, yolk sac endothelium or trachea (Eggleton et al., 1976). However, they are fastidious and slow-growing organisms, and it could take up to three weeks before any significant growth is seen. Mycoplasmas require complex protein-rich growth media that includes animal serum, glucose and/or arginine, yeast extracts, DNA, nicotinamide adenine dinucleotide (required by MS), L-cysteine and other vitamins (Eggleton et al., 1976, Whitford et al., 1994); and some inhibitors such as amoxycillin to inhibit other bacteria and actidione to inhibit fungi (Eggleton et al., 1976).

Non-pathogenic species will usually overgrow slow-growing pathogenic mycoplasmas (Eggleton et al., 1976).

1.2.8.2 Serology

The most common serological techniques are the rapid serum agglutination (RSA) test, enzyme-linked immunosorbent assay (ELISA), and haemagglutination inhibition (HI) test. Serology tests are useful in the diagnosis of suspected mycoplasma infections and monitoring of flocks in control programs; however, they have low sensitivity and are time-consuming and laborious (Swayne and Glisson, 2013, Umar et al., 2017). Problems with reagents cross-reacting as well as non-specific reactions are common (Umar et al., 2017).

1.2.8.3 Molecular

Polymerase chain reaction (PCR) testing is used to detect mycoplasma DNA directly from samples. It is specific, sensitive, and rapid, and is often used as a confirmation tool to detect specific mycoplasma DNA instead of culture (Umar et al., 2017, Yadav et al., 2021). Multiplex PCR assays are also available for the simultaneous detection of different mycoplasma species (Swayne and Glisson, 2013). A summary of the molecular methods available for mycoplasmas is shown in the table in Appendix A1.

1.2.8.3.1 PCR

Polymerase chain reaction (also referred to as conventional PCR) is a cyclic process that amplifies nucleic acid sequences (DNA) in a given sample to produce a large number of copies of the target sequence for later analysis (Kubista et al., 2006). In just a few hours, a DNA segment or gene could be amplified a million times (Figure 1.1) (Pestana, 2010). The first step of the process is denaturation, where the double-stranded DNA (dsDNA) template to be amplified is denatured/separated by heating (94°C), to break the hydrogen bonds connecting the two DNA strands; and this is also true for the primers (Pestana, 2010, Viljoen et al., 2005). This is followed by annealing, where the temperature is lowered to 45 - 60°C to allow the primers, which are synthetic pieces of DNA that complement specific sequences of the target DNA segment, to select and bind/hybridize themselves to the single DNA strands (Pestana, 2010,

Viljoen et al., 2005). Lastly elongation, where the temperature is raised to 72°C and a heat stable DNA polymerase elongates the attached primers along the single DNA strand (Pestana, 2010, Viljoen et al., 2005).

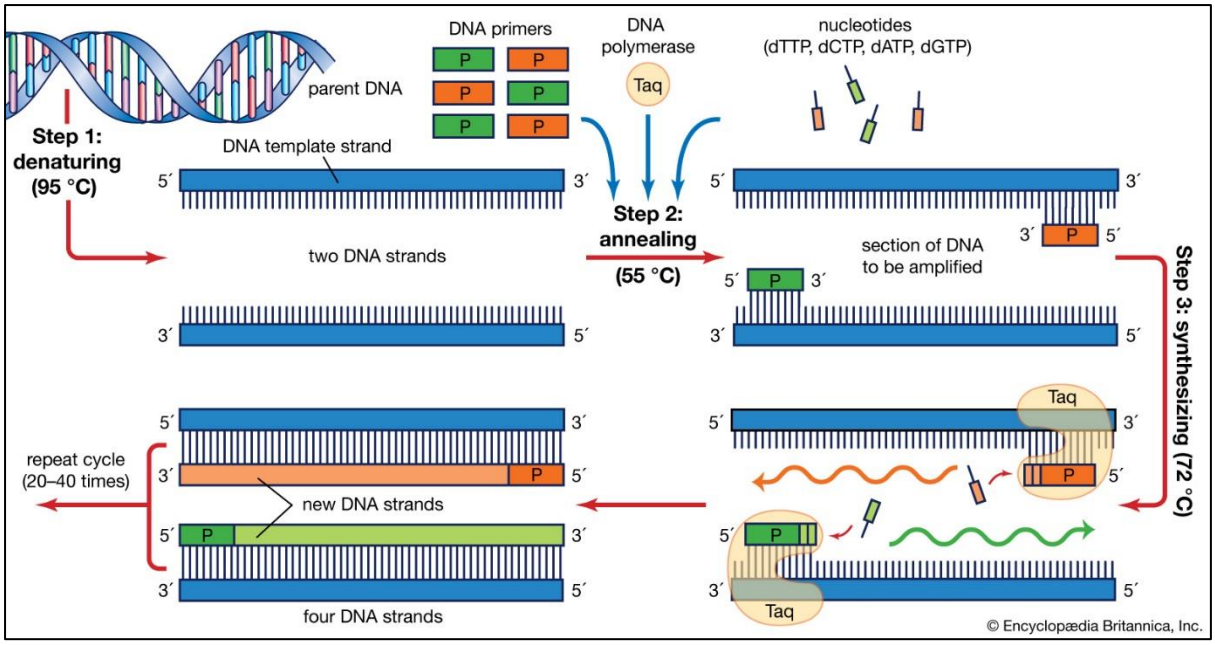

Figure 1.1: The polymerase chain reaction process.

Source: https://www.britannica.com/science/polymerase-chain-reaction#/media/1/468736/18071

Conventional PCR involves endpoint analysis were the PCR product (amplified DNA) is stained with a marker, that can seep between the bases of the DNA strand, and allowed to migrate through agarose gels using electrophoresis (separation by band size and weight) and visualised using ultraviolet illumination (Lorenz, 2012). The bands are then compared to a standard molecular weight maker, also referred to as a GeneRuler™ / DNA ladder, of a particular size.

1.2.8.3.2 Real-time PCR

Real-time PCR extends the capabilities of conventional PCR by simultaneously amplifying and monitoring the products generated in each cycle of the process in real-time, through different fluorometric chemistries (Pabla and Pabla, 2008, Pestana, 2010, Wilhelm and Pingoud, 2003). In real-time PCR the same elements are required as in conventional PCR, with the addition of one or more oligonucleotides with a fluorescent reporter that will signal and reflect the amount of formed product (Kubista

et al., 2006). Initially the signal is weak and cannot be separated from the background, but as the cycles continue running and more products are amplified, the signal starts increasing exponentially after which it levels off and saturates (Kubista et al., 2006, Pestana, 2010). Saturation is caused by the depletion of required elements in the reaction, these can be the primers, the dNTPs, or the reporter; or due to a limiting number of polymerase molecules (Kubista et al., 2006).

1.2.9 Control

The most ideal control strategy is starting with breeding flocks that are free of mycoplasma infections and rearing them in a clean environment with proper isolation and biosecurity measures to eliminate introduced infections (Kleven, 2008, Swayne and Glisson, 2013). Since all pathogenic mycoplasmas are vertically transmitted, breaking the cycle of egg-transmission has made it possible to eliminate infections from breeder flocks (Kleven, 2008, Whithear, 1996), and the specific pathogen-free status that falls on the progeny brings with it some economic benefits (Whithear, 1996). Maintenance of mycoplasma-free flocks is not always possible, particularly on sites with multi-age flocks or continuous production sites, therefore treatment with antimicrobials and vaccination are alternative measures that can also be used (Pattison et al., 2007, Whithear, 1996).

Although mycoplasmas are resistant to penicillins and cephalosporins (antibiotics affecting the cell wall), they are susceptible to quinolones (enrofloxacin, danofloxacin, imequil and norfloxacin), tetracyclines (chlortetracycline, doxycycline and oxytetracycline), macrolides (tylosin, kitasamycin, lincomycin, erythromycim and spiramycin), aminoglycosides and pleuromutilin (tiamulin) (Kleven, 2008, Nascimento et al., 2005). Using antimicrobials that affect other bacteria would be most beneficial, but it should be noted that not all mycoplasmas will be eliminated, and overuse could result in the development of resistant strains (Pattison et al., 2007).

Current commercial vaccine options include live vaccines, killed vaccines, oil-adjuvant vaccines (bacterins) and recombinant live poxvirus vaccines that contain and express MG antigens (Kleven, 2008). Three live vaccines are commercially available for MG; the F strain, 6/85 strain, and ts-11 strain, and only one live vaccine for MS which is the MS-H strain. The F strain vaccine is highly virulent to turkeys, of moderate virulence

to chickens and has been used in immunization programs for pullets (Swayne and Glisson, 2013, Whithear, 1996). The 6/85 vaccine is of low virulence, available in a lyophilized form and vaccination is administered by aerosol (Kleven, 2008, Pattison et al., 2007). The ts-11 strain, originally obtained from an Australian field isolate that was mutagenized chemically and selected for its temperature-sensitivity, is a vaccine that has low virulence, is safe in chickens and turkeys and has poor transmissibility from bird to bird (Kleven, 2008). It is administered via eye drop, induces immunity that is long-lived in the birds by persisting for long periods in the upper respiratory tract and allows for use with other respiratory viral vaccines (Kleven, 2008, Pattison et al., 2007). The MS-H strain, also originating from Australia, is a temperature-sensitive vaccine proven to be effective in the reduction of apical egg shell abnormalities caused by MS (Swayne and Glisson, 2013, Whithear, 1996). Oil-adjuvant vaccines (killed mycoplasmas that are administered by injection) will not prevent infection but will provide protection against losses in egg production (caused by MG) in layers (Pattison et al., 2007).

1.3 JUSTIFICATION

Globally, consumers are more attracted to poultry meat due to its affordability compared to other types of meat. Positive trends seen in the increase in consumption have held true for the past 35 years (Valceschini, 2006, Windhorst, 2006) and consumption has been shown to steadily increase over time (Whitnall and Pitts, 2019). As the second most consumed meat in the world after pork, poultry meat and eggs are of significant economic importance worldwide (Chai et al., 2017, Roenigk, 1999). In Africa, it is estimated that 20% of consumed protein originates from poultry, and more than 80% of rural households keep poultry in their backyards (Mtileni et al., 2012). In South Africa (SA), poultry production accounted for a total gross value (consumption) of 20.9% and an animal product gross value (consumption and usage) of 43% in 2018 (SAPA, 2018). From 2020 to 2021 the poultry industry for meat production accounted for 74.5 % gross value, while egg production accounted to 25.5 % gross production (SAPA, 2021). In addition to GDP, the poultry industry is also an important source of employment opportunities and a large part consists of Small, Medium and Micro Enterprises (SMMEs) (Mkhabela and Nyhodo, 2011). Under the Fertilizers, Farm Feeds, Seeds and Remedies Act 36 of 1947, poultry litter can be used as animal feed for livestock, provided that regulations are correctly adhered to (Van Ryssen, 2001, Wiese and Bot, 1971).

Unfortunately, keeping poultry comes with many drawbacks and the most significant are diseases. Many diseases that affect poultry and can be grouped as viral-, bacterial-, fungal-, parasitic-, and non-infectious diseases (Swayne and Glisson, 2013). Avian mycoplasmosis is a serious and chronic bacterial disease that can greatly impact the sustainability and profits of poultry farming in any economy, and because the disease does not manifest clinical symptoms that can be used for diagnosis, testing is required (Wang et al., 1990). The gold standard of identification is culture, but the time required (up to three weeks) to get significant growth is a major limitation, and thus molecular techniques are used.

The molecular techniques available are mainly focused on mycoplasmas considered to be of economic importance to poultry health and production (Sprygin et al., 2010) and as such non-pathogenic mycoplasmas are ignored. Previous diagnostic testing of samples from poultry flocks in SA mainly involved culture followed by growth inhibition

using species-specific hyperimmune anti-serum, which would only identify mycoplasmas as either MS, MG, or *M. species* (unidentified) (Beylefeld, 2018). Beylefeld, (2018) then classified these samples into six mycoplasma species, viz. *M. gallisepticum*, *M. synoviae*, *M. pullorum*, *M. gallinarum*, *M. iners* and *M. gallinaceum* and one *Acheloplasma laidlawii* strain, through full-genome sequencing and phylogenetic analysis. *Mycoplasma meleagridis* and *M. iowae* were not found, this is largely because they are mainly pathogens of turkeys (Nascimento et al., 2005, Umar et al., 2017) and there is no commercial turkey farming in South Africa (Beylefeld, 2018). The putative transposase gene found in the 16S-23S rRNA intergenic spacer region of *M. imitans* (Harasawa et al., 2004), used to differentiate it from MG, was not found in the samples (Beylefeld, 2018).

Minimum inhibitory concentration tests were performed on the six mycoplasma species, and evidence of multidrug resistance in some of the non-pathogenic mycoplasma species was demonstrated (Beylefeld et al., 2018). It was found that *M. gallinarum*, *M. pullorum* and *M. gallinaceum* strains were more likely to be resistant to a range of antimicrobials than *M. gallisepticum* and *M. synoviae* (Beylefeld et al., 2018). If these species were to transfer antibiotic-resistant traits to other mycoplasma species (although this is currently unknown), it would have major implications for the control and management of avian mycoplasmosis in the country, especially since there are no vaccines available for non-pathogenic mycoplasmas (Beylefeld et al., 2018). Thus, the development of a molecular test to detect all the mycoplasmas circulating in the national flock and not only the significant pathogens could bring a better understanding of the persistence of mycoplasma infections.

1.4 RESEARCH AIM AND OBJECTIVES

This study aimed to develop and test a multiplex real-time PCR that would simultaneously detect six *Mycoplasma* species that were previously identified in the national flock, namely *M. gallisepticum*, *M. synoviae*, *M. pullorum*, *M. gallinarum*, *M. iners* and *M. gallinaceum*, and a mycoplasma group target as an internal control to indicate species that may not have been detected previously. Since the complete genome sequences for *M. gallinarum* and *M. iners* are not yet available, these two species could not be included in the analysis and primer design. The objectives of the study were as follows:

1. Design and optimisation of PCRs for mycoplasma detection and differentiation
    - Conduct a literature review to compare published primer and probe sequences for mycoplasmas
    - Conduct a pan-genome analysis to identify genes in conserved regions using available whole *Mycoplasma* genomes
    - Select or design group- and species-specific primers for synthesis
    - Test synthesised primers on cPCR for specificity and sensitivity
    - Synthesise respective probes

2. Design, diagnostic efficiency, and testing of a Multiplex Real-time PCR assay
    - Test primers and probes in single locus real-time PCRs (Assay linear range, efficiency and LOD)
    - Test known positive axenic and mixed samples in multiplex real-time PCRs
    - Conduct a field survey with sample collection from chicken flocks suspected of mycoplasma infection
    - Test field samples on the multiplex real-time PCR assay in conjunction with cultivation and identification (gold standard).

CHAPTER 2: DESIGN AND OPTIMISATION OF PCRS FOR MYCOPLASMA DETECTION AND DIFFERENTIATION

2.1 INTRODUCTION

Mycoplasmas have exceedingly small genomes when compared to other bacterial microorganisms. The genome sizes in the genus *Mycoplasma* are between 580-1350 kb (Razin et al., 1998), making them much easier to work with using bioinformatics programs and tools. Bioinformatics is a new and continuously growing section of science in which biological questions are tackled using computational approaches/tools and the associated information used takes advantage of large and complex sets of data (Baxevanis et al., 2020, Luscombe et al., 2001). In general, these computational tools are used, for example, in protein structure and physiochemical property determination, phylogenetic analyses, gene characterisation, etc. (Amer Mohamed et al., 2019) and are useful in guiding researchers in selecting better and informed approaches when designing experiments in the laboratory (Amer Mohamed et al., 2019, Baxevanis et al., 2020). In summary, bioinformatics aids in data organisation where researchers can not only get easy access to existing information but can also provide new information as they received it and have resources and tools developed to assist in analysis of the data and interpretation of acquired results in a biological fashion (Luscombe et al., 2001).

The tools of interest for this study were the rapid prokaryote genome annotation tool Prokka (available at https://github.com/tseemann/prokka), the pan-genome pipeline tool Roary (available at https://github.com/sanger-pathogens/Roary), and the intergenic regions (IGRs) analysis tool Piggy (available at https://github.com/harry-thorpe/piggy). These tools can all be installed and accessed through an operating system called Ubuntu, a distribution of Linux based on Debian (Raggi et al., 2011) (available at https://ubuntu.com/download).

Prokka is a software tool that involves whole genome annotation, where in a set of genomic DNA sequences, relevant features are identified and labelled with useful information (Seemann, 2014). It is a rapid tool that will produce annotation results in at least 10 minutes of inputting a command (Seemann, 2014), as compared to, for example, the Rapid Annotation using Subsystem Technology (RAST) server

(available at https://rast.nmpdr.org/) where a similar function would take up 24 hours to produce results (Aziz et al., 2008). The input data/files used in Prokka should be preassembled DNA sequences presented in FASTA format (Seemann, 2014) and the output files produced are listed in Table 2.1.

Table 2.1: Output files produced by Prokka.

Extension	Description
.gff	Master annotation in GFF3 format, containing both sequences and annotations.
.gbk	Standard Genbank file derived from the master .gff.
.fna	Nucleotide FASTA file of input contig sequences.
.faa	Protein FASTA file of translated CDS sequences.
.ffn	Nucleotide FASTA file of all prediction transcripts (CDS, rRNA, tRNA, tmRNA, misc_RNA)
.sqn	ASN1 format "Sequin" file for submission to Genbank.
.fsa	Nucleotide FASTA file of the input contig sequences, used by "tbl2asn" to create the .sqn file.
.tbl	Feature Table file, used by "tbl2asn" to create the .sqn file.
.err	Unacceptable annotations - the NCBI discrepancy report.
.log	Contains all the output that Prokka produced during its run.
.txt	Statistics relating to the annotated features found.
.tsv	Tab-separated file of all features: locus_tag,ftype,len_bp,gene,EC_number,COG,product

Source: https://github.com/tseemann/prokka#output-files

Roary is a computational tool that rapidly constructs large-scale pan genomes to identify core and accessory genes from samples of the same species (Page et al., 2015) through analysis of tens or hundreds of lineages, producing results in a decent amount of computational time (Sitto and Battistuzzi, 2020). A pan-genome is a collection of all the genes that are shared by various strains of a particular species (Sitto and Battistuzzi, 2020, Tettelin et al., 2005). Using Roary, a pan-genome analysis can be performed to obtain a list of core genes with 95% identity in 99% of all the strains of a particular species. To input data/files into Roary, an annotated assembly of each sample in GFF3 format, obtained from the National Center for Biotechnology Information (NCBI) genomes database, should be used (Page et al., 2015, Sitto and Battistuzzi, 2020). Files (FSA files/contigs files) obtained from the RAST server would first need to be annotated using Prokka (Page et al., 2015). There is a list of options when setting the commands of interest in Roary (see Table 2.2). The output results produced are a series of statistical files on genes that are shared by most or all

lineages (core or soft core), or by some genomes (accessory) (Sitto and Battistuzzi, 2020).

Like Roary, Piggy is also a rapid, large-scale pan-genome analysis tool, used for the IGRs of bacterial genomes rather than genes. Piggy is a tool that compares entire sequences, and the outcome presents as lists of clusters for each species. It is used in conjunction with Roary; where the output data produced by Roary is used as the input data.

Table 2.2: List of different command settings used in any specific Roary run.

*Options		Description
-p	INT	Number of threads [1]
-o	STR	Clusters output filename [clustered_proteins]
-f	STR	Output directory [.]
-e		Create a multiFASTA alignment of core genes using PRANK
-n		Fast core gene alignment with MAFFT, use with -e
-i		Minimum percentage identity for blastp [95]
-cd	FLOAT	Percentage of isolates a gene must be in to be core [99]
-qc		Generate QC report with Kraken
-k	STR	Path to Kraken database for QC, use with -qc
-a		Check dependancies and print versions
-b	STR	Blastp executable [blastp]
-c	STR	Mcl executable [mcl]
-d	STR	Mcxdeblast executable [mcxdeblast]
-g	INT	Maximum number of clusters [50000]
-m	STR	Makeblastdb executable [makeblastdb]
-r		Create R plots, requires R and ggplot2
-s		Don’t split paralogs
-t	INT	Translation table [11]
-ap		Allow paralogs in core alignment
-z		Dont delete intermediate files
-v		Verbose output to STDOUT
-w		Print version and exit
-y		Add gene inference information to a spreadsheet, does not work with -e
-iv	STR	Change the MCL inflation value [1.5]
-h		Help message

*Vary according to which commands are of interest for each specific run and are shown with the command: `roary -h`

Source: https://github.com/sanger-pathogens/Roary#usage

In this chapter, Roary was used to obtain a list of core genes with 95 % identity, found in 99 % of all the strains/samples of each species. Four lists of genes were obtained, and the genes used for primer design, synthesis, and testing, were identified, and selected from each list. Piggy was used to obtain data as an alternative if the data obtained from Roary could not be used. A literature review was also conducted to compare published primer and probe sequences for mycoplasma detection and differentiation.

2.2 MATERIALS AND METHODS

2.2.1 Poultry mycoplasma isolates in the UP-DVTD repository

The isolates used in this chapter originated from farms in Gauteng, Limpopo, North-West and the Western Cape provinces. The isolates were cultured, genomic DNA extracted, and whole genome sequencing was performed with the data placed in a mycoplasma database for future studies (results of a previous study) (Beylefeld, 2018). The cultures were kept in mycoplasma broth at -80°C in the Department of Veterinary Tropical Diseases at the University of Pretoria, and were used with permission from Dr Annelize Jonker, the subject specialist and veterinarian of the Bacteriology section. Table 2.3 shows the list of samples used.

Table 2.3: Mycoplasma samples used for testing of primers and diagnostic efficiency.

*Axenic mycoplasma samples		Mixed mycoplasma samples	
B1102-03_MG	B313-05_Mgal	Bedson 04/09.07	B359/15 (2)
B726-06_MG	B733-05_Mgal	B1072/08	B359/15 (3)
B852-06_MG	B2176-13_Mgal	B730/09	B359/15 (4)
B943-06_MG	B878-14-M1_Mgal	B2076/13 (3)	B359/15 (8)
B1102-06_MG	B878-14-M4_Mgal	B2888/13 1A	B457/15 (3)
B1028-07_MG	B878-14-M5_Mgal	B1101/14 (10)	B458/15 (10)
B642-08_MG	B1101-14-7_Mgal	B1342/14 (18)	B458/15 (5)
B758-08_MG	B1173-14-2A_Mgal	B1342/14 (4)	B464/15 (3)
B2159-13_MG	B1173-14-2B_Mgal	B1342/14 (9)	B540/15 (4)
B1395-14-1_MG	B1173-14-4A_Mgal	B1393/14 (4)	B540/15 (5)
B1552-14-19_MG	B1173-14-4B_Mgal	B1395/14 (5)	
B2771-14-1A_MG	B1173-14-5_Mgal	B1396/14 (6)	
B2771-14-1B_MG	B1173-14-6_Mgal	B2771/14 (15) A	
B878-14-L3_MG	B1173-14-7_Mgal	B878/14 -Moria 2	
B457-15-5_MG	B1173-14-8_Mgal	B1931/15 (6A)	
B2214-07_MS	B1342-14-8_Mgal	B1932/15 (2)	
B1064-14-H3_MS	B1342-14-10_Mgal	B2053/15 (1)	
B1064-14-H4_MS	B1342-14-13_Mgal	B2053/15 (3)	
B1064-14-H5_MS	B1342-14-14_Mgal	B2053/15 (5)	
B1393-14-10_MS	B1395-14-2_Mgal	B2063/15 (3)	
B1394-14-2_MS	B1396-14-7_Mgal	B2777/15A (7)	
B1394-14-5_MS	B1396-14-8_Mgal	B2777/15A (8)	
B458-15-1_MS	B1396-14-9_Mgal	B293/15 (14)	
B458-15-5M_MS	B1414-14-1_Mgal	B293/15 (18)	
B458-15-6_MS	B2096-14-2_Mgal	B293/15 (7)	
B458-15-11_MS	B2096-14-4_Mgal	B293/15 (9)	
B2096-14-3_Mpul	B2096-14-7_Mgal	B3443/15 (1)	
B293-15-12_Mpul	B2096-14-8_Mgal	B3443/15 (2)	
B293-15-13_Mpul	B293-15-16_Mgal	B3443/15 (3)	
B293-15-15_Mpul	B3381-15-1_Mgal	B3443/15 (4)	
B293-15-17_Mpul	B3381-15-2_Mgal	B3443/15 (5)	
B359-15-5_Mpul	B3381-15-3_Mgal	B3443/15 (6)	
B359-15-6_Mpul	B3381-15-4_Mgal	B3443/15 (7)	
B540-15-2_Mpul	B3381-15-5_Mgal	B3443/15 (8)	

*MG - *M. gallisepticum*
*MS = *M. synoviae*
*Mpul = *M. pullorum*
*Mgal = *M. gallinaceum*

2.2.2 Pan-genome analysis

The complete unpublished genomes of MG (n=15), MS (n=11), *M. gallinaceum* (n=34) and *M. pullorum* (n=8) species circulating in the South African poultry flock (Beylefeld, 2018) were used along with complete published genomes of MG (n=8), MS (n=5), and *M. gallinaceum* (n=2) strains, available in the NCBI genomes database (Table 2.4). Genomes for MG strain R(low) (accession number AE015450), MS strain 53 (accession number AE017245), *M. gallinaceum* strain B2096-14-8 (accession number CP011021) and *M. pullorum* strain B359-15-6 (accession number CP01781) were used as references. Data of the unpublished genomes for the four species listed were downloaded from the RAST server, and the downloaded files were .gff and .fsa (Table 2.1). The .fsa files were then annotated using Prokka and output data were stored for further use. Using the output .gff files obtained from Prokka, the .gff files downloaded from the RAST website were adjusted with 1) ## *sequence-region Joined_contig* at the beginning and 2) ## *FASTA* + *nucleotide sequence* at the end. This was done to get the RAST files in the correct readable format for Roary.

Using these two sets of data (Prokka and RAST), the pan-genome analysis was done. As Roary takes the first sample in the list of samples used upon input and compares it to the rest of the samples in the list, the samples and references were used interchangeably. The command settings used for the Roary runs were:

1. `~$ roary -f [output directory] -e -n -t 4 -v [folder location/*.gff]`,
2. `~$ roary -f [output directory] -e -n -r -s -t 4 -z -v [folder location/*.gff]`, and
3. `~$ roary -f [output directory] -e -n -i 90 -r -s -t 4 -z -v [folder location/*.gff]`.

In command setting 1, the options selected were `-e` to create a multiFASTA alignment, `-n` for a fast core gene alignment, and `-t 4` for the mycoplasma translation table. In command setting 2, `-s` and `-z` were added to prevent Roary from splitting paralogs and deleting intermediate files. Lastly in setting 3, `-i 90` was used to keep the minimum percentage identity for blastp at 90%, and the output data in this command setting would be used as input data for Piggy.

Table 2.4: Complete Mycoplasma genome sequences available in Genbank.

#Organism Name	Strain	Assembly	Level	Size(Mb)	GC%	Release Date	GenBank FTP
M. gallisepticum	R(high)	GCA_000025365.1	Complete	1,01203	31,5	2010-01-29T00:00:00Z	ftp://ftp.ncbi.nlm.nih.gov/gen
M. gallisepticum	F	GCA_000025385.1	Complete	0,977612	31,4	2010-01-29T00:00:00Z	ftp://ftp.ncbi.nlm.nih.gov/gen
M. gallisepticum	R(low)	GCA_000092585.1	Complete	1,0128	31,5	2003-06-09T00:00:00Z	ftp://ftp.ncbi.nlm.nih.gov/gen
M. gallisepticum	S6	GCA_000211545.6	Complete	0,985433	31,5	2013-12-11T00:00:00Z	ftp://ftp.ncbi.nlm.nih.gov/gen
M. gallisepticum	f99 Avipro vaccine	GCA_004771095.1	Complete	0,975069	31,4	2019-04-11T00:00:00Z	ftp://ftp.ncbi.nlm.nih.gov/gen
M. gallisepticum	6/85	GCA_008728895.1	Complete	0,994372	31,6	2019-09-30T00:00:00Z	ftp://ftp.ncbi.nlm.nih.gov/gen
M. gallisepticum	ts-11	GCA_008728915.1	Complete	0,963058	31,4	2019-09-30T00:00:00Z	ftp://ftp.ncbi.nlm.nih.gov/gen
M. gallisepticum	NCTC10115	GCA_900476085.1	Complete	0,981408	31,5	2018-08-19T00:00:00Z	ftp://ftp.ncbi.nlm.nih.gov/gen
M. synoviae ATCC 25204	WVU1853	GCA_000969765.1	Complete	0,846495	28,3	2015-04-07T00:00:00Z	ftp://ftp ftp://ftp.ncbi.nlm.nih.
M. synoviae	NCTC10124	GCA_900475235.1	Complete	0,848181	28,4	2018-08-19T00:00:00Z	ftp://ftp ftp://ftp.ncbi.nlm.nih.
M. synoviae	MS-H	GCA_003147565.1	Complete	0,818848	28,2	2018-05-22T00:00:00Z	ftp://ftp ftp://ftp.ncbi.nlm.nih.
M. synoviae	HN01	GCA_009671165.1	Complete	0,817087	28,3	2019-11-19T00:00:00Z	ftp://ftp ftp://ftp.ncbi.nlm.nih.
M. synoviae	53	GCA_000008245.1	Complete	0,799476	28,5	2005-08-05T00:00:00Z	ftp://ftp ftp://ftp.ncbi.nlm.nih.
M. gallinaceum	NCTC10183	GCA_900660495.1	Complete	1,07484	28,7	2019-01-28T00:00:00Z	ftp://ftp ftp://ftp.ncbi.nlm.nih.
M. gallinaceum	B2096 8B	GCA_000965765.1	Complete	0,845307	28,4	2015-03-30T00:00:00Z	ftp://ftp.ncbi.nlm.nih.gov/gen
M. gallinarum DSM 19816	Bacteria;Terraba	PRJNA221017	GCA_000621	Scaffold	0,8347	646	2014-04 ftp://ftp.ncbi.nlm.nih.
M. gallinarum	Bacteria;Terraba	PRJNA247551	GCA_001637	Contig	0,7977	619	2016-05 ftp://ftp.ncbi.nlm.nih.
M. iners ATCC 19705	https://www.ncbi.nlm.nih.gov/nuccore/657716126						
Field strain							
Vaccine strain							

2.2.3 Gene selection and primer/probe design

The sequences of the genes obtained from the pan-genome were aligned using CLC Genomics Workbench version 8.5.1 (CLC Bio-Qiagen, Aarhus, Denmark). For the selection of good candidate genes, alignment was done initially for each of the four-mycoplasma species of interest to identify similarities, differences, deletions, and insertions, and then between species for overlapping similarities. The full sequences of the selected candidate genes were then retrieved from the references, in the NCBI genomes database.

The oligonucleotides of the primers and probes were designed using the Realtime PCR tool from Integrated DNA Technologies™ (IDT). Parameters such as annealing temperature, self-annealing and hairpin formation were tested using the online tool Oligo Calc (available at http://biotools.nubic.northwestern.edu/OligoCalc.html), and nucleotide Basic Local Alignment Search Tool (BLASTn) (available at https://blast.ncbi.nlm.nih.gov/Blast.cgi) screening was performed to test for non-specific binding *in silico*. The probes were labelled at the 5'-end with fluorescent reporter dyes: 6-FAM™ for MG & *M. gallinaceum*, VIC® for MS & *M. pullorum*, and NED® for the group-specific and all probes were labelled at the 3'-end with minor groove binder nonfluorescent quenchers (MGB-NFQ). One published primer set for *M. gallinaceum* and two published primer sets for MG and MS were selected from the table in Appendix A2, and synthesized along with the newly designed primers, to assess the possibility of using a published assay and/or support the need for a novel assay.

The group-specific primers were selected and used as described by McAuliffe et al., (2005) and Kiss et al., (1997a). The forward primer used by McAuliffe et al., (2005) was selected from the amplification of the V3 region of the 16S RNA gene with minor modifications, using the universal bacterial primer GC-341F; and the reverse primer (R543), specifically for *Mollicutes* (*Mycoplasma* species), was designed using Primrose (Ashelford, 2002) through alignment of 102 *Mycoplasma* species (McAuliffe et al., 2005). Kiss et al., (1997a) designed species-specific primers that allow for the general detection of mycoplasma species.

2.2.4 Primer testing

2.2.4.1 Samples

Two representatives of each species were selected from the axenic culture repository to test the primers. The primers were first tested using the respective target species, then tested against non-target species for specificity and finally tested using the group-specific primers. A 16S primer pair from (Beylefeld, 2018, Van Kuppeveld et al., 1992) that can detect mycoplasma species was used as a positive control. The samples used to test the primers are listed in Table 2.5 below.

Table 2.5: Sample list for conventional PCR test.

Species	Sample ID
MG	B457-15-5, B1395-14-1 & B1552-14-19
MS	B458-15-5A & B1393-14-10
M. gallinaceum (Mgal)	B1173-14-2A & B1173-14-2B
M. pullorum (Mpul)	B540-15-2 & B293-15-15

*PCR-grade water was used as a negative control.

2.2.4.2 Conventional PCR

The primers (Table 2.8) were reconstituted to 100 µM stock concentrations and then further diluted to 10 µM working solutions, using PCR-grade water. For the conventional PCR, the Thermo Scientific™ Phusion™ Flash High-Fidelity PCR Master Mix (ThermoFisher Scientific) was used and prepared according to the manufacturer's recommendations. Each 20 µl PCR reaction contained 10 µl of 2X Phusion Flash PCR Master Mix, 1 µl each of forward and reverse primer, 7 µl of PCR-grade water, and 1 µl of nucleic acid extracts. Amplification was done in the Veriti™ 96-Well Thermal Cycler, using the conditions in Table 2.6.

Table 2.6: Conditions used for conventional PCR.

PCR step	Temperature	Duration	Number of cycles
Initial denaturation	98 °C	10 sec	1
Denaturation	98 °C	1 sec	30
Primer annealing	Various	5 sec	
Extension	72 °C	15 sec	
Final extension	72 °C	1 min	1
	4 °C	∞	

To visualise the results, 10 µl of PCR product mixed with 2 µl of loading dye (ThermoFisher Scientific) was loaded on a 1 % agarose gel prepared in 1X tris acetate ethylenediaminetetraacetic acid (TAE) buffer with 0.5 µg/ml of ethidium bromide. The band sizes were compared to a GeneRuler^{TM} 100 bp DNA ladder (ThermoFisher Scientific), on the gels (pages 29-36) with the lowest band representing 100 bp. To evaluate the specificity of the primers, DNA fragments excised from the agarose gels were extracted and purified in TAE buffer using the QIAquick Gel Extraction Kit (ThermoFisher Scientific) and sent to Inqaba Biotechnical Industries (Pty) Ltd for Sanger sequencing.

2.3 RESULTS

2.3.1 Pan-genome analysis

The summary_statistics.txt files of the output data (total number of genes) obtained from command settings 1 and 2 for both Prokka and RAST were compared (Table 2.7). The RAST data yielded a higher number of genes overall and was therefore selected for use in the rest of the study. For data analysis, the gene_presence_absence.scv files from command settings 1 and 2 were combined in each of the four species. A total of 6 059 genes for MG, 2 266 genes for MS, 8 235 genes for *M. gallinaceum*, and 2 314 genes for *M. pullorum* were obtained (Data submitted electronically, available upon request). Although many genes were listed by their proper gene names, some were listed as either ‘group_number’, ‘FIG_number-hypothetical’, or just as ‘hypothetical’ these were deleted together with several repetitions. Genes not present in all the samples for each species were deleted, and genes that were too short (less than 500 bp) or too long (over 3000 bp) were also deleted (Pestana et al., 2010). This resulted in 189 genes remaining for MG, 256 genes for MS, 81 genes for *M. gallinaceum*, and 200 genes for *M. pullorum*.

2.3.2 Gene selection and primer/probe design

Following alignment, the number of differences in the sequences of the samples in each species was recorded as either no differences (identical), less than 5 differences, up to 10 differences and over 10 differences (Appendix B1-B4). Genes with over 10 differences were excluded thus five genes were selected for primer design/synthesis for each species. *Mycoplasma gallinaceum* was the only exception as it only had one gene with less than five differences whilst the rest had over 10 differences. The primers and probes were designed using the IDT Realtime PCR tool (available at https://eu.idtdna.com/scitools/Applications/RealTimePCR/default.aspx) according to the following criteria: length 18-30 bp, melting temperature 60-65 °C (8 °C higher for probes), GC content 40-60%, amplicon size 80-120 bp. BLASTn results of the primers and probes showed high specificity of each species to target species on the NCBI database and no non-specific detection. Almost none of the primers had self-annealing and/or hairpin formation properties. The primers were synthesised by LTC Tech South Africa (PTY) LTD (Applied Biosystems®). All primers are listed in Table 2.8.

Table 2.7: The summary statistics data of the different sample lists for command settings 1 and 2 used for the pan-genome analysis.

			Summary statistics				
	Samples	Command settings 1 & 2	Core genes	Soft core gene	Shell genes	Cloud genes	Total genes
			99% <=strains <=100%	95% <=strains <=99%	15% <=strains <=95%	0% <=strains <=15%	0% <=strains <=100%
Prokka	MG_Roary_ref	Roary_ref-S	349	0	571	856	1776
		Roary_ref+S	411	0	341	211	963
	MG_Roary_refs	Roary_re-S	322	170	453	981	1925
		Roary_ref+S	409	176	167	2489	1000
	MG_Roary_samples	Roary_re-S	342	0	574	886	1802
		Roary_ref+S	406	0	348	215	969
RAST	MG_Roary_ref	Roary_re-S	353	610	610	996	1959
		Roary_ref+S	421	359	359	288	1068
	MG_Roary_refs	Roary_re-S	353	605	605	1006	1964
		Roary_ref+S	421	359	359	288	1068
	MG_Roary_samples	Roary_re-S	354	614	614	975	1943
		Roary_ref+S	417	364	364	293	1074

Table 2.7 shows the total number of genes found in all the different strains of the species, separating them into categories. A single reference was placed at the top of the list in MG_Roary_ref, all references were placed before the samples in MG_Roary_refs, and all the samples were placed before the references in MG_Roary_samples.

Table 2.8: List of published and newly synthesised primers for testing and use in the study.

Primer	Sequence	T_m (°C)	T_a (°C)	GC %	Amplicon (bp)
MG ABC For	CGCTGTTGACTCACAAGGAAT	62	57	48	133
MG ABC Rev	CACATCTCCAGCAAACGATCTAC				
MG Purine For	ACAGTTATGGGTCACGGGAT	63	58	50	113
MG Purine Rev	ACTAAAGCGCCACATGAACC				
MG Putative For	TGTTCCCTTTGCACGATCAG	62	57	49	113
MG Putative Rev	CCTTGAAGAGCATCTAGGGTTTG				
[1]MG mgc2 For	TTGGGTTTAGGGATTGGGATT	65	60	44	<100
[1]MG mgc2 Rev	CCAAGGGATTCAACCATCTT				
[2]MG pvpA For	GCCAMTCCAACTCAACAAGCTGA	62	57	57	>400
[2]MG pvpA Rev	GGACGTSGTCCTGGCTGGTTAGC				
MS FtsZ For	GACTACGCTGACGTGGTTAAA	62	57	48	83
MS FtsZ Rev	GCTCTATCTTTACCGGTAGCTTG				
MS Inorganic For	GGATTTGTGCCAAGCACTTTAG	62	57	46	134
MS Inorganic Rev	TCGCCATCGTCAACCATTT				
MS NAD-dep For	GATACCGTGATTTCTGCTGCAT	63	58	45	119
MS NAD-dep Rev	GCAGTATATGAGTGAACGGTTGTC				
[3]MS vlhA For	CCAGGAGGTGGTACAGTTGAC	63	58	43	<100
[3]MS vlhA Rev	TTAATGCTTCTTTAACTGAATCTGA				
[4]MS vlhA MS For	GATGCGTAAAATAAAAGGAT	63	58	38	>300
[4]MS vlhA MS Rev	GCTTCTGTTGTAGTTGCTTC				
Mgal Glycyl For	CCAAGGAGTAACTGAAGATAGC	60	55	45	131
Mgal Glycyl Rev	CGACTTGACCAACTCCAAAT				
Mgal oligo For	CGTGGTAGATTACGTTCAAATGGG	63	58	46	114
Mgal oligo Rev	GCTAACTTGGTCTCAGCTAGAAAG				
Mgal Peptide For	TTGCTGGTTCAGGAGCTAGA	63	58	48	108
Mgal Peptide Rev	CCATAACTGGTGAAAGGCTAGTTG				
Mgal Endo For	CGTGCTGGAGTGACTGTTT	61	56	49	149
Mgal Endo Rev	GCTACATCAGGATTTACCACAC				
[5]Mgal MGC For	CGTGCCCCCTTGATTGGGATAACGCTG	60	55	55	>100
[5]Mgal MGC Rev	TAGCTAATGTTACGCACCCCGATCCCCTTGT				
Mpul Transl For	ACAGACGTAACAGGTGGAGTT	63	58	49	102
Mpul Transl Rev	AACCGCGATAGGAGCGATTA				
Mpul RpiR For	ATGAAGGACTTGGTGGGTT	61	58	46	85
Mpul RpiR Rev	GCACTAACTGGACTGTCAATTC				
Mpul N6-L For	GTGGTGGTGTTAGTGCGAATA	62	57	48	120
Mpul N6-L Rev	ATGCGGTCTGAGCAATCATAG				
Mpul 4deoxy For	CACTCAGGTGTTGGAACATCAA	62	57	45	84
Mpul 4deoxy Rev	AACACCTTGCATGTCGGTAT				
Mpul Serine For	TCCTGAACTAGCGGATCCAAA	63	58	48	117
Mpul Serine Rev	CGTGCATCGCTACCTTGAAAT				
[6]16S RNA GC-341 F	CCTACGGGAGGCAGCAG	61	56	59	>100
[6]16S RNA R543 R	ACCTATGTATTACCGCG				
[7]16S rRNA-spp For	AACACCAGAGGCAAGGCGAGG	60	55	51	>400
[7]16S rRNA-spp Rev	ACGGATTTGCAACTGTTTGTATTGG				

[1](Raviv and Kleven, 2009), [2](Felice et al., 2020, Hashemi et al., 2018), [3](Felice et al., 2020, Hashemi et al., 2018), [4](Amer et al., 2019, Felice et al., 2020, Hong et al., 2004, Moscoso et al., 2004), [5](Adeyemi et al., 2018), [6](McAuliffe et al., 2005), [7](Kiss et al., 1997b).

2.3.3 Primer testing using conventional PCR

Amplification of the targeted sequences was successful, as visualized by agarose gel electrophoresis. Figure 2.1-2.4 show agarose gel electrophoresis results for primer testing on target species, Figure 2.5-2.8 show agarose gel electrophoresis results for primer testing on non-target species, and Figure 2.9 shows agarose gel electrophoresis results of the group-specific primers.

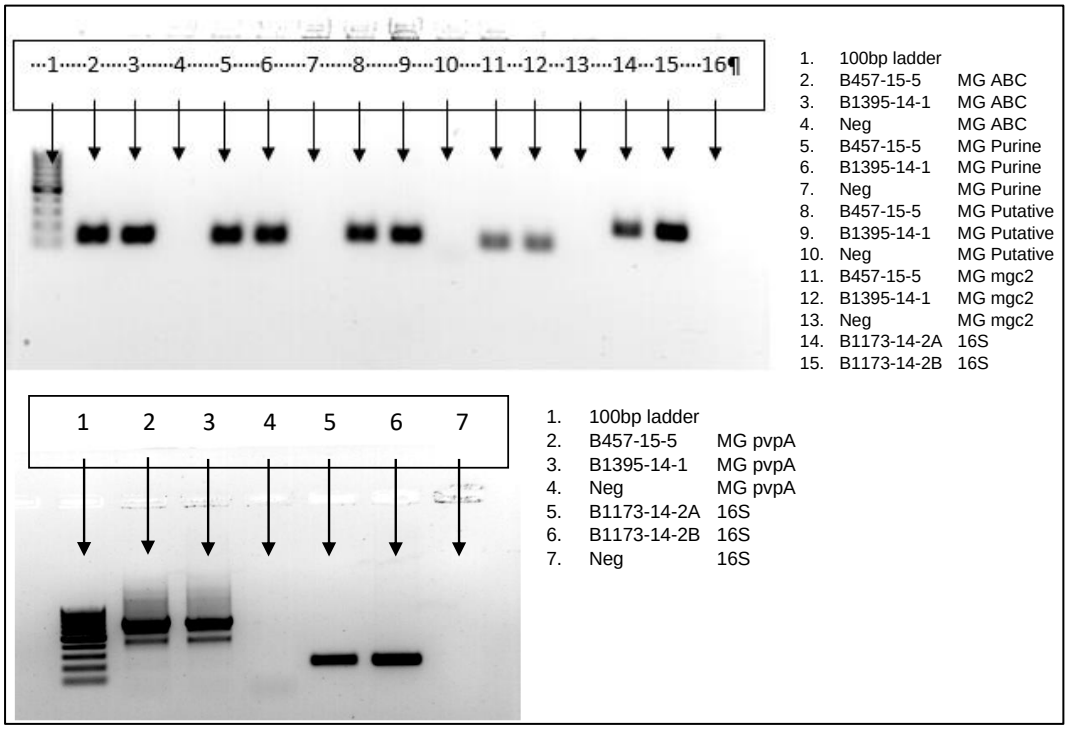

Figure 2.1: Agarose gel electrophoresis results for target species: MG.

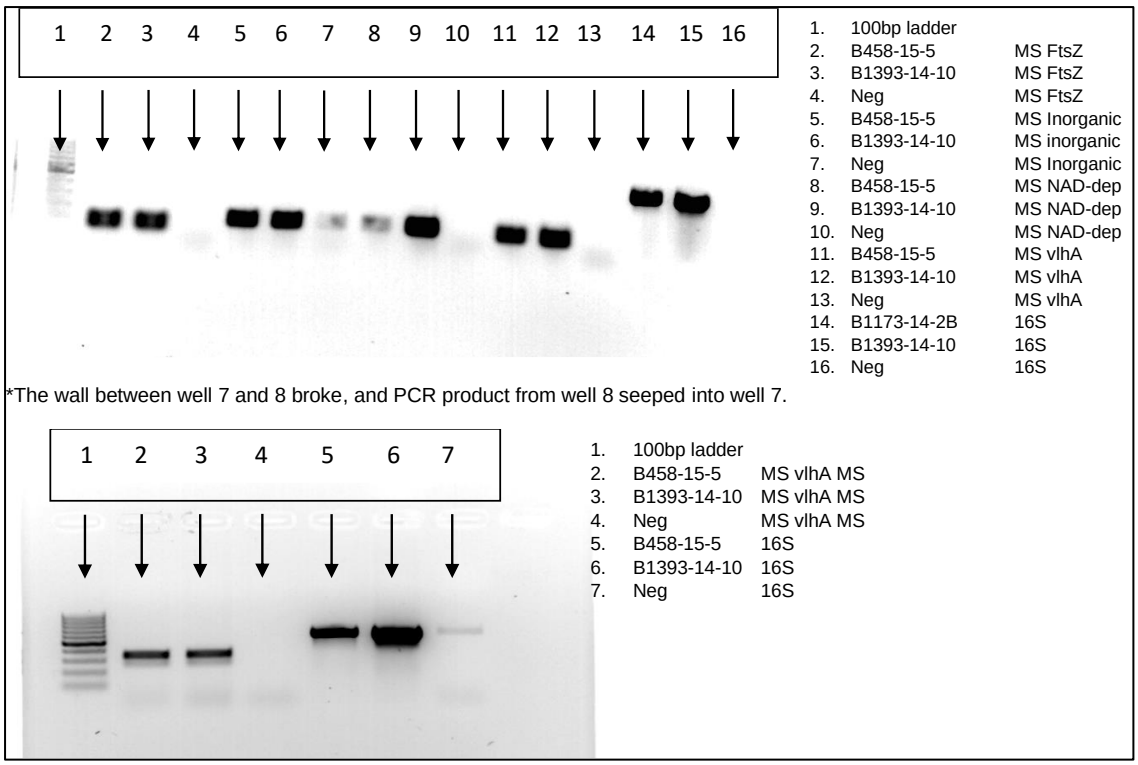

Figure 2.2: Agarose gel electrophoresis results for target species: MS.

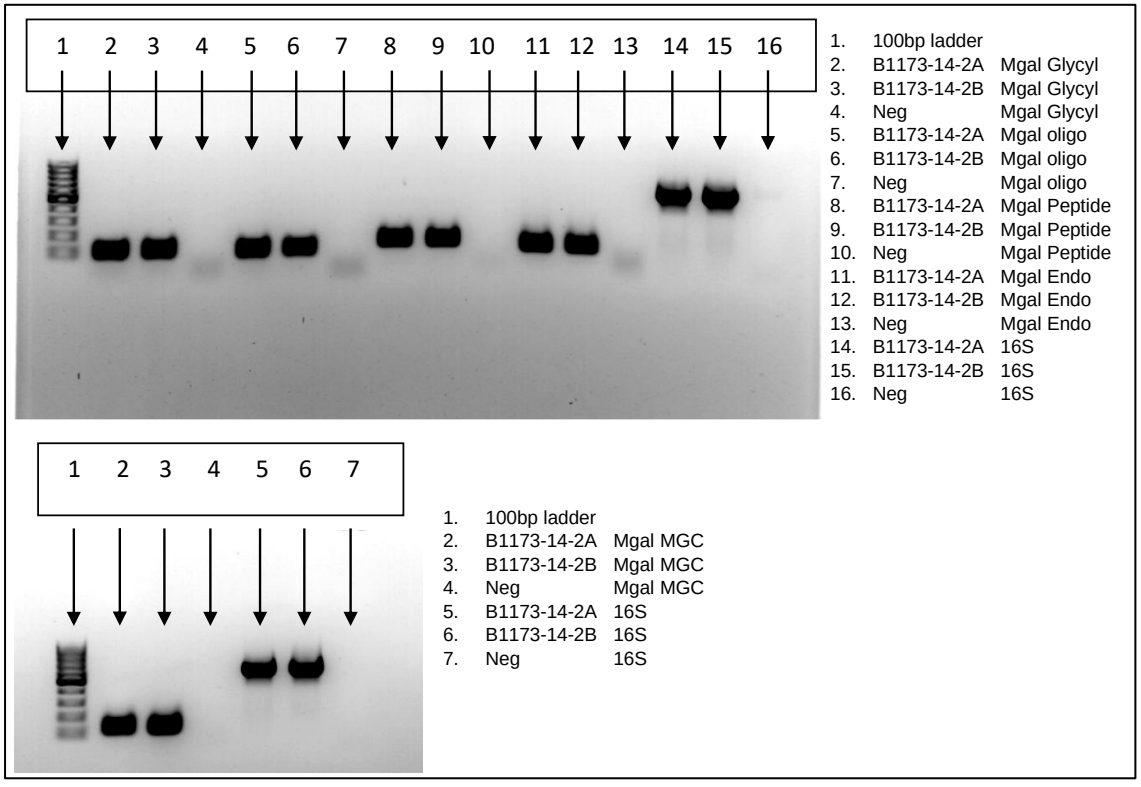

Figure 2.3: Agarose gel electrophoresis results for target species: *M. gallinaceum*.

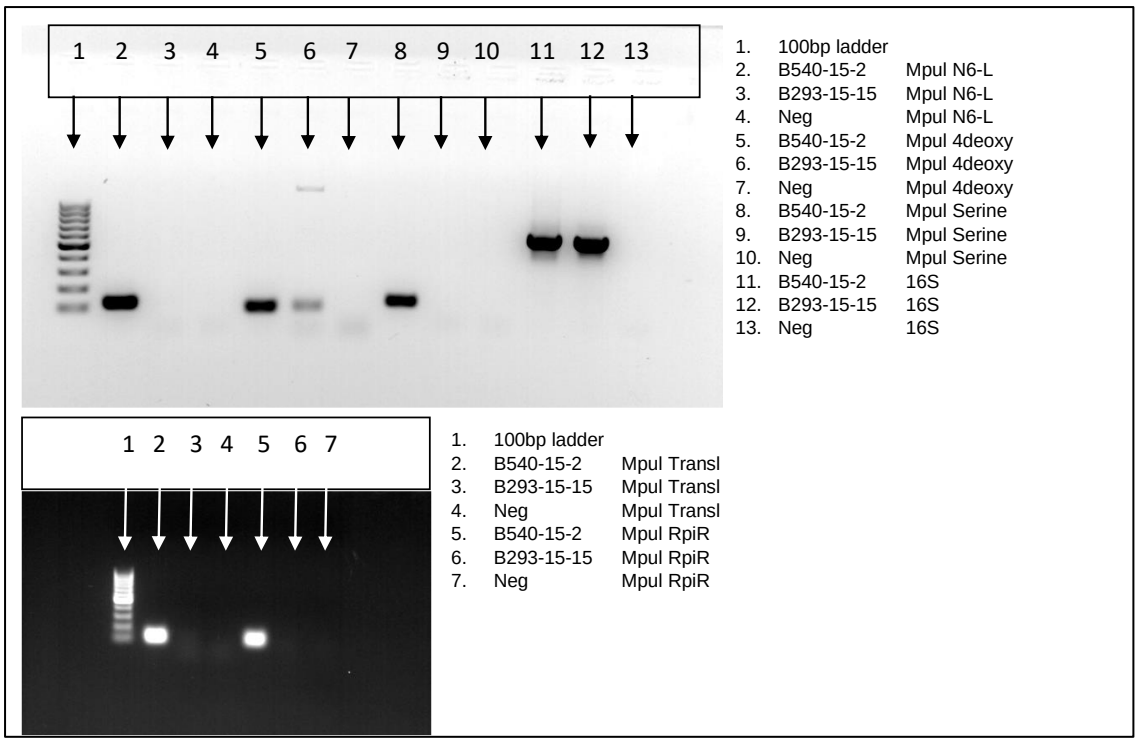

Figure 2.4: Agarose gel electrophoresis results for target species: *M. pullorum*.

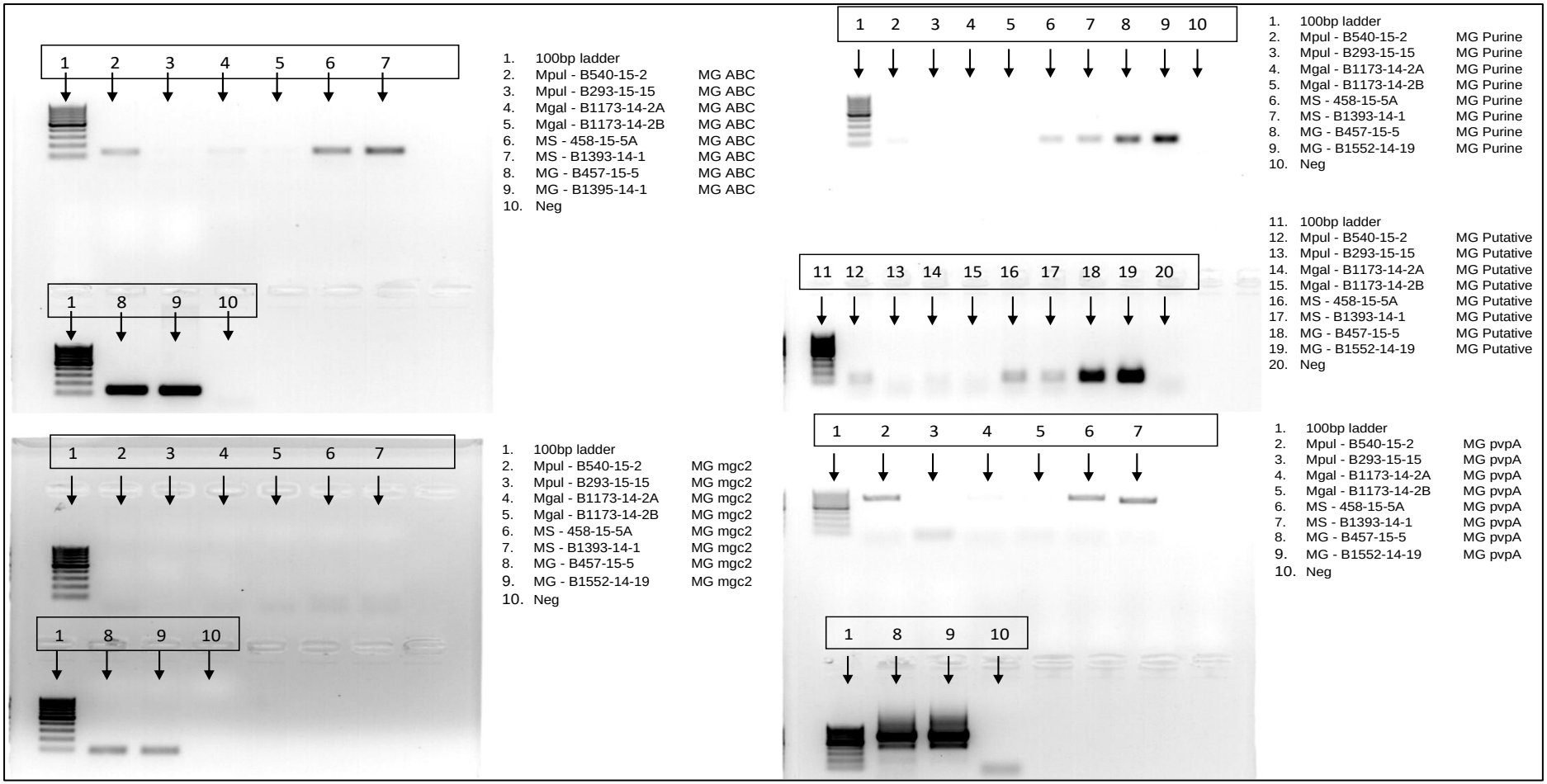

Figure 2.5:Agarose gel electrophoresis results for non-target species: MG.

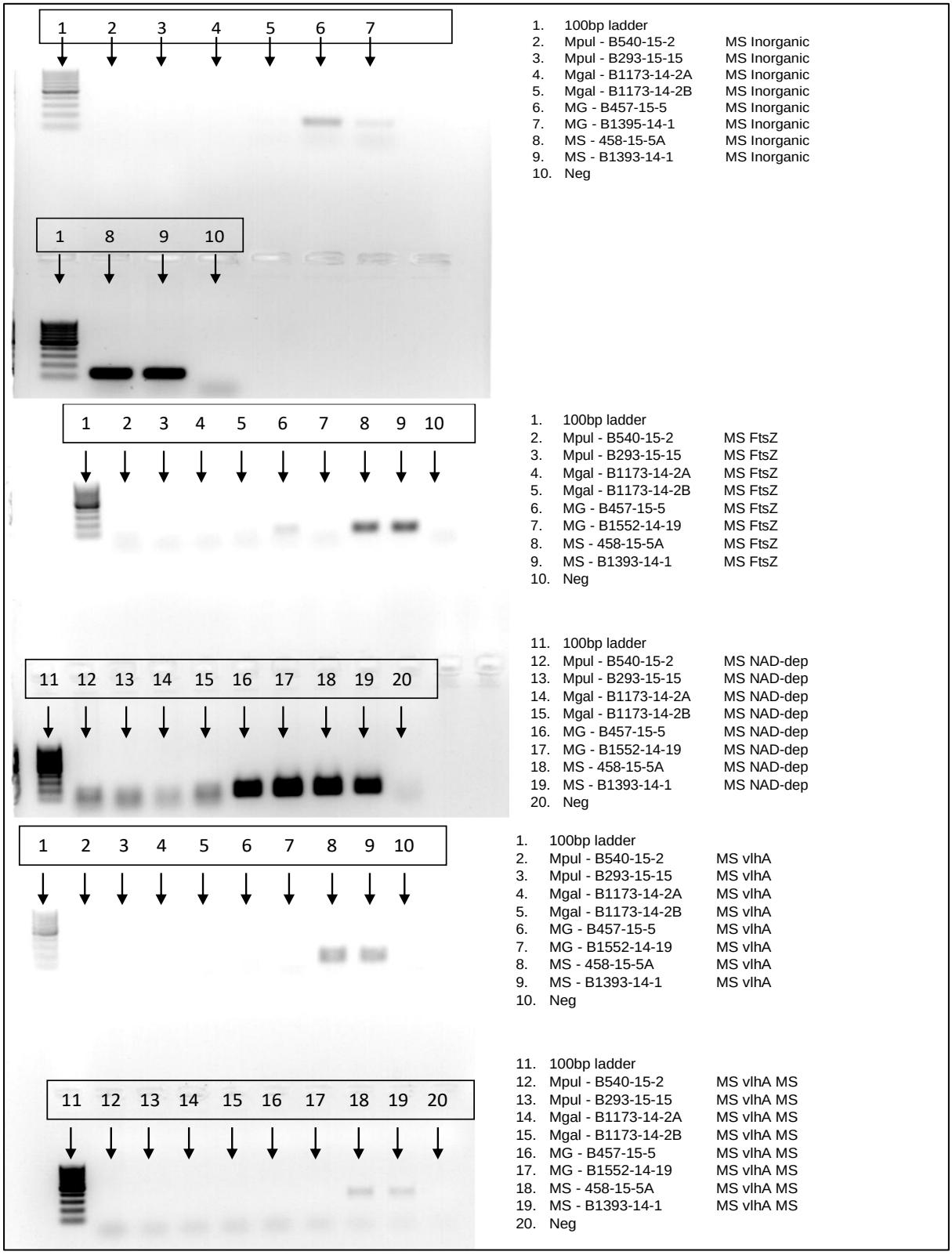

Figure 2.6: Agarose gel electrophoresis results for non-target species: MS.

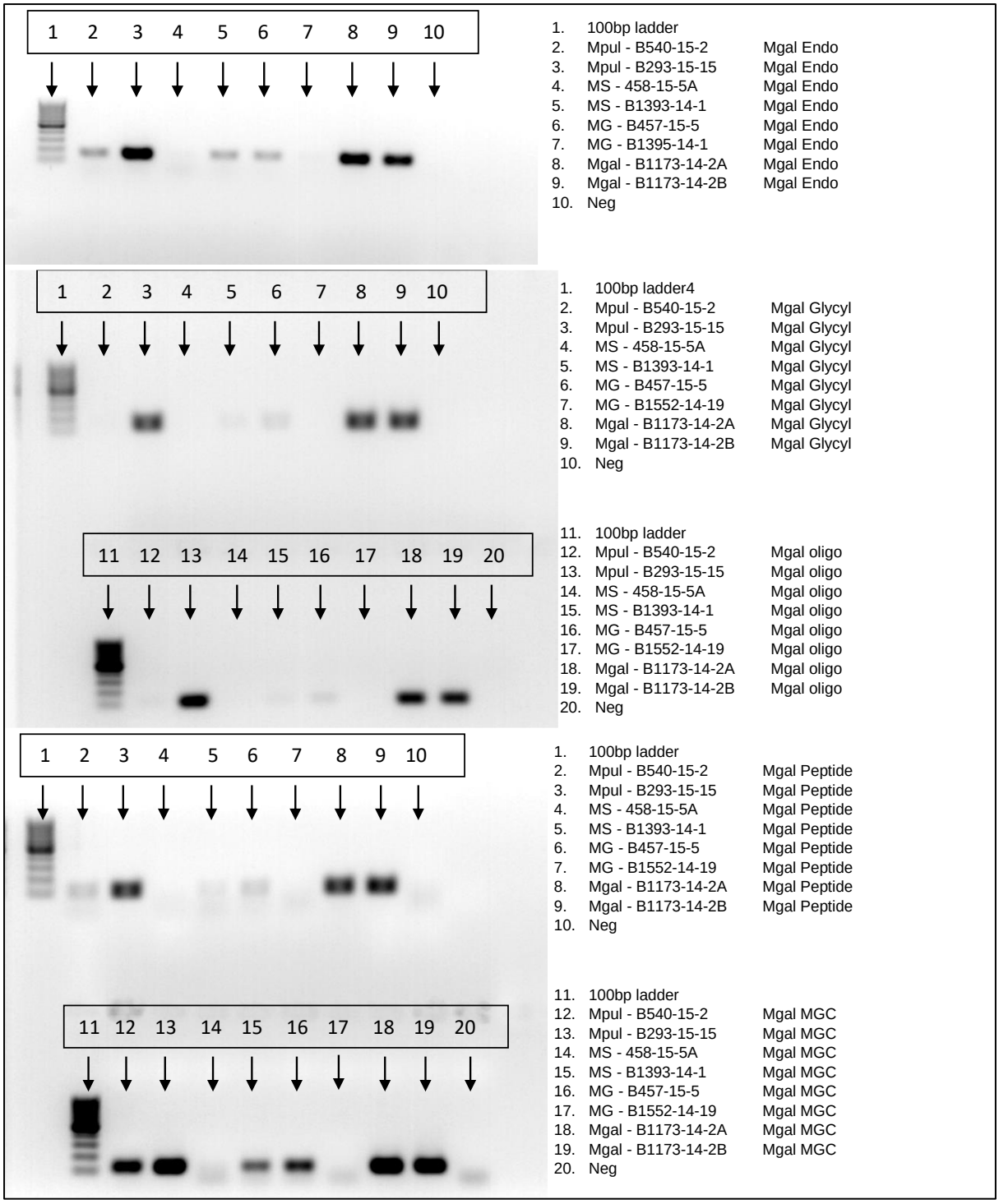

Figure 2.7: Agarose gel electrophoresis results for non-target species: *M. gallinaceum*.

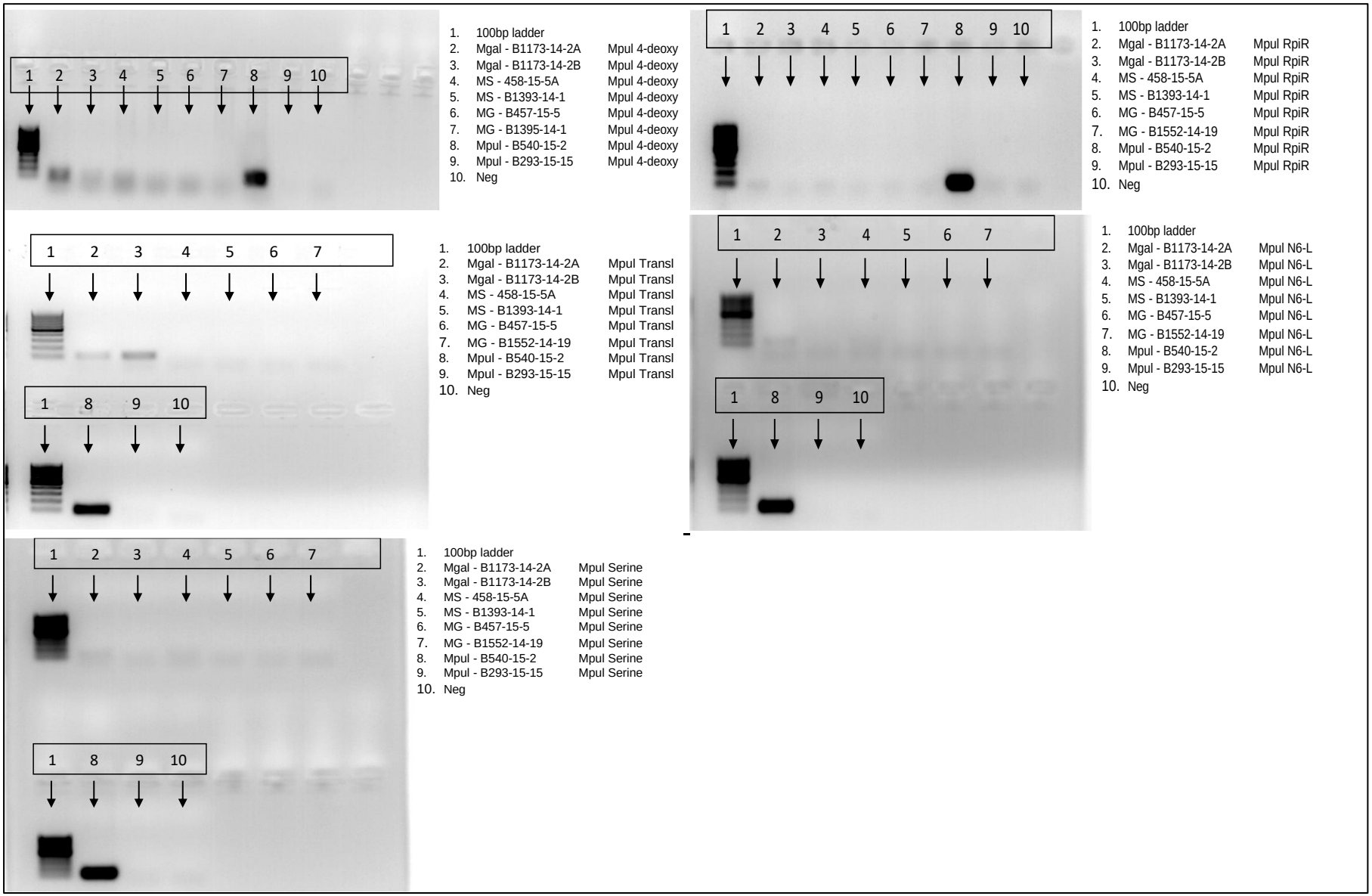

Figure 2.8: Agarose gel electrophoresis results for non-target species: *M. pullorum*.

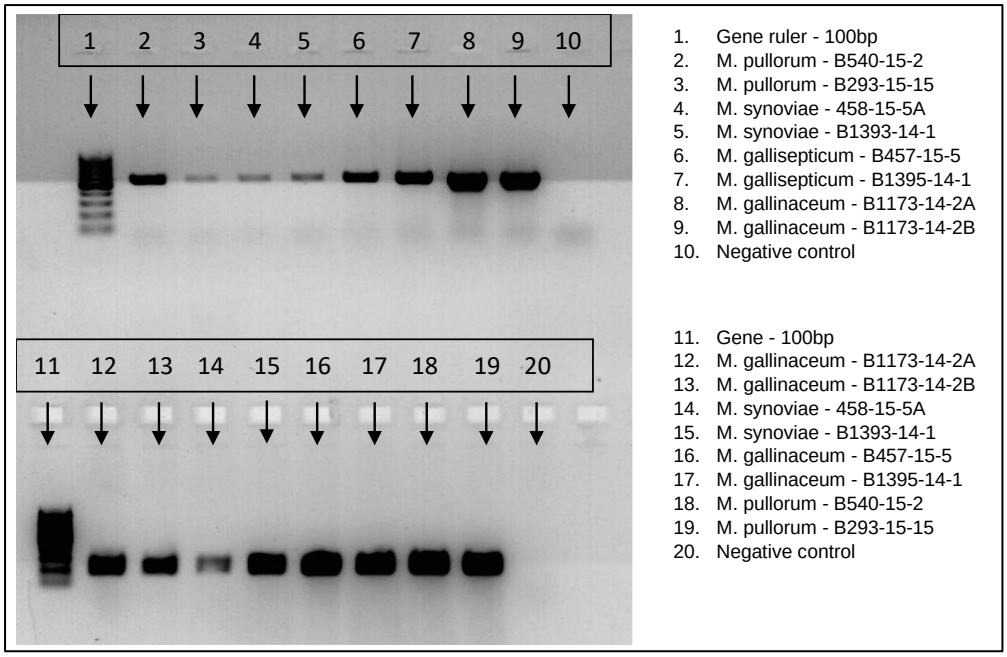

Figure 2.9: Agarose gel electrophoresis results for the group-specific primers.

The 16S rRNA-spp primer pair was used in numbers 1 through 10; and the 16SRNA GC-341&R543 primer pair was used in numbers 11 through 20. Both primer sets showed amplification in all the samples tested.

One group-specific primer set, and one primer set specific for each species, were then selected for TaqMan® Minor Groove Binder (MGB) probe synthesis. The selected primer sets were mgc2 (MG), vlhA (MS), oligo (*M. gallinaceum*), RpiR (*M. pullorum*) and the 16S RNA GC-341&R543 (group-specific).

2.4 DISCUSSION

Bioinformatics involves working with large amounts of data giving researchers access to existing information as well as allowing them to input their findings as they receive it; and this is achieved within a short period using the applicable computational tool(s). Although these tools do produce results rapidly, this is only the case if one is using a high-performance computer. One of the challenges met in completing Roary runs for the pan-genome analysis was that the computer used was a standard computer (laptop), in which case, runs that could take as little as 10 minutes took a minimum of 30 minutes to about 90 minutes to complete. Table 2.9 lists a few specifications that can be considered when selecting a computer to run these bioinformatics programs/tools faster, compared to the standard computer (laptop) used initially. The

same is also true for the alignments done in CLC Genomics Workbench. This programme was purchased and installed on a computer at campus (Faculty of Veterinary Science) of the University and cannot be accessed after hours or for evening work; as such, the alignments were done remotely using TeamViewer™ (available at https://www.teamviewer.com). The challenge, in this case, was that the longer the CLC Genomics Workbench was accessed through TeamViewer™, the slower the computer (laptop) became.

Table 2.9: Minimum recommended computer specifications compared to the computer used initially.

Specifications		
	Standard (Laptop)	Recommended
Processor	Intel Celeron N4000	Intel Core I5, I7 or similar
Memory (RAM)	8 Gb	8 Gb
Operating system	64-bit	64-bit, x64-based processor
Storage	500 Gb	500 Gb (Solid state drive)

Following the pan-genome analysis, the RAST data was selected over the Prokka data as there were a higher number of genes present overall. This preference was used on the basis that there could have been more genes identified in the RAST data that would have possibly been missed in the Prokka data. The data obtained from command 3 was kept for use in Piggy, providing that not enough genes were obtained from commands 1 and 2.

For data analysis, MG and *M. gallinaceum* data yielded a significantly higher number of genes than MS and *M. pullorum* because they each had more than one reference sample. The sample list '*_Roary_ref*' and '*_Roary_refs*' used in each species for command settings 1 and 2 meant that more data was input which would then result in more genes detected.

Genes that were not listed by their proper names were deleted on the basis that they have yet to be identified and listed in the Prokka/NCBI database, meaning they are still only 'hypothetical' genes. Gene length was used as another criterion for deletion based on the nature of the PCR to be designed. In real-time PCR, target DNA strands are usually short fragments, and the product length is about 150 bp. The short genes were deleted to ascertain that target regions, from which the primers and probes would

be designed, could be easily identified without the worry of insufficient nucleotides. The long genes were deleted on the same basis, but in this case, to prevent the identification of too many regions for the primers/probes which could lead to cross-reaction (among the different species).

Since the whole genomes of the refence strains (MG strain R(low), MS strain 53, *M. gallinaceum* strain B2096-14-8 and *M. pullorum* strain B359-15-6) were already available in the NCBI database, they were used as the template from which the selected genes would be designed. Gene names were used to search and download the complete sequences of the genes, from the respective strain in each species. It was interesting to note that while searching for the sequences, not all the genes selected could be found in the genomes, meaning that at most the annotation might indicate possible protein product.

Primer synthesis for each species was as follows: there were three newly synthesised primer sets, and two published primer sets for both MG and MS; four newly synthesised primer sets, and one published primer set for *M. gallinaceum*; and all five primer sets for *M. pullorum* were newly synthesised. After testing the primers, specific amplification products were obtained for almost all the samples. When testing for primers on target species, only one *M. pullorum* sample (B293-15-15) showed no amplification.

However, when initially testing the primers on non-target species, amplification was obtained for non-target species. The non-specific binding was as follows: MS and *M. pullorum* for MG, MG for MS, *M. gallinaceum* for *M. pullorum* and MG, MS, and *M. pullorum* for *M. gallinaceum*. The PCR conditions were adjusted to reduce cross-reactivity and samples that still showed strong cross-reactions were sent for Sanger sequencing. Only sample (B293-15-15) was identified as an *M. gallinaceum* species after Sanger sequencing. It was therefore evident that the sample had been contaminated with an *M. gallinaceum* strain before PCR. This would then explain why there was no amplification observed when this sample was tested for *M. pullorum*. Figure 2.5-2.8 show agarose gel electrophoresis results before optimisation of PCR conditions.

Selection of the group- and species-specific primers for probe synthesis was based on successful amplification, coupled with small product size (between 50-150 bp (Pestana, 2010)) and little to no cross-reactivity.

Initially, *M. gallinarum* and *M. iners* were included in this study but were later excluded because the culturing of these two species was unsuccessful, possibly because the stock in the DVTD sample repository, is quite old. The samples kept in the repository were first isolated between 2003 and 2015, followed by sequencing and identification between 2015 and 2018 (Beylefeld, 2018). Culturing from samples that may have undergone numerous amounts of freezing and thawing has proven very difficult. Another challenge that was met was the financial constraint of simultaneous synthesis of primers and probes. Probes are generally more expensive to synthesise than primers. This then led to testing the primers initially, using conventional PCR, thereafter, only the probes for the primers to be used throughout the study would be synthesised.

In conclusion, primers and probes for the PCR detection and differentiation of *M. gallisepticum*, *M. synoviae*, *M. gallinaceum*, *M. pullorum,* as well as *Mycoplasma* species (group-specific) were successfully designed/synthesised from literature, and PCR conditions optimised.

CHAPTER 3: DESIGN, DIAGNOSTIC EFFICIENCY, AND TESTING OF A MULTIPLEX REAL-TIME PCR ASSAY

3.1 INTRODUCTION

A multiplex PCR is an assay designed to detect multiple pathogens using a single primer set or more in a reaction (Mackay, 2004, Viljoen et al., 2005). Mackay, (2004) describes multiplex real-time PCR as “the use of multiple fluorogenic oligoprobes for the discrimination of amplicons that may have been produced by one or several primer pairs” (Mackay, 2004). Developing a multiplex real-time PCR can be challenging due to the limited number of available fluorophores (Mackay, 2004), but not impossible. Wittwer et al., (2001) propose multiplexing by T_m, where the multiple fluorophores used are combined with the discrimination of additional targets allowing the detection of a larger number of target amplicons (Mackay, 2004, Wittwer et al., 2001). Henegariu et al., (1997) developed a protocol that can be used when designing a multiplex PCR, and although the focus is on conventional PCR, it is possible to adapt the parameters to real-time PCR. There are five steps in the protocol: (1) choice of primers, (2) testing/aligning primer sequences, (3) single locus PCR, (4) multiplex PCR, and (5) optimisation of the multiplex (Henegariu et al., 1997).

Validation of real-time PCR involves two main parameters, assay specificity and assay performance. Assay specificity refers to primer/probe target specificity and this is evaluated by *in silico* validation (Broeders et al., 2014). Assay performance involves several steps which include, but are not limited to, assay efficiency, linearity, analytical specificity, analytical sensitivity, and accuracy (Broeders et al., 2014). Assay efficiency is defined as the doubling of target molecules after the completion of each amplification cycle, is derived from a serial dilution and is expressed as a percentage (Rebrikov and Trofimov, 2006). Linearity is the measurement of the response of an assay and is also determined from a serial dilution, where the data plotted in a spreadsheet is used to calculate the correlation coefficient (R^2) value of the curve (Broeders et al., 2014). Analytical specificity (selectivity) is the level to which the assay distinguishes the target agent from other infectious agents in the sample matrix (Kralik and Ricchi, 2017). Analytical sensitivity, also referred to as the limit of detection (LOD), is the smallest detectable amount of analyte in a sample that can be measured with a

defined certainty but not quantified as an exact value; and represented as either infectious dose, plaque-forming units, number of genome copies, or colony-forming units (CFUs) etc., that is detected/distinguished from a zero result (Health, 2014a). Lastly, accuracy is also referred to as diagnostic specificity (DSp) and diagnostic sensitivity (DSe). DSp is the proportion of known uninfected reference animals or target analytes that correctly test negative in the tested assay; and DSe is the proportion of known infected reference animals or target analytes that correctly test positive in the tested assay (Health, 2014a).

In Chapter 2, twenty primer sets for the PCR detection and differentiation of *M. gallisepticum*, *M. synoviae*, *M. gallinaceum*, *M. pullorum*, as well as two primer sets for *Mycoplasma* species (group-specific) were designed/synthesised from literature and tested, whereafter, one set of oligonucleotide primers for each species was selected and the respective probes synthesised. The group-specific probe was designed to detect most mycoplasma species.

In this chapter, a multiplex real-time PCR assay using the selected and new oligonucleotide primers and probes was optimised and validated and the diagnostic efficiency was determined. The assay was then used to test (a) a panel of known cultured samples from the UP repository and (b) field samples collected from farms known to have persisting mycoplasma infections, to determine how well the assay would be able to detect and distinguish between *Mycoplasma* species.

3.2 MATERIALS AND METHODS

3.2.1 Positive controls and real-time PCR optimisation

A plasmid containing a synthetic construct encoding the target sequences for the specific forward, probe, and reverse oligonucleotides of all four species for use as a laboratory positive control (Figures 3.1 and 3.2) was designed and synthesized by Inqaba Biotec™ (Pretoria), where the cloning vector used was pUC57-Simple (Biomatik Corporation - Canada) with the position of the insert from 445 bp to 802 bp (368 bp in total). The starting concentration of the plasmid was 1.49×10^{11} plasmid copies/µl. Plasmid concentration for real-time PCR testing was the same as the primers in 2.2.4.2.

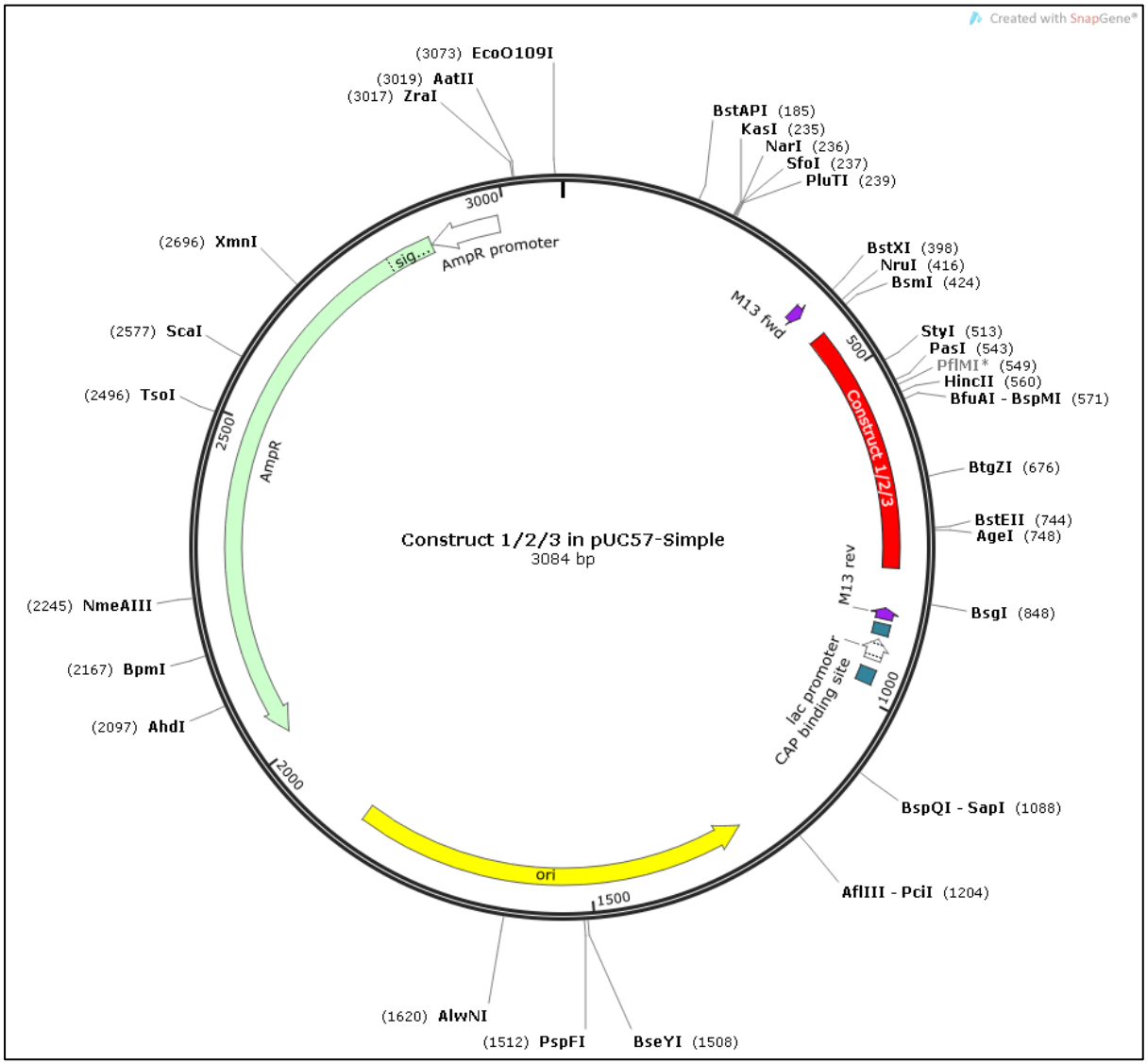

Figure 3.1: Plasmid map for the synthetic positive control.

```
301 TACGCCAGCT GGCGAAAGGG GGATGTGCTG CAAGGCGATT AAGTTGGGTA
351 ACGCCAGGGT TTTCCCAGTC ACGACGTTGT AAAACGACGG CCAGTGAATT
401 GGAGATCGGT ACTTCGCGAA TGCGTCGAGA TATCTTCTTT ACTCTTGGGT
451 TTAGGGATTG GGATTTGATG ATCCAAGAAC GTGAAGAACA CCTATTTTGT
501 TCTTCGATTA TTCCAAGGGA TTCAACCATC TTTCCAGGAA ACCCAGGGGG
551 TGGTACAGTT GACCTGCTAA AACAGAAGCT AAAACCGCTA TACTTGTCTT
601 TTTAATGCTT CTTTAACTAA ATCTGATGGA AATGTTCGTG GTAGATTACG
651 TTCAAATGGG TGGTGGCGGA CATGACAATG CTGCTAAGAA CATCGCTAAC
701 TTGGTCTCAG CTAGAAAGTC AAAATATTAT GAAGGACTTG GTGGGTTACC
751 GGTTGGTTCA AGTGGACGCA GAGTTTAAAA GAATTGACAG TCCAGTTAGT
801 GCGATATCGG ATGCCGGGAC CGACGAGTGC AGAGGCGTGC AAGCGAGCTT
851 GGCGTAATCA TGGTCATAGC TGTTTCCTGT GTGAAATTGT TATCCGCTCA
901 CAATTCCACA CAACATACGA GCCGGAAGCA TAAAGTGTAA AGCCTGGGGT
```

Figure 3.2: Positions of the oligonucleotide sequences used for synthetic construct 1/2/3.

The forward primers are highlighted in yellow, probes in blue and reverse primers in green. The individual constructs for each species are placed in the following positions: MG: nt 445 to 542, MS: nt 543 to 626, *M. gallinaceum:* nt 627 to 718, *M. pullorum*: nt 719 to 802.

Reference strains used as positive controls were MG NCTC 10115, MS ATCC 25204, and *M. gallinaceum* B2096-14-8, while the plasmid was used for *M. pullorum*. Extraction of nucleic material from the reference strains was done using the user-developed protocol: Purification of genomic DNA from cultured cells using the QIAamp® DNA Micro Kit (available at https://www.qiagen.com/us/resources/). The concentration for MG NCTC was 6.77 × 10^{10} colony forming units/µl (CFUs/µl), while the concentration for both MS ATCC 25204 and *M. gallinaceum* B2096-14-8 was 6.72 × 10^{10} CFUs/µl. The group-specific primers and probe were tested using the sample representatives listed in Table 2.5.

3.2.2 Conventional PCR (cPCR)

The Ampliqon Taq DNA Polymerase 2X Master Mix RED (ThermoFisher Scientific) was used and prepared according to manufacturer's recommendations. Each PCR reaction contained 12.5 µl of Taq 2X Master Mix, 1 µl each of forward and reverse primer, 8.5 µl of PCR-grade water, and 2 µl of nucleic acid extracts (PCR-grade water for negative control). Amplification was done in the Veriti™ 96-Well Thermal Cycler, starting with 1 cycle of initial denaturation at 95 °C for 2 minutes, followed by 30 cycles

of denaturation at 95 °C for 20 seconds, annealing at 59 °C (MG)/ 64.5 °C (Mspp) for 20 seconds and extension at 72 °C for 30 seconds; and a final extension step at 72 °C for 5 minutes (1 cycle). Visualisation of results was done as in 2.2.4.

For the cPCRs two sets of primers were used. The Universal 16S Mycoplasma GPO3F and MGSO (Van Kuppeveld et al., 1994); the forward primer was used as in the paper, the reverse was replaced with a newly designed primer that could detect all known avian mycoplasmas except MG and *M. imitans* (not published). The MGSO were used as references for the cPCR testing because of the product size (700 bp) which yielded better quality Sanger sequencing results as compared to the McAuliffe group-specific primers that yeilded a product size of about 100 bp long. The reverse primer was replaced due poor binding performance and the replacement primer (also yeilding a 700 bp product size) could not detect MG and *M. imitans*. To include MG, the MG 16S PCR MG-13R and MG-14F (Lauerman, 1998) was used (WOAH: https://www.woah.org/fileadmin/Home/eng/Health_standards/tahm/3.03.05_%20AVIAN_MYCO.pdf).The primer sequences are listed in Table 3.1.

Table 3.1: Primes used for cPCR testing

Species	Primer	Sequence
Universal 16S Mycoplasma	GPO3F	TGGGGAGCAAACAGGATTAGATACC
	Mspp-R	AGACCCGAGAACGTATTCAAC
MG 16S PCR	MG-14F	GAGCTAATCTGTAAAGTTGGTC
	MG-13R	GCTTCCTTGCGGTTAGCAAC

3.2.3 Real-time PCR

The TaqMan® Minor Groove Binder (MGB) probes, synthesised by LTC Tech South Africa (PTY) LTD (Applied Biosystems®), were based on the selection made in 2.3.3 and are listed with the respective primers in Table 3.2.

Table 3.2: List of TaqMan® MGB probes and corresponding primers used for the multiplex real-time PCR testing.

Species	Probe name	Probe Sequence	Primer Sequence
MG	MG-mgc2 AE015450	TGATGATCCAAGAACGTGA AGAACACC	F-TTGGGTTTAGGGATTGGGATT R-CCAAGGGATTCAACCATCTT
MS	MS-vlhA AF035624	CTGCTAAAACAGAAGCTAAA ACCGCTAT	F-CCAGGAGGTGGTACAGTTGAC R-TTAATGCTTCTTTAACTGAATCTGA
M. gallinaceum	Mgal-oligo	TGGTGGCGGACATGACAAT GCTGC	F-CGTGGTAGATTACGTTCAAATGGG R-GCTAACTTGGTCTCAGCTAGAAAG
M. pullorum	Mpul-RpiR	ACCGGTTGGTTCAAGTGGA CGCA	F-ATGAAGGACTTGGTGGGTT R- GCACTAACTGGACTGTCAATTC
M. species (group-specific)	16S McAuliffe	TGATGGAGCGACACAGCGT G	F-CCTACGGGAGGCAGCAG R- ACCTATGTATTACCGCG

A 10X working solution of primer/probe mix was made for each species, by adding 9 µl each of the forward- and reverse primers (100 µM), and 2.5 µl of the respective TaqMan® MGB probe (100 µM) to 79.5 µl of PCR-grade water. For the real-time qPCR reactions, the VetMax™-Plus qPCR Master Mix kit (Applied Biosystems®) was used and prepared according to the manufacturer's recommendations. Each PCR reaction contained 6.25 µl of VetMax™-Plus qPCR master mix, 1 µl of the respective primer/probe mix, 1.25 µl of PCR-grade water, and 3 µl of nucleic acid extracts (PCR-grade water for the negative control). The real-time PCR reactions were performed using the StepOnePlus™ Real-Time PCR system (Applied Biosystems®) using the recommended conditions. Denaturation was set at 95 °C for 10 minutes, followed by 15 seconds of annealing at 95 °C (adjusted), and the region of interest extended at 60 °C for 45 seconds. For the multiplex real-time PCR, MG and MS were pooled together in one mix, while *M. gallinaceum* and *M. pullorum* were pooled together in another mix.

3.2.4 Assay efficiency, measurement response linearity and LOD

Assay efficiency and linear range were determined using serial dilutions for each species. Ten-fold serial dilutions (10^{0} to 10^{-10}) of the references/controls were prepared in biotechnology grade Tris-EDTA (TE) buffer (VWR® Life Science - Avantor™ USA). Each dilution series was tested in single locus real-time PCR runs; this was repeated a minimum of three times. Standard curves were then plotted to show cycle threshold (C_t) values versus log plasmid copies/CFUs per µl in each reaction. The slope of the regression line for each standard curve was then used to calculate assay efficiency using the formula: PCR efficiency (%) = $(10^{-1/slope} - 1) \times 100$. The serial dilutions were also used to conservatively estimate the limit of detection (LOD) at the last dilution series where 100% detection was observed, and the threshold cut-offs were determined from the mean of the C_t values from all the replicates at this dilution.

3.2.5 *Mycoplasma* species detection

DNA extracted from cultured samples, previously identified as either MG, MS, *M. gallinaceum*, *M. pullorum*, *M. gallinarum*, and/or *M.iners*, (axenic [n=52] and mixed [n=33]) were used as known positive samples (Table 2.3). They were tested using the multiplex real-time PCR to determine how well the assay would detect and distinguish known positive samples that are either axenic or mixed. The *Mycoplasma* species in the mixed samples were determined by 16S rRNA gene identification in a previous study (Beylefeld, 2018), and the different species combinations (n=14) are listed in Table 3.3.

Table 3.3: Different species combinations in the known mixed samples

No.	*Mycoplasma* species combination
1	MG & MS
2	MG & *M. gallinaceum*
3	MG, *M. gallinaceum* & *M. gallinarum*
4	MG & *M. pullorum*
5	MG, *M. pullorum* & *M. gallinarum*
6	MG & *M. gallinarum*
7	MS & *M. gallinarum*
8	*M. gallinaceum* & *M. pullorum*
9	*M. gallinaceum*, *M. pullorum* & *M. gallinarum*
10	*M. gallinaceum*, *M. pullorum* & *M. iners*
11	*M. gallinaceum* & *M. gallinarum*
12	*M. pullorum* & *M. gallinarum*
13	*M. pullorum*, *M. gallinarum* & *M. iners*
14	*M. pullorum* & *M. iners*

3.2.6 Field sample testing

A total of 203 samples were collected and tested to compare how well the multiplex assay would accurately detect and distinguish *Mycoplasma* species. The sample size was determined with the assumption that the expected assay sensitivity would be 95 % (+/- 5 %) at the 95 % level of confidence, thus a total sample size of 102 infected chickens was determined; and based on the same assumption, 102 uninfected chickens were also required. The prevalence of Mycoplasma infections in sick birds was expected to be 50% and therefore the total sample size was the simple addition, for a total of 204 birds. Choanal cleft or tracheal dry swabs were collected by veterinarians, transported on ice, and sent to the laboratory within 24 hours of collection. Sample information is listed in Table 3.4.

Table 3.4: Information of the field sample collected for multiplex real-time PCR testing.

Sample ID	FR 5, 9,10, 16, 22, 45 & 47	FR 27, 28 & 29	V 1, 2, 3, 4, 6, 7 & 10	VL 1-20	NFS 1-13
Collection date	June 2022	June 2022	June 2022	June 2022	August 2022
Type of farm	Unknown	Unknown	Unknown	Unknown	Unknown
District/Province	Gauteng (Skeerkrans)	Gauteng (Skeerkrans)	Free State	Free State	Gauteng
Antibiotic treatment	Unknown	Unknown	*V1, 2 & 4 (treatment unknown)	Unknown	Unknown

*Flocks in the farms from which samples V1, 2 & 4 were taken had undergone antibiotic treatment. The specific antibiotics used are unknown.

Three methods of testing were employed: 1) cultivation and isolation (gold standard), 2) conventional PCRs (cPCR), and 3) multiplex real-time PCRs (qPCR). Figure 3.3 shows an outline of the steps taken in the testing methods used.

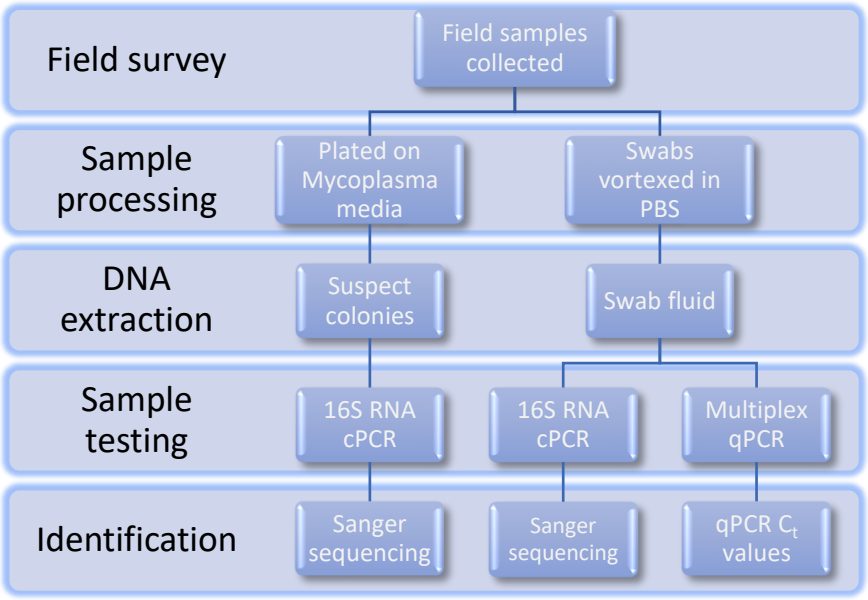

Figure 3.3 Flow diagram outlining the testing methods used.

In cPCR, results are visualised on agarose gels using electrophoresis (separation by band size), where a 1 % agarose gel, prepared in 1X tris acetate ethylenediaminetetraacetic acid (TAE) buffer and stained with ethidium bromide, is used to separate DNA fragments. The fragments (bands) are then compared to a GeneRuler™ 100 bp DNA ladder (ThermoFisher Scientific). In qPCR, results are observed in real-time and are expressed in cycle threshold (C_t) values.

3.2.6.1 Sample preparation

Choanal cleft or tracheal dry swabs were received at the University of Pretoria, Poultry Section, were stored at -80 °C and processed within 48 hours from collection. Mycoplasma agar (Oxoid, ThermoFisher Scientific) was prepared according to the manufacturer's instructions. The agar plates were prepared by suspending 35.5 g of Mycoplasma agar base in 1 L of distilled water, boiled to dissolve the agar, divided into 80 ml volumes, autoclaved at 121 °C for 15 minutes and then cooled down to 50°C. Mycoplasma selective supplement G (20 ml reconstitute) (Oxoid, ThermoFisher Scientific) was added to 80 ml sterilised agar, poured into petri dishes and allowed to set. The swab samples were then inoculated and streaked on the plates and incubated at 37 °C with 5% carbon dioxide (CO_2). The mycoplasma plates were incubated for

five days initially and monitored once a week thereafter. The swab samples were also vortexed in 1 ml of phosphate-buffered saline (PBS) (Sigma-Aldrich) and stored at -20°C for later use. A 50% (v/v) Mycoplasma storage broth was prepared using Mycoplasma broth (Oxoid, ThermoFisher Scientific) and 50% (v/v) glycerol. 100 µl of swab fluid was transferred to 1 ml of the Mycoplasma storage solution containing 50% (v/v) glycerol; and stored at -80°C for later use.

3.2.6.2 Mycoplasma isolation and DNA purification

Suspected colonies that showed the characteristic fried-egg morphology of mycoplasma were selected and the DNA was extracted directly from the agar. Using a sterile scalpel blade, an agar plug containing an entire colony was excised and placed into a sterile 1.5 ml Eppendorf tube. One hundred µl of Buffer QG from the QIAquick® Gel Extraction Kit (Qiagen) was added to the tube and incubated at 50°C for 10 minutes with occasional vortexing, until the agar melted. The Invitrogen™ PureLink™ Genomic DNA kit (ThermoFisher Scientific) for Gram-positive bacteria was then used with modified volumes to purify the DNA as follows: 30 µl of Proteinase K, 360 µl PureLink® Genomic Lysis/Binding buffer, and 360 µl of 100% alcohol. The lysozyme digestion buffer was prepared to a final concentration of 20 mg/ml and 180 µl (modified volume) was used. The extracted DNA was tested using the 16S RNA cPCR (Table 3.1) and then sent to Inqaba Biotechnical Industries (Pty) Ltd for Sanger sequencing. The results were analysed using Chromas© (DNA sequencing and analysis software) and the sequences were confirmed on the NCBI database using the Basic Local Alignment Search Tool (BLASTn) (available at https://blast.ncbi.nlm.nih.gov/Blast.cgi).

3.2.6.3 DNA extraction for PCR

Extraction of DNA from the samples stored in PBS was performed using the IndiMag Pathogen Kit on the Indical IndiMag automated system (Whitehead Scientific) using the manufacturer's instructions. Each extraction reaction contained 20 µl of Proteinase K, 200 µl of the sample, 100 µl of Buffer VXL, 400 µl of Buffer ACB, 25 µl of MagAttract Suspension, and 1 µl of Carrier RNA (1 µg/µl). The extracted DNA was then used for both conventional PCR (3.2.2) and real-time PCR (3.2.3). Conventional PCR

visualisation and preparation for identification (Sanger sequencing) was done as in 2.2.4.2, and analysis of sequencing results as in 3.2.6.2.

3.3 RESULTS

3.3.1 Positive controls and real-time PCR optimisation

The plasmid was designed and synthesised. Upon receipt, the plasmid was reconstituted according to the manufacturer's instructions and tested against the specific *Mycoplasma* species. Only *M. pullorum* was amplified successfully, MG and *M. gallinaceum* produced C_t values above 34, and no amplification was observed for MS (Figure 3.4). Different plasmid concentrations were also tested (1.49×10^9-, 1.49×10^{10}-, and 1.49×10^{11} plasmid copies/µl), however, the result remained unchanged.

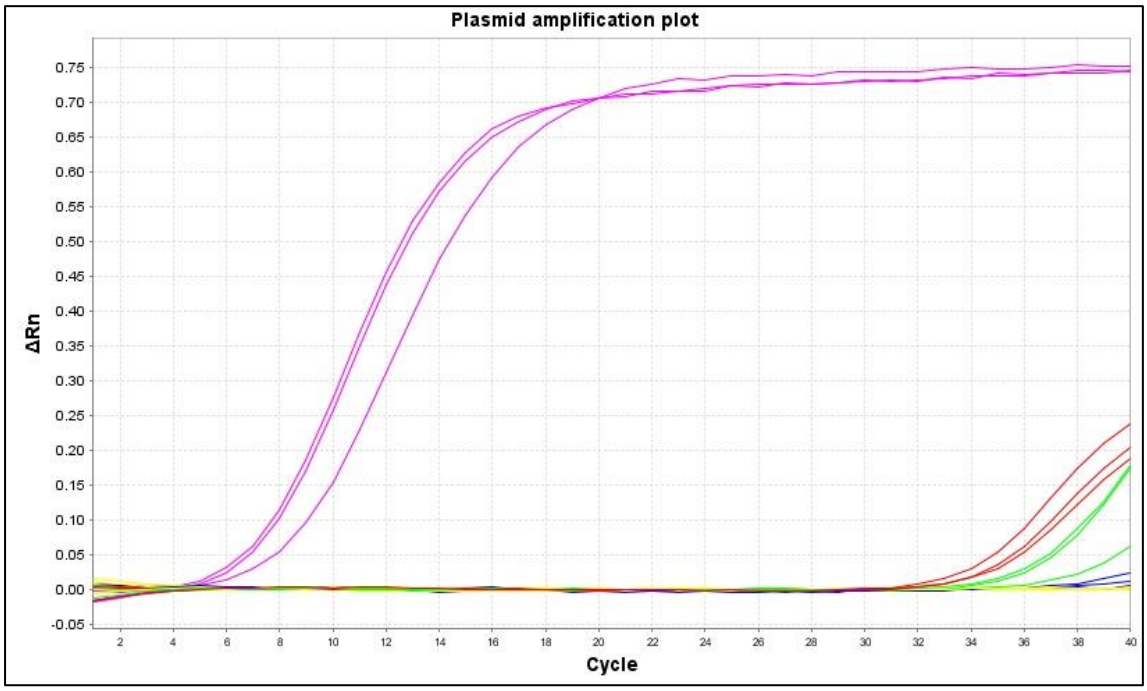

Figure 3.4: Amplification plot for MG, MS, *M. gallinaceum*, and *M. pullorum*.

The amplification curves are highlighted in red for MG, blue for MS, green for *M. gallinaceum*, pink for *M. pullorum*, and yellow for the negative controls.

When visualised using agarose gel electrophoresis, the plasmid showed a faint band above the gene ruler, indicating that there could be inhibitors present preventing amplification (Figure 3.5).

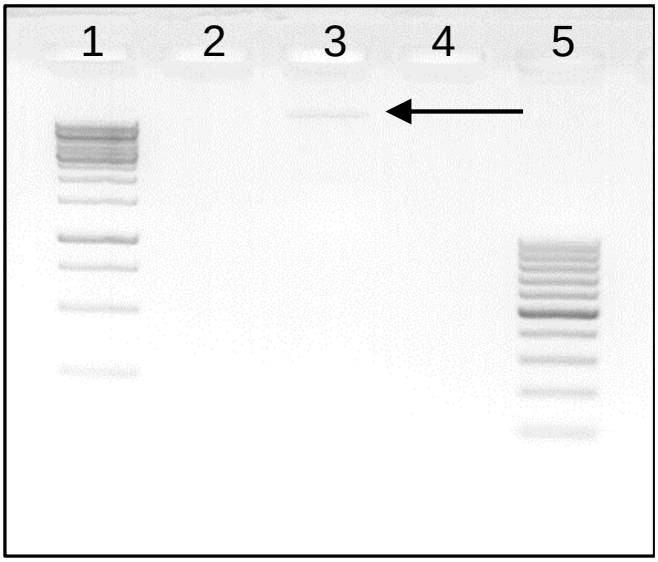

Figure 3.5: Agarose gel electrophoresis image of the plasmid laboratory control.

The plasmid is indicated by the arrow in lane 3. A GeneRuler™ DNA Ladder (ThermoFisher Scientific) was used in lane 1 (1 kb), and lane 5 (100 bp).

To mitigate the presence of possible inhibitors still present following the synthesis, attempts were made to further purify the plasmid, using the wash/purification steps in the following methods/kits: 1) The Invitrogen™ PureLink™ Genomic DNA kit (ThermoFisher Scientific), 2) DNA precipitation from diluted solutions using Isopropanol as well as Glycogen (ThermoFisher Scientific), 3) QIAquick® Gel Extraction Kit (Qiagen), and 4) QIAamp® DNA Micro Kit (ThermoFisher Scientific). Amplification of the other species (MG, MS, and *M. gallinaceum*) from the plasmid was still unsuccessful. Upon closer inspection of the sequences in the 368 bp insert, it was noted that there were errors in the sequences (Figure 3.6).

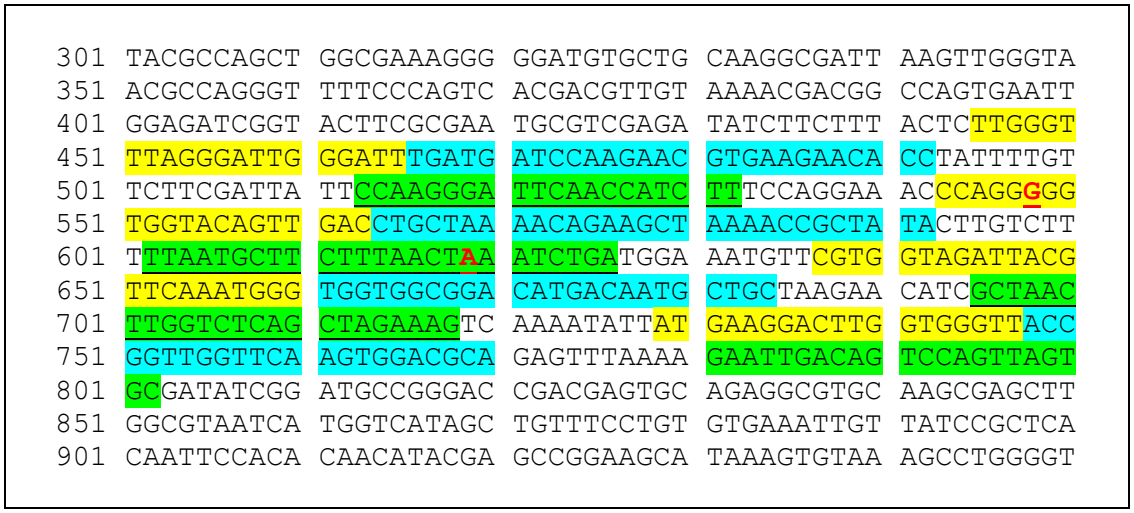

Figure 3.6: Plasmid insert errors for MG, MS, and *M. gallinaceum*.

Both the forward and reverse sequences of the primers for MS had nucleotide errors (highlighted in red) which occurred during the synthesis of the plasmid and the reverse sequences (underlined) for all three species were not correctly reverse complemented.

The errors were only noted on the sequences for MG, MS, and *M. gallinaceum* but no errors on the sequences for *M. pullorum*. This explains why amplification was achieved only for *M. pullorum*, as well as why only *M. pullorum* results were expressed as plasmid copies/µl. No amplification was observed in any of the samples tested using the group-specific primers and probe.

3.3.2 Assay efficiency, measurement response linearity and LOD

DNA amplification in each standard curve showed linearity for results in the ranges that were tested (Figure 3.7). The accepted correlation coefficient (R^2) of each target was ≥ 0.98, the slope of the regression line was between -3.9 and -2.9, and the efficiency ranged between 80% and 120% (Broeders et al., 2014). The results for assay efficiency, linearity, coefficient of variation (R^2), LOD, and C_t value cut-offs are listed in Table 3.5.

Table 3.5: Real-time PCR validation and optimisation results

Species	Linearity	(R^2)	Efficiency	100% LOD	*C_t cut-off
MG	$10^{3.83}$ - $10^{9.83}$	0.9945	103%	10^3 CFUs/µl	30
MS	$10^{3.82}$ - $10^{9.82}$	0.9808	99%	10^4 CFUs/µl	32
M. gallinaceum	$10^{3.82}$ - $10^{9.82}$	0.98	92%	10^3 CFUs/µl	31
M. pullorum	$10^{1.17}$ - $10^{10.17}$	0.978	98%	10^2 plasmid copies/µl	32

*Samples that yielded a C_t value greater than the cut-off would be deemed negative.

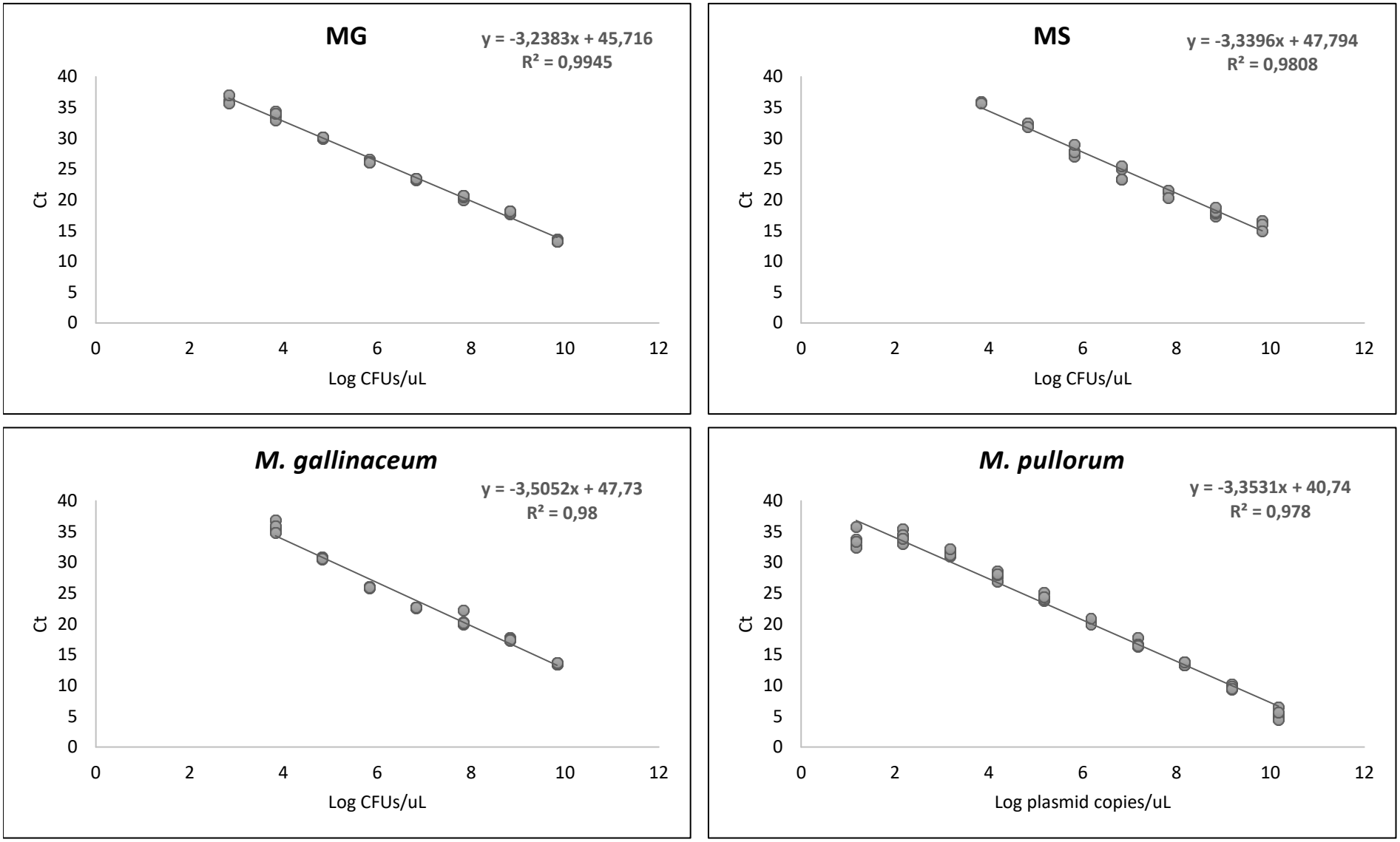

Figure 3.7: Standard curves of each TaqMan® MGB assay for the detection of MG, MS, *M. gallinaceum* and *M. pullorum*.

Assay efficiency is 103%, 99%, 92% and 98% respectively.

3.3.3 *Mycoplasma* species detection

A total of 52 axenic and 33 mixed samples were tested on the multiplex PCR assay. The assay successfully detected the axenic samples (MG [n=8], MS [n=9], *M. gallinaceum* [n=19] and *M. pullorum* [n=4]), as well as the mixed samples (MG [n=14], MS [n=3], *M. gallinaceum* [n=6] and *M. pullorum* [n=13]) accurately (Tables 3.6 and 3.7). A comparison of the real-time PCR results, of the axenic and mixed samples, to the previous results can be seen in Figure 3.8. The reason that there were samples where no amplification was detected may have been due to DNA degradation over time, as the samples were quite old and may have undergone numerous amounts of freezing and thawing (see Tables 3.6 and 3.7).

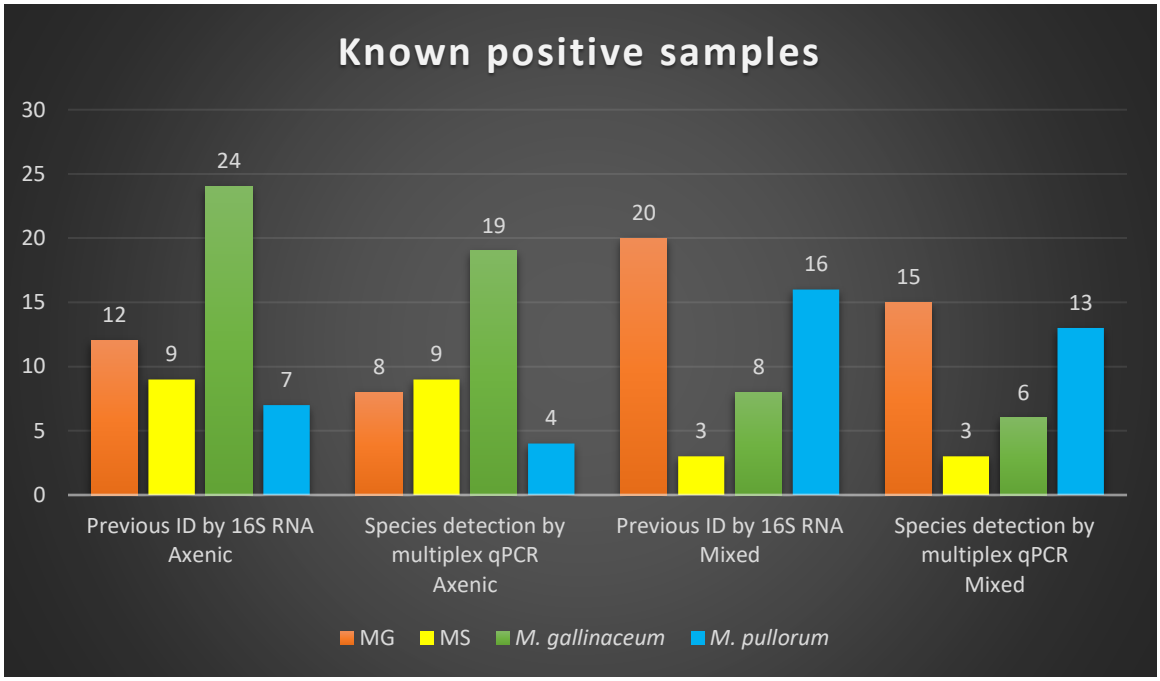

Figure 3.8: Comparison of results observed from testing known positive samples using the multiplex real-time PCR to results previously obtained by 16S RNA identification.

Table 3.6: Multiplex real-time PCR C_t values of axenic samples.

No.	Sample ID	Year of ID	Species ID by 16S RNA	qPCR C_t value
1	B1102-03	2003	MG	Undetermined
2	B726-06	2006	MG	12,90
3	B852-06	2006	MG	10,55
4	B943-06	2006	MG	11,18
5	B642-08	2006	MG	12,81
6	B758/08	2007	MG	5,92
7	B2159-13	2008	MG	Undetermined
8	B1552-14-19	2008	MG	Undetermined
9	B2771-14-1A	2013	MG	8,47
10	B2771-14-1B	2014	MG	Undetermined
11	B878-14-L3	2014	MG	6,70
12	B457-15-5	2014	MG	11,91
13	B2214-07	2007	MS	14,94
14	B1064-14-H3	2014	MS	26,06
15	B1064-14-H5	2014	MS	13,55
16	B1393-14-10	2014	MS	11,93
17	B1394-14-2	2014	MS	10,23
18	B458-15-1	2014	MS	14,58
19	B458-15-5M	2014	MS	13,54
20	B458-15-6	2015	MS	25,31
21	B458-15-11	2015	MS	13,13
22	B878-14-M1	2014	*M. gallinaceum*	14,22
23	B878-14-M4	2014	*M. gallinaceum*	8,64
24	B878-14-M5	2014	*M. gallinaceum*	9,56
25	B1101-14-7	2014	*M. gallinaceum*	14,14
26	B1173-14-2a	2014	*M. gallinaceum*	16,59
27	B1173-14-2b	2014	*M. gallinaceum*	Undetermined
28	B1173-14-4a	2014	*M. gallinaceum*	10,96
29	B1173-14-4b	2014	*M. gallinaceum*	11,60
30	B1173-14-5b	2014	*M. gallinaceum*	Undetermined
31	B1173-14-6b	2014	*M. gallinaceum*	15,20
32	B1173-14-7b	2014	*M. gallinaceum*	11,37
33	B1173-14-8b	2014	*M. gallinaceum*	12,86
34	B1342-14-8	2014	*M. gallinaceum*	13,91
35	B1342-14-10	2014	*M. gallinaceum*	13,75
36	B1342-14-13	2014	*M. gallinaceum*	Undetermined
37	B1396-14-8	2014	*M. gallinaceum*	14,23
38	B1396-14-9	2014	*M. gallinaceum*	Undetermined
39	B2096-14-2	2014	*M. gallinaceum*	13,92
40	B2096-14-7	2015	*M. gallinaceum*	14,66

No.	Sample ID	Year of ID	Species ID by 16S RNA	qPCR C_t value
41	B2096-14-8	2015	*M. gallinaceum*	11,48
42	B3381-15-1	2015	*M. gallinaceum*	11,12
43	B3381-15-3	2015	*M. gallinaceum*	Undetermined
44	B3381-15-4	2015	*M. gallinaceum*	12,52
45	B3381-15-5	2015	*M. gallinaceum*	19,98
46	B2096-14-3	2014	*M. pullorum*	13,29
47	B293-15-12	2014	*M. pullorum*	25,85
48	B293-15-13	2014	*M. pullorum*	Undetermined
49	B293-15-17	2015	*M. pullorum*	35,98
50	B359-15-5	2015	*M. pullorum*	14,42
51	B359-15-6	2015	*M. pullorum*	14,44
52	B540-15-2	2015	*M. pullorum*	Undetermined

Table 3.7: Multiplex real-time PCR C_t values of mixed samples.

No.	Sample ID	Year of ID	qPCR C_t value			
			MG	MS	*M. gallinaceum*	*M. pullorum*
1	B2888-13-1A	2013	30,09	N/A	15,15	N/A
2	B1101-14-10	2014	N/A	N/A	36,63	33,16
3	B1342-14-18	2014	N/A	N/A	13,31	N/A
4	B1342-14-4	2014	N/A	N/A	30,86	N/A
5	B1393-14-4	2014	N/A	*16,32	N/A	N/A
6	B1395-14-5	2014	N/A	N/A	14,3	14,67
7	B1396-14-6	2014	Undetermined	N/A	N/A	N/A
8	B2771-14-15A	2014	Undetermined	N/A	N/A	Undetermined
9	B878-14-M2	2014	*11,37	N/A	N/A	N/A
10	B1931-15-6A	2015	*8,72	N/A	N/A	N/A
11	B1932-15-2	2015	*6,73	N/A	N/A	N/A
12	B2063-15-3	2015	*7,6	N/A	N/A	N/A
13	B293-15-14	2015	N/A	N/A	N/A	*15,76
14	B293-15-18	2015	N/A	N/A	13,76	16,48
15	B293-15-7	2015	N/A	*12,69	N/A	N/A
16	B293-15-9	2015	Undetermined	N/A	N/A	N/A
17	B3443-15-1	2015	*8,33	N/A	N/A	N/A
18	B3443-15-2	2015	15,92	N/A	N/A	9,99
19	B3443-15-3	2015	N/A	N/A	N/A	9,58
20	B3443-15-4	2015	10,86	N/A	N/A	9,37
21	B3443-15-5	2015	10,01	N/A	N/A	9,39
22	B3443-15-6	2015	11,35	N/A	N/A	12,01
23	B3443-15-7	2015	N/A	N/A	N/A	*11,92
24	B3443-15-8	2015	13,88	N/A	N/A	12,07
25	B359-15-2	2015	N/A	N/A	33,89	14,44
26	B359-15-3	2015	N/A	N/A	N/A	*35,97
27	B359-15-4	2015	N/A	N/A	N/A	*16,27
28	B457-15-3	2015	*11,82	N/A	N/A	N/A
29	B458-15-10	2015	*14,97	N/A	N/A	N/A
30	B458-15-5	2015	Undetermined	N/A	N/A	N/A
31	B464-15-3	2015	Undetermined	13,16	N/A	N/A
32	B540-15-4	2015	11,32	N/A	N/A	30,83
33	B540-15-5	2015	24,74	N/A	29,29	N/A

*Samples mixed only with *M. gallinarum*, and/or *M.iners*. N/A = Not applicable

3.3.4 Field sample testing

From culture, only nine *Mycoplasma* species were isolated, and tested as described in 3.2.6.2. These were identified as MG (n=1), *M. gallinaceum* (n=1), *M. pullorum* (n=2), *M. gallinarum* (n=1), *M. glycophilum* (n=1) and *M. iners* (n=3). From direct cPCR testing, amplification was observed only in eleven samples. All eleven suspect samples could not be identified to species level, because of overlapping peaks in the DNA sequences that were likely caused by the presence of more than one sequence/ *Mycoplasma* species. Culture and direct cPCR results can be seen in Figure 3.9.

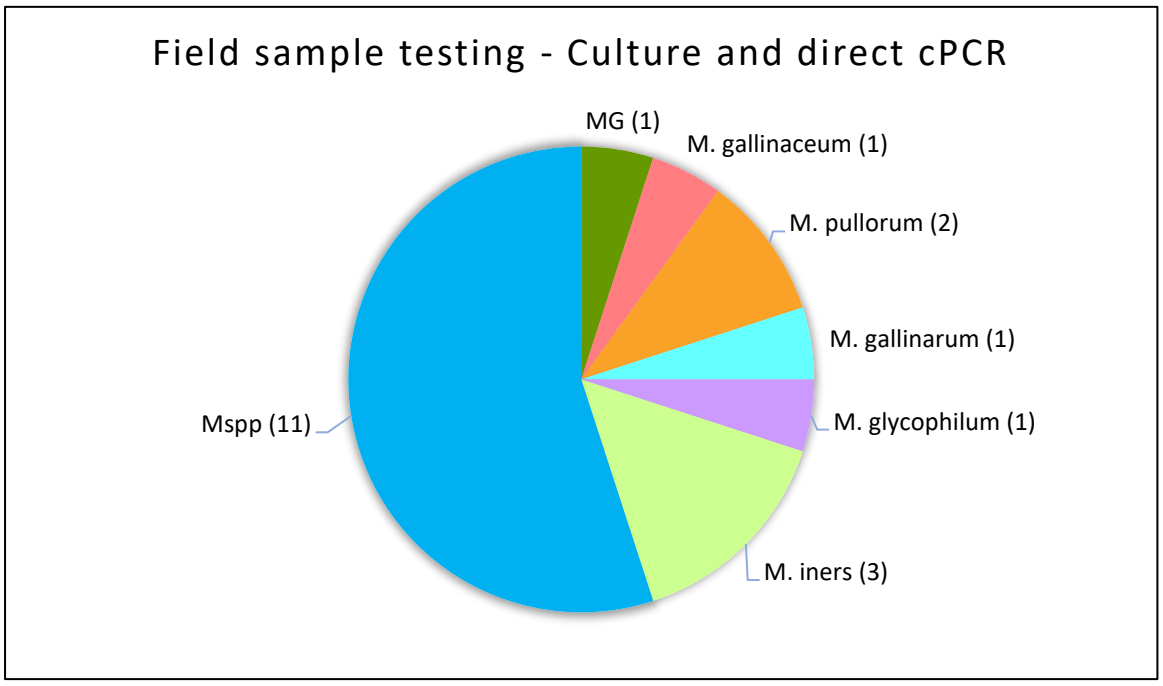

Figure 3.9: *Mycoplasma* species detected from culture and direct cPCR testing.

For the multiplex real-time PCR, the field samples were tested in batches with MG and MS in one mix, and *M. gallinaceum* and *M. pullorum* in another. Amplification was successfully achieved, and visualisation of the curves (amplification plots) can be seen in Figure 3.10. From real-time PCR testing, MG (n = 125), MS (n = 169), *M. gallinaceum* (n = 31) and *M. pullorum* (n = 65) were detected, and the C_t values are listed in Table 3.8 and are summarised in Figure 3.11. All the samples had amplification for one or more *Mycoplasma* species; this corroborates the results observed from direct cPCR testing.

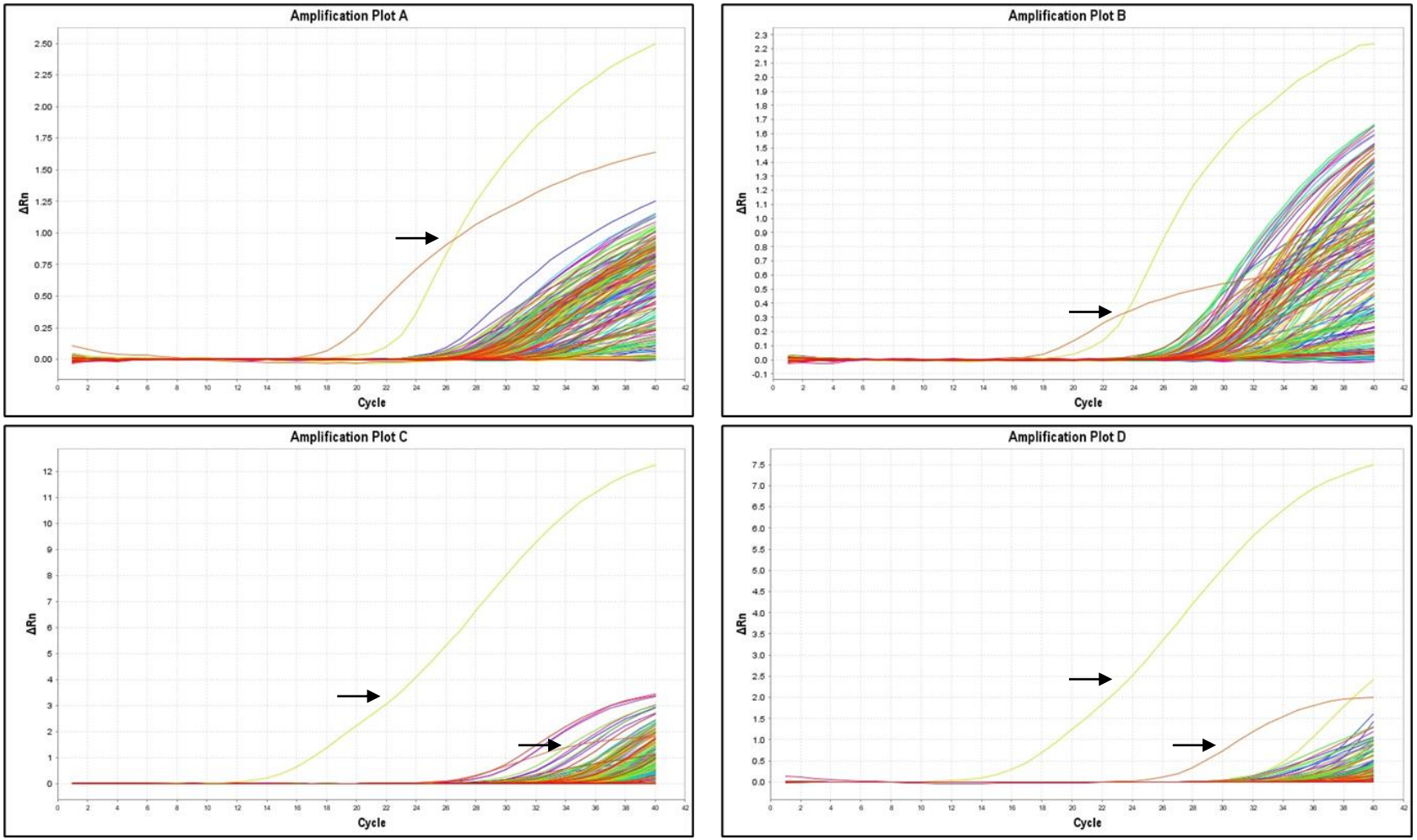

Figure 3.10: Amplification plots of field samples tested on the multiplex real-time PCR.

The figures show the two sets of amplification plots of the field samples tested. Amplification plots A and B are MG and MS, and plots C and D are *M. gallinaceum* and *M. pullorum*. The arrows in all four plots indicate the positive controls.

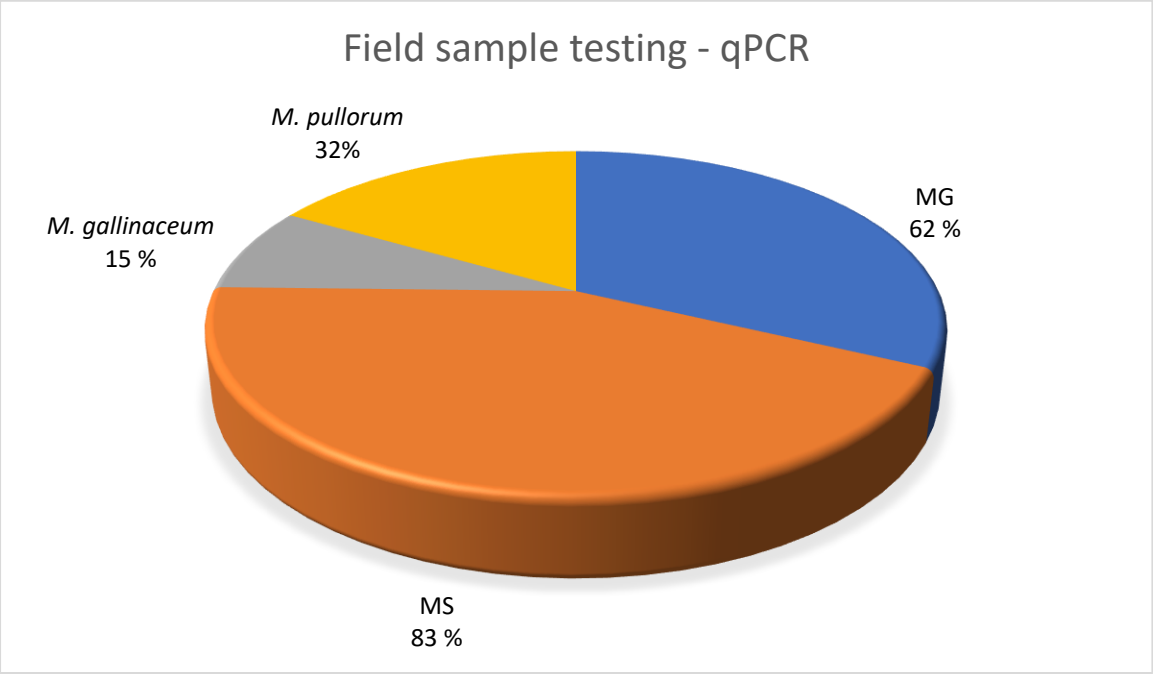

Figure 3.11: *Mycoplasma* species detected from multiplex qPCR testing.

Out of the 203 samples tested, MG was observed in 62 % of the samples, MS in 83 %, *M. gallinaceum* in 15 % and *M. pullorum* in 32 %.

Table 3.8: Multiplex real-time PCR C_t values of field samples.

No.	Sample ID	MG	MS	*M. gallinaceum*	*M. pullorum*
1	FR27-01-22	34.31	32.51	36.81	28.96
2	FR27-02-22	26.89	29.25	Undetermined	36.75
3	FR27-03-22	28.79	33.80	Undetermined	29.88
4	FR27-04-22	28.50	31.83	Undetermined	33.79
5	FR27-05-22	35.18	34.70	Undetermined	32.37
6	FR27-06-22	27.26	33.80	33.92	30.98
7	FR27-07-22	28.14	30.09	Undetermined	33.80
8	FR27-08-22	29.44	33.11	34.37	31.36
9	FR27-09-22	26.79	30.82	Undetermined	33.47
10	FR27-10-22	26.83	27.90	Undetermined	32.34
11	FR28-01-22	26.04	27.65	34.16	32.97
12	FR28-02-22	28.28	33.28	35.34	31.15
13	FR28-03-22	26.95	32.77	34.88	30.85
14	FR28-04-22	25.19	27.03	34.57	32.99
15	FR28-05-22	25.36	29.75	Undetermined	36.55
16	FR28-06-22	27.18	28.72	Undetermined	34.47
17	FR28-07-22	27.41	31.69	33.76	32.19
18	FR28-08-22	25.60	32.07	Undetermined	Undetermined
19	FR28-09-22	31.55	28.51	Undetermined	34.60
20	FR28-10-22	26.27	31.48	Undetermined	36.92
21	FR29-01-22	30.82	25.70	35.00	33.85
22	FR29-02-22	25.74	27.50	33.64	30.02
23	FR29-03-22	26.66	27.24	34.75	28.65
24	FR29-04-22	28.09	25.17	30.19	25.18
25	FR29-05-22	26.15	25.28	27.33	24.74
26	FR29-06-22	24.25	26.18	36.64	31.38
27	FR29-07-22	23.93	24.64	29.24	26.79
28	FR29-08-22	30.91	27.89	32.58	32.61
29	FR29-09-22	29.99	27.57	Undetermined	31.24
30	FR29-10-22	28.86	31.18	34.79	31.37
31	V1-01-22	28.62	25.75	33.12	28.61
32	V1-02-22	27.49	31.28	Undetermined	30.62
33	V1-03-22	26.72	28.77	30.88	30.81
34	V1-04-22	27.61	28.63	32.24	30.76
35	V1-05-22	26.33	26.04	28.31	29.32
36	V1-06-22	27.13	27.51	Undetermined	32.88
37	V1-07-22	33.20	29.34	Undetermined	29.97
38	V1-08-22	29.88	32.53	Undetermined	30.63
39	V1-09-22	23.83	26.26	32.40	30.19
40	V1-10-22	Undetermined	29.53	35.08	27.53
41	V2-01-22	26.35	29.67	Undetermined	37.14

No.	Sample ID	MG	MS	*M. gallinaceum*	*M. pullorum*
42	V2-02-22	26.64	28.78	34.5	30.94
43	V2-03-22	28.28	26.80	38.62	28.38
44	V2-04-22	32.96	34.24	34.32	34.35
45	V2-05-22	28.80	30.84	34.94	30.67
46	V2-06-22	27.81	28.38	Undetermined	36.28
47	V2-07-22	24.99	25.10	28.76	25.42
48	V2-08-22	26.08	26.15	36.67	33.90
49	V2-09-22	29.11	28.13	36.31	32.18
50	V2-10-22	26.81	27.80	33.82	32.42
51	V3-01-22	31.48	29.37	34.56	33.70
52	V3-02-22	30.62	29.24	34.37	32.83
53	V3-03-22	32.9	33.00	35.45	32.62
54	V3-04-22	Undetermined	28.43	35.90	36.57
55	V3-05-22	30.26	31.21	37.30	30.55
56	V3-06-22	Undetermined	26.82	34.81	30.23
57	V3-07-22	30.37	30.07	Undetermined	32.25
58	V3-08-22	33.70	33.48	Undetermined	32.38
59	V3-09-22	25.02	26.74	34.53	30.99
60	V3-10-22	Undetermined	26.83	Undetermined	32.12
61	V4-01-22	Undetermined	28.62	35.27	31.47
62	V4-02-22	26.24	27.86	35.14	30.89
63	V4-03-22	27.85	27.41	Undetermined	32.42
64	V4-04-22	28.25	29.69	Undetermined	Undetermined
65	V4-05-22	26.38	27.49	Undetermined	33.98
66	V4-06-22	Undetermined	29.01	34.56	31.97
67	V4-07-22	33.57	31.57	Undetermined	31.15
68	V4-08-22	26.99	26.67	27.90	31.65
69	V4-09-22	33.18	30.46	Undetermined	33.00
70	V4-10-22	24.23	23.96	37.03	34.21
71	V6-01-22	32.84	29.11	Undetermined	33.26
72	V6-02-22	32.92	29.68	34.95	31.71
73	V6-03-22	Undetermined	27.69	Undetermined	Undetermined
74	V6-04-22	36.29	31.17	Undetermined	33.36
75	V6-05-22	33.34	32.32	Undetermined	33.67
76	V6-06-22	Undetermined	28.19	35.82	33.94
77	V6-07-22	Undetermined	30.09	34.19	34.16
78	V6-08-22	Undetermined	31.02	Undetermined	33.08
79	V6-09-22	30.86	29.55	34.86	35.52
80	V6-10-22	33.25	29.56	34.51	32.43
81	V7-01-22	29.78	28.34	35.99	31.63
82	V7-02-22	Undetermined	35.80	30.97	30.65
83	V7-03-22	Undetermined	27.38	33.77	32.91

No.	Sample ID	MG	MS	*M. gallinaceum*	*M. pullorum*
84	V7-04-22	30.28	29.33	35.14	31.75
85	V7-05-22	32.29	27.80	33.61	31.16
86	V7-06-22	Undetermined	25.70	34.02	27.53
87	V7-07-22	29.23	31.31	Undetermined	33.61
88	V7-08-22	27.59	27.12	34.92	30.42
89	V7-09-22	28.10	28.72	Undetermined	36.12
90	V7-10-22	24.64	29.65	Undetermined	Undetermined
91	V10-01-22	27.71	32.23	33.50	31.27
92	V10-02-22	31.78	32.52	35.32	31.93
93	V10-03-22	30.29	33.10	Undetermined	33.22
94	V10-04-22	32.49	34.72	31.29	31.44
95	V10-05-22	32.66	35.21	34.63	36.02
96	V10-06-22	32.62	38.68	38.24	36.84
97	V10-07-22	31.22	26.08	36.35	37.26
98	V10-08-22	29.63	28.86	35.43	35.61
99	V10-09-22	31.20	27.36	36.25	37.64
100	V10-10-22	29.75	28.21	35.82	38.38
101	VL 01-22	34.50	31.55	Undetermined	Undetermined
102	VL 02-22	31.51	29.43	Undetermined	35.96
103	VL 03-22	30.53	28.21	32.53	33.61
104	VL 04-22	31.68	26.74	Undetermined	33.52
105	VL 05-22	33.23	28.28	Undetermined	Undetermined
106	VL 06-22	31.80	29.74	Undetermined	Undetermined
107	VL 07-22	32.99	29.96	Undetermined	37.93
108	VL 08-22	32.11	31.56	37.7	Undetermined
109	VL 09-22	Undetermined	32.46	Undetermined	Undetermined
110	VL 10-22	33.08	28.43	Undetermined	Undetermined
111	VL 11-22	34.60	Undetermined	35.85	35.30
112	VL 12-22	Undetermined	26.36	Undetermined	38.44
113	VL 13-22	34.19	29.09	Undetermined	Undetermined
114	VL 14-22	32.82	27.78	35.96	37.38
115	VL 15-22	33.44	31.56	Undetermined	Undetermined
116	VL 16-22	34.83	30.37	37.07	36.81
117	VL 17-22	Undetermined	25.86	Undetermined	Undetermined
118	VL 18-22	33.05	28.99	36.18	37.11
119	VL 19-22	30.27	29.02	35.59	Undetermined
120	VL 20-22	32.54	28.35	Undetermined	35.35
121	FR5-01-22	26.87	30.23	28.88	32.98
122	FR5-02-22	30.76	32.62	32.17	35.54
123	FR5-03-22	31.89	28.88	29.67	32.23
124	FR5-04-22	33.24	Undetermined	34.16	33.62
125	FR5-05-22	25.69	26.09	33.76	38.31

No.	Sample ID	MG	MS	*M. gallinaceum*	*M. pullorum*
126	FR5-06-22	27.59	27.66	31.61	35.47
127	FR5-07-22	27.96	27.08	30.37	33.51
128	FR5-08-22	26.73	29.69	32.84	32.85
129	FR5-09-22	26.60	28.85	34.14	Undetermined
130	FR5-10-22	29.53	33.12	30.83	34.06
131	FR9A-01-22	26.39	31.31	32.99	34.25
132	FR9A-02-22	31.53	28.81	33.61	33.84
133	FR9A-03-22	32.03	28.98	39.09	Undetermined
134	FR9A-04-22	33.51	Undetermined	Undetermined	33.49
135	FR9A-05-22	Undetermined	Undetermined	30.44	Undetermined
136	FR9A-06-22	28.82	26.08	Undetermined	36.49
137	FR9A-07-22	32.44	27.86	35.33	34.06
138	FR9A-08-22	29.83	27.74	Undetermined	38.42
139	FR9A-09-22	31.64	31.72	Undetermined	Undetermined
140	FR9A-10-22	31.74	Undetermined	39.11	35.5
141	FR10A-01-22	30.64	26.00	Undetermined	36.39
142	FR10A-02-22	25.41	30.95	Undetermined	Undetermined
143	FR10A-03-22	27.88	31.41	Undetermined	38.25
144	FR10A-04-22	24.26	28.25	33.73	38.25
145	FR10A-05-22	26.32	27.36	Undetermined	36.95
146	FR10A-06-22	26.70	30.22	Undetermined	37.90
147	FR10A-07-22	30.46	29.28	Undetermined	Undetermined
148	FR10A-08-22	Undetermined	28.85	Undetermined	Undetermined
149	FR10A-09-22	28.03	32.65	Undetermined	Undetermined
150	FR10A-10-22	28.79	32.11	Undetermined	34.81
151	FR16-01-22	29.24	27.72	33.17	35.40
152	FR16-02-22	25.37	27.77	35.52	36.42
153	FR16-03-22	23.36	26.91	37.81	Undetermined
154	FR16-04-22	31.06	29.37	37.50	36.00
155	FR16-05-22	25.81	28.70	34.79	Undetermined
156	FR16-06-22	27.28	29.08	Undetermined	38.35
157	FR16-07-22	27.70	30.00	34.66	37.13
158	FR16-08-22	Undetermined	30.46	Undetermined	37.77
159	FR16-09-22	26.30	29.49	32.55	35.27
160	FR16-10-22	32.37	31.14	36.84	Undetermined
161	FR22-01-22	27.46	31.37	32.21	Undetermined
162	FR22-02-22	25.15	30.00	32.32	Undetermined
163	FR22-03-22	25.07	27.21	29.15	Undetermined
164	FR22-04-22	27.51	25.48	32.59	35.75
165	FR22-05-22	24.83	30.26	29.67	35.49
166	FR22-06-22	23.70	29.50	28.57	Undetermined
167	FR22-07-22	24.40	28.83	33.77	Undetermined

No.	Sample ID	MG	MS	*M. gallinaceum*	*M. pullorum*
168	FR22-08-22	28.41	33.10	Undetermined	Undetermined
169	FR22-09-22	32.23	29.90	30.61	34.85
170	FR22-10-22	24.62	27.57	28.84	33.24
171	FR45-01-22	23.75	25.37	28.41	33.13
172	FR45-02-22	25.85	25.52	31.59	36.62
173	FR45-03-22	25.00	28.25	32.37	36.72
174	FR45-04-22	25.01	26.12	30.19	34.38
175	FR45-05-22	Undetermined	25.68	34.11	35.11
176	FR45-06-22	27.26	29.26	35.94	35.44
177	FR45-07-22	25.43	26.34	30.71	34.35
178	FR45-08-22	26.79	28.76	33.3	35.18
179	FR45-09-22	28.30	27.90	31.85	34.87
180	FR45-10-22	27.13	30.78	28.57	35.95
181	FR47-01-22	26.75	25.64	35.68	34.65
182	FR47-02-22	24.65	24.79	33.40	30.19
183	FR47-03-22	22.33	25.82	28.69	32.59
184	FR47-04-22	32.93	28.74	34.23	32.74
185	FR47-05-22	24.96	26.35	37.08	34.26
186	FR47-06-22	24.33	25.43	33.95	33.03
187	FR47-07-22	26.20	28.47	28.53	33.58
188	FR47-08-22	26.35	27.30	28.72	28.99
189	FR47-09-22	25.41	26.19	31.47	31.08
190	FR47-10-22	30.11	27.78	35.13	33.46
191	NFS 1	34.67	35.54	33.36	30.68
192	NFS 2	27.98	32.74	33.97	29.40
193	NFS 3	Undetermined	36.47	33.30	31.53
194	NFS 4	27.81	33.12	28.55	30.12
195	NFS 5	29.51	33.28	33.61	31.16
196	NFS 6	22.58	27.40	33.21	31.13
197	NFS 7	31.72	34.32	32.72	31.42
198	NFS 8	Undetermined	34.44	33.11	31.21
199	NFS 9	Undetermined	Undetermined	33.68	31.49
200	NFS 10	35.99	Undetermined	33.80	31.57
201	NFS 11	Undetermined	34.72	33.45	31.97
202	NFS 12	35.10	36.99	33.54	31.42
203	NFS 13	Undetermined	Undetermined	32.93	31.41

The four *Mycoplasma* species detected, were observed to appear in 10 different combinations. Table 3.9 shows a summary of the combinations observed, and Figure 3.12 shows a summary of all the samples tested.

M. gallisepticum and MS were the most frequent combinations observed (65 samples). *M. pullorum* was detected in 24 samples combined with MG and MS; and in 12 samples combined only with MS. *M. gallinaceum* was detected in a higher number of samples in combination with MG and MS (14 samples) but was the least frequent species observed. The least frequent combinations observed were MG, *M. gallinaceum* & *M. pullorum*, and MG & *M. gallinaceum* (1 sample each). Although MG appeared in 6 of the 9 different combinations observed, only 3 samples had single MG detections. MS appeared in 5 of the 9 different combinations and was the most frequent species observed in single detections (42 samples). Ten samples were observed where all four *Mycoplasma* species (MG, MS, *M. gallinaceum* and *M. pullorum*) were not detected, and 10 samples were observed where all four *Mycoplasma* species were detected.

Table 3.9: Summary of the different *Mycoplasma* species combinations

No.	*Species combination	Number of samples	Proportion (%)
1	MG, MS, Mgal & Mpul	10	4.92 %
2	MG & MS	65	32.01 %
3	MG, MS & Mgal	14	6.89 %
4	MG, MS & Mpul	24	11.82 %
5	MG, Mgal & Mpul	1	0.49 %
6	MG & Mgal	1	0.49 %
7	MG & Mpul	7	3.44 %
8	MS & Mgal	2	0.98 %
9	MS & Mpul	12	5.91 %
10	Mgal & Mpul	2	0.98 %

*MG - *M. gallisepticum*
*MS = *M. synoviae*
*Mgal = *M. gallinaceum*
*Mpul = *M. pullorum*

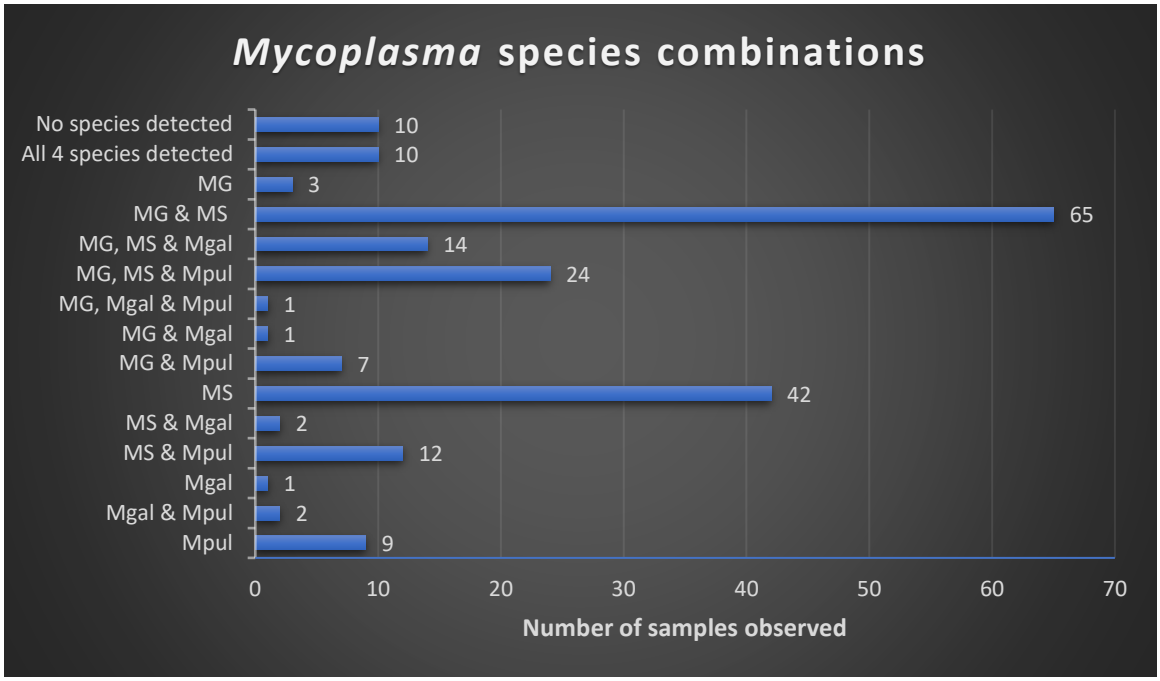

Figure 3.12: Different combinations of the four *Mycoplasma* species detected from field samples (n=203) after multiplex qPCR testing.

3.4 DISCUSSION

Mycoplasmas require complex protein-rich growth media with supplements for successful cultivation, and often the non-pathogenic *Mycoplasma* species will usually overgrow the slow-growing pathogenic species (Nascimento et al., 2005). As discussed in Chapter 2, culture was a very big challenge throughout the study given the nature of the samples that were available for use. To conduct real-time PCR validation (diagnostic) with more accuracy, positive controls that are of known concentration are required. The ATCC/NCTC reference material in the UP repository was too old to culture successfully, and due to global logistical challenges, could not be imported at the time of this study. This led to the decision of synthesising a laboratory-positive control for use in the validation.

Using an adapted protocol for the direct purification of genomic DNA from mycoplasma colonies I obtained sufficient material for the amplification of MG, MS, and *M. gallinaceum*. However, since the DNA extracted from these references was of unknown concentration, an estimated value was assigned. This meant that the entire diagnostic validation would not be carried out as initially intended but only the diagnostic efficiency of the assay would be determined. To determine assay efficiency, serial dilutions were made but amplification was not accurately achieved and the group-specific primers and probes were undetermined (no amplification), even after newly synthesized oligonucleotides were tested. There could likely be a flaw in the design of the probe, as the respective primers were previously tested and showed successful amplification.

For four *Mycoplasma* species (MG, MS, *M. gallinaceum* and *M. pullorum*), assay efficiency was successfully determined, while the LOD was only conservatively estimated. To determine a more precise estimation, two-fold serial dilutions at intervals between 0% and 100% detection are usually done with the endpoint often set at 95% and the LOD estimated using logistic regression or probit analysis (Health, 2014b). This was not done in the study because the starting concentrations of the references were assigned estimates.

In testing the known positive samples, challenges that were met included depletion of the existing DNA material, DNA had become too degraded to test (determined by

agarose gel visualisation), and samples missing from storage (this occurs when freezers break down and the contents are moved to temporary storage freezers). Initially, there were 68 samples available for use in the study (Table 2.3), but only 52 samples were tested due to the reasons stated. The same was also true for the mixed samples, out of the 44 samples available for use (Table 2.3), only 33 samples were tested. When testing the mixed samples, only the MG, MS, *M. gallinaceum*, and *M. pullorum* results were analysed, as the multiplex could only detect those four species. The group-specific probe was included in the study to act as an internal 'catch-all' control, for the general detection of *Mycoplasma* species. This meant that all four *Mycoplasma* species (MG, MS, *M. gallinaceum* and *M. pullorum*) would also be detected; thus, samples where all four *Mycoplasma* species would not be detected, but show amplification on the group-specific probe, would be tested further to determine the presence of *M. gallinarum*, *M. iners* and possibly other poultry *Mycoplasma* species.

Testing of the field samples involved three different methods, firstly to culture the samples to determine which *Mycoplasma* species grew, then test the samples directly on conventional PCR, then finally testing the samples on the multiplex real-time PCR to determine if the assay would be able to detect field isolates; and then finally corroborate all the results obtained when compared. The different results obtained did not match in all three different testing methods. In culture, only the fast-growing *Mycoplasma* species were observed (apart from one MG isolate). As part of the fast-growing *Mycoplasma* species, *Mycoplasma glycophilum* was isolated. This species is not commonly found in South Africa, and has only been reported once before in a study conducted in the Free State province (Moretti, 2012). The samples used by Moretti, (2012) were collected from various poultry farms across the country. Other reports of *M. glycophilum* are from studies that were conducted in Europe (Benčina et al., 1987, Bradbury et al., 2001, Forrest and Bradbury, 1984, Loria et al., 2008, Ongor et al., 2008). In the direct cPCR testing, the 11 suspect samples could not be identified to species level, this was likely due to the samples containing more than one mycoplasma species. In the multiplex real-time PCR testing, the assay detected all four *Mycoplasma* species, and in most cases, they were mixed (Table 3.8). Out of the 203 samples that were tested, MG was observed in 62 % of the samples, MS in 83 %, *M. gallinaceum* in 15 % and *M. pullorum* in 32 %. This indicates that there are

proportionally more MG and MS species circulating in poultry populations. *Mycoplasma* species coinfections (combinations) were also observed, and MG and MS coinfections were the most frequent (32.01 % of samples tested). The least frequent coinfections observed were combinations of MG, *M. gallinaceum* & *M. pullorum* and MG & *M. gallinaceum* (0.49 % of samples tested). *Mycoplasma pullorum* was observed to appear in high numbers both as single infections and coinfections, whereas *M. gallinaceum* only appeared in high numbers in coinfections.

In conclusion, the multiplex real-time PCR was successfully developed and optimised. There was no evidence of cross-reaction between the selected primer mixes when multiplexing. The assay can detect the four *Mycoplasma* species accurately and simultaneously. With good-quality references and a known starting concentration, a more precise validation can be done on the assay in future to increase accuracy and meet statistical validation guidelines (Health, 2014b).

CHAPTER 4

4.1 GENERAL CONCLUSIONS

Bioinformatics involves working with large amounts of data by not only giving researchers access to existing information but also allowing them to input their findings as they receive them. The small size and complexity of mycoplasma genomes grants researchers the ability to successfully use *Mycoplasma* species in bioinformatics, but this does not come without challenges. Currently, the biggest challenge is that the results obtained (i.e., from conducting a pan genome analysis) are largely dependent on data that already exists. In this study a large portion of genes that were obtained from the pan genome analysis were 'hypothetical', meaning that data for those genes is not yet available. This was also noted for genes that were correctly identified from the pan genome analysis but could not be found in the whole genomes of the *Mycoplasma* species available in the NCBI database. The only way to mitigate this challenge is for researchers to continue to test and submit more results on to the Prokka and RAST servers. Another shortfall of using bioinformatics was that, even though the size of *Mycoplasma* species makes them better suited for use, the data (list of genes) obtained still ranged in the thousands. To successfully design novel oligonucleotides that stringently match the regions of interest, the large data sets had to be manually analysed, and the process was very time consuming and monotonous.

The biggest limitation of the study was the cultivation of previously stored references and field samples. The isolates took exceptionally long to grow on agar media, and when they did grow would immediately die out when transferred to broth media. The bacteria could have been transferred into new media too soon before the colonies could fully grow onto the surface of the agar. The volume of broth used (10 ml) could also have been too much for initial transfer of a single colony from agar. In studies where large quantities of bacteria are not necessarily required, extracting DNA directly from the colony on agar medium, as per the technique used in this study, would be very effective and time-saving.

Oligonucleotide primers and probes for the PCR detection and differentiation of four Mycoplasma species viz., *Mycoplasma gallisepticum*, *Mycoplasma synoviae*, *Mycoplasma gallinaceum*, *Mycoplasma pullorum* were successfully designed, tested and PCR conditions were optimised. A multiplex real-time PCR assay using these

oligonucleotides was developed and optimised. The assay can accurately and simultaneously detect and differentiate between all four *Mycoplasma* species and is the first to be validated for the real-time detection of *M. gallinaceum* and *M. pullorum*. After a more precise statistical validation, the assay could potentially be used commercially as a diagnostic test. The complete genome of *M. glycophilum* is currently available (https://www.ncbi.nlm.nih.gov/nuccore/NZ_LR215024.1), once the complete genomes for *M. gallinarum* and *M. iners* are available, oligonucleotides could be similarly designed and tested for these species too.

Field samples were tested using the assay I developed, and out of the 203 samples collected, MG was detected in 61 %, MS in 83 %, *M. gallinaceum* in 15 % and *M. pullorum* in 32 % of cases. Coinfections were observed in 68 % of the samples, where about 67 % were coinfections with either MG, MS, or both and about 1% were coinfections without MG and MS. These results give an indication that although there are proportionately more MG and MS species circulating in poultry populations, the non-pathogenic *Mycoplasma* species are widely present and larger numbers appear mostly in coinfections with either MG, MS, or both. From culture and identification, 6 *Mycoplasma* species viz., MG, *M. gallinaceum*, *M. pullorum*, *M. gallinarum*, *M. glycophilum*, and *M. iners* were isolated; all of which are fast-growing *Mycoplasma* species, excluding MG. *Mycoplasma glycophilum* was detected for only the second time in South Africa, with the initial report made in 2012.

Prospective research arising from this study includes:

- Conducting a more precise statistical validation on the assay.
- Conducting a pan-genome analysis to compare the genes found in field strains to the genes in vaccine strains, as well as genes of species known to carry antibiotic resistance.
- Possible synthesis of more oligonucleotide primers and probes for the specific detection of *M. gallinarum* and *M. iners*, and possibly other poultry *Mycoplasma* species, that can be used in the optimised real-time PCR assay.
- Investigation of the rate at which non-pathogenic *Mycoplasmas* species appear in coinfections.

- Investigation of the rate at which non-pathogenic *Mycoplasmas* species are possibly spreading in poultry populations
- Investigation of the frequency of *M. glycophilum* in South African poultry flocks.

References

ADEYEMI, M., BWALA, D. G. & ABOLNIK, C. 2018. Comparative Evaluation of the Pathogenicity of *Mycoplasma gallinaceum* in Chickens. *Avian diseases,* 62**,** 50-56.

AMER, M. M., MEKKY, H. M. & FEDAWY, H. S. 2019. Molecular identification of *Mycoplasma synoviae* from breeder chicken flock showing arthritis in Egypt. *Veterinary World,* 12**,** 535-541.

AMER MOHAMED, M. M., MEKKY, H. M. & FEDAWY, H. S. 2019. Molecular identification of *Mycoplasma synoviae* from breeder chicken flock showing arthritis in Egypt. *Veterinary World,* 12**,** 535-541.

ARMOUR, N. K., LAIBINIS, V. A., COLLETT, S. R. & FERGUSON-NOEL, N. 2013. The development and application of a *Mycoplasma gallisepticum* sequence database. *Avian Pathology,* 42**,** 408-415.

ASHELFORD, K. E. 2002. PRIMROSE: a computer program for generating and estimating the phylogenetic range of 16S rRNA oligonucleotide probes and primers in conjunction with the RDP-II database. *Nucleic Acids Research,* 30**,** 3481-3489.

AZIZ, R. K., BARTELS, D., BEST, A. A., DEJONGH, M., DISZ, T., EDWARDS, R. A., FORMSMA, K., GERDES, S., GLASS, E. M., KUBAL, M., MEYER, F., OLSEN, G. J., OLSON, R., OSTERMAN, A. L., OVERBEEK, R. A., MCNEIL, L. K., PAARMANN, D., PACZIAN, T., PARRELLO, B., PUSCH, G. D., REICH, C., STEVENS, R., VASSIEVA, O., VONSTEIN, V., WILKE, A. & ZAGNITKO, O. 2008. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics,* 9**,** 75.

BAXEVANIS, A. D., BADER, G. D. & WISHART, D. S. 2020. *Bioinformatics*, John Wiley & Sons.

BEHBOUDI, S. 2022. *Avian mycoplasmosis (Mycoplasma gallisepticum)*, CABI International.

BENČINA, D., DORRER, D. & TADINA, T. 1987. Mycoplasma species isolated from six avian species. *Avian Pathology,* 16**,** 653-664.

BEYLEFELD, A. 2018. Genomic comparison of *Mycoplasma* species isolated from commercial chickens in South Africa. *University of Pretoria*.

BEYLEFELD, A., WAMBULAWAYE, P., BWALA, D. G., GOUWS, J. J., LUKHELE, O. M., WANDRAG, D. B. R. & ABOLNIK, C. 2018. Evidence for multidrug resistance in nonpathogenic Mycoplasma species isolated from South African poultry. *Applied and Environmental Microbiology,* 84**,** e01660-18.

BOYLE, J. S., GOOD, R. T. & MORROW, C. J. 1995. Detection of the turkey pathogens Mycoplasma meleagridis and M. iowae by amplification of genes coding for rRNA. *Journal of Clinical Microbiology,* 33**,** 1335-1338.

BRADBURY, J. 2005. Poultry mycoplasmas: sophisticated pathogens in simple guise. *British poultry science,* 46**,** 125-136.

BRADBURY, J. M., ABDUL-WAHAB, O. M. S., YAVARI, C. A., DUPIELLET, J.-P. & BOVÉ, J. M. 1993. *Mycoplasma imitans* sp. nov. is related to *Mycoplasma gallisepticum* and found in birds. *International Journal of Systematic and Evolutionary Microbiology,* 43**,** 721-728.

BRADBURY, J. M., YAVARI, C. A. & DARE, C. M. 2001. Mycoplasmas and respiratory disease in pheasants and partridges. *Avian Pathology,* 30**,** 391-396.

BROEDERS, S., HUBER, I., GROHMANN, L., BERBEN, G., TAVERNIERS, I., MAZZARA, M., ROOSENS, N. & MORISSET, D. 2014. Guidelines for validation of qualitative real-time PCR methods. *Trends in Food Science & Technology,* 37**,** 115-126.

BUIM, M. R., METTIFOGO, E., TIMENETSKY, J., KLEVEN, S. & FERREIRA, A. J. P. 2009. Epidemiological survey on *Mycoplasma gallisepticum* and *M. synoviae* by multiplex PCR in commercial poultry. *Pesquisa Veterinária Brasileira,* 29**,** 552-556.

CARLI, K. T. & EYIGOR, A. 2003. Real-time polymerase chain reaction for detection of *Mycoplasma gallisepticum* in chicken trachea. *Avian diseases,* 47**,** 712-717.

CHAI, S. J., COLE, D., NISLER, A. & MAHON, B. E. 2017. Poultry: the most common food in outbreaks with known pathogens, United States, 1998–2012. *Epidemiology and Infection,* 145**,** 316-325.

CHAIDEZ-IBARRA, M. A., VELAZQUEZ, D. Z., ENRIQUEZ-VERDUGO, I., CASTRO DEL CAMPO, N., RODRIGUEZ-GAXIOLA, M. A., MONTERO-PARDO, A., DIAZ, D. & GAXIOLA, S. M. 2022. Pooled molecular occurrence of *Mycoplasma gallisepticum* and *Mycoplasma synoviae* in poultry: A systematic review and meta-analysis. *Transboundary and emerging diseases,* 69**,** 2499-2511.

CHANOCK, R. M., HAYFLICK, L. & BARILE, M. 1962. Growth on artificial medium of an agent associated with atypical pneumonia and its identification as a PPLO. *Proceedings of the National Academy of Sciences of the United States of America,* 48**,** 41.

CHEN, T. A. & LIAO, C. H. 1975. Corn stunt spiroplasma: isolation, cultivation, and proof of pathogenicity. *Science,* 188**,** 1015-1017.

CITTI, C. & BLANCHARD, A. 2013. Mycoplasmas and their host: emerging and re-emerging minimal pathogens. *Trends in microbiology,* 21**,** 196-203.

CLARK, M. I. 2019. 30 - Management of Breeding in Small Poultry Production Units. *In:* NOAKES, D. E., PARKINSON, T. J. & ENGLAND, G. C. W. (eds.) *Veterinary Reproduction and Obstetrics (Tenth Edition).* St. Louis (MO): W.B. Saunders.

COUCH, R., CATE, T. R. & CHANOCK, R. 1964. Infection with artificially propagated Eaton agent (*Mycoplasma pneumoniae*): implications for development of

attenuated vaccine for cold agglutinin-positive pneumonia. *JAMA,* 187**,** 442-447.

DAVIS, R. E. & WORLEY, J. F. 1973. Spiroplasma: motile, helical microorganism associated with corn stunt disease. *Phytopathology,* 63**,** 403-408.

DAVIS, R. E., WORLEY, J. F., WHITCOMB, R. F., ISHIJIMA, T. & STEERE, R. L. 1972. Helical filaments produced by a mycoplasma-like organism associated with corn stunt disease. *Science,* 176**,** 521-523.

DYBVIG, K. & VOELKER, L. L. 1996. Molecular biology of mycoplasmas. *Annual review of microbiology,* 50**,** 25-57.

EGGLETON, D., LEWIS, P. F., HALL, W. T. K., AUSTRALIAN AGRICULTURAL COUNCIL. ANIMAL HEALTH COMMITTEE. WORKING PARTY OF PRINCIPAL LABORATORY, O. & AUSTRALIAN BUREAU OF ANIMAL, H. 1976. *Avian mycoplasmosis,* [Canberra], Australian Bureau of Animal Health.

EVANS, J., LEIGH, S., PURSWELL, J., JACOB, R., PEEBLES, E., COLLIER, S. & BRANTON, S. 2012. A comparative study of live attenuated F strain–derived *Mycoplasma gallisepticum* vaccines. *Avian diseases,* 56**,** 396-401.

FALLON, R. 1967. Mycoplasmas and their role as rodent pathogens. *Laboratory Animals,* 1**,** 43-54.

FELICE, V., LUPINI, C., MESCOLINI, G., SILVEIRA, F., GUERRINI, A., CATELLI, E. & DI FRANCESCO, A. 2020. Molecular detection and characterization of *Mycoplasma gallisepticum* and *Mycoplasma synoviae* strains in backyard poultry in Italy. *Poultry science,* 99**,** 719-724.

FORREST, M. & BRADBURY, J. M. 1984. *Mycoplasma glycophilum*, a New Species of Avian Origin. *Microbiology,* 130**,** 597-603.

FUJISAWA, S., MURATA, S., TAKEHARA, M., KATAKURA, K., HMOON, M. M., WIN, S. Y. & OHASHI, K. 2019. Molecular detection and genetic characterization of *Mycoplasma gallisepticum*, *Mycoplasma synoviae*, and infectious bronchitis virus in poultry in Myanmar. *BMC Veterinary Research,* 15.

GRÓZNER, D., SULYOK, K. M., KREIZINGER, Z., RÓNAI, Z., JÁNOSI, S., TURCSÁNYI, I., KÁROLYI, H. F., KOVÁCS, Á. B., KISS, M. J., VOLOKHOV, D. & GYURANECZ, M. 2019. Detection of *Mycoplasma anatis*, *M. anseris*, *M. cloacale* and *Mycoplasma* sp. 1220 in waterfowl using species-specific PCR assays. *PLOS ONE,* 14**,** e0219071.

HAMMOND, P. P., RAMÍREZ, A. S., MORROW, C. J. & BRADBURY, J. M. 2009. Development and evaluation of an improved diagnostic PCR for *Mycoplasma synoviae* using primers located in the haemagglutinin encoding gene vlhA and its value for strain typing. *Veterinary Microbiology,* 136**,** 61-68.

HARASAWA, R., PITCHER, D. G., RAMÍREZ, A. S. & BRADBURY, J. M. 2004. A putative transposase gene in the 16S–23S rRNA intergenic spacer region of *Mycoplasma imitans*. *Microbiology,* 150**,** 1023-1029.

HASHEMI, S., MAHZOUNIEH, M., SHEIKHI, N. & EBRAHIMI, A. 2018. Application of high-resolution melting-curve analysis on pvpA gene for detection and classification of *Mycoplasma gallisepticum* strains. *Microbial Pathogenesis,* 124**,** 365-371.

HEALTH, W. O. F. A. 2014a. Chapter 1.1. 6. Principles and methods of validation of diagnostic assays for infectious diseases. OIE Paris, France.

HEALTH, W. O. F. A. 2014b. Validation guideline, chapter 3.6. 5: statistical approaches to validation. OIE Paris, France.

HENEGARIU, O., HEEREMA, N., DLOUHY, S., VANCE, G. & VOGT, P. 1997. Multiplex PCR: critical parameters and step-by-step protocol. *Biotechniques,* 23**,** 504-511.

HONG, Y., GARCÍA, M., LEITING, V., BENČINA, D., DUFOUR-ZAVALA, L., ZAVALA, G. & KLEVEN, S. H. 2004. Specific detection and typing of *Mycoplasma synoviae* strains in poultry with PCR and DNA sequence analysis targeting the hemagglutinin encoding gene vlhA. *Avian Diseases,* 48**,** 606-616.

JARQUIN, R., SCHULTZ, J., HANNING, I. & RICKE, S. C. 2009. Development of a real-time polymerase chain reaction assay for the simultaneous detection of *Mycoplasma gallisepticum* and *Mycoplasma synoviae* under industry conditions. *Avian diseases,* 53**,** 73-77.

JORDAN, F. T. W. 1985. Gordon Memorial Lecture: People, Poultry and Pathogenic Mycoplasmas. *World's Poultry Science Journal,* 41**,** 226-239.

KAHYA ÖZGE YILMAZ, S., ÇARLI, K. T., EYİGÖR, A. & TEMELLİ, S. 2015. Detection of *Mycoplasma gallisepticum* and *Mycoplasma synoviae* by Real-Time PCRs and *Mycoplasma gallisepticum*-antibody Detection by an ELISA in Chicken Breeder Flocks. *Damızlık Tavuk Kümeslerinde Mycoplasma gallisepticumve Mycoplasma synoviae'nin Gerçek Zamanlı PCR'lar ile ve Mycoplasma gallisepticum Antikorlarının ELISA ile Tespiti.,* 21**,** 361-366.

KAHYA, S., TEMELLI, S., EYIGOR, A. & CARLI, K. T. 2010. Real-time PCR culture and serology for the diagnosis of *Mycoplasma gallisepticum* in chicken breeder flocks. *Veterinary Microbiology,* 144**,** 319-324.

KEMPF, I., BLANCHARD, A., GESBERT, F., GUITTET, M. & BENNEJEAN, G. 1994. Comparison of antigenic and pathogenic properties of *Mycoplasma iowae* strains and development of a pcr-based detection assay. *Research in Veterinary Science,* 56**,** 179-185.

KISS, I., MATIZ, K., KASZANYITZKY, É., CHÁVEZ, Y. & JOHANSSON, K. E. 1997a. Detection and identification of avian mycoplasmas by polymerase chain reaction and restriction fragment length polymorphism assay. 58**,** 23-30.

KISS, I., MATIZ, K., KASZANYITZKY, É., CHÁVEZ, Y. & JOHANSSON, K. E. 1997b. Detection and identification of avian mycoplasmas by polymerase chain reaction and restriction fragment length polymorphism assay. *Veterinary Microbiology,* 58**,** 23-30.

KLEVEN, S. 2008. Control of avian mycoplasma infections in commercial poultry. *Avian diseases,* 52**,** 367-374.

KRALIK, P. & RICCHI, M. 2017. A Basic Guide to Real Time PCR in Microbial Diagnostics: Definitions, Parameters, and Everything. *Frontiers in Microbiology,* 8.

KUBISTA, M., ANDRADE, J. M., BENGTSSON, M., FOROOTAN, A., JONÁK, J., LIND, K., SINDELKA, R., SJÖBACK, R., SJÖGREEN, B., STRÖMBOM, L., STÅHLBERG, A. & ZORIC, N. 2006. The real-time polymerase chain reaction. *Molecular Aspects of Medicine,* 27**,** 95-125.

KUMAR, A., RAHAL, A., CHAKRABORTY, S., VERMA, A. K. & DHAMA, K. 2014. *Mycoplasma agalactiae*, an Etiological Agent of Contagious Agalactia in Small Ruminants: A Review. *Vet Med Int,* 2014**,** 286752.

KUO, H.-C., LO, D.-Y., CHEN, C.-L., TSAI, Y.-L., PING, J.-F., LEE, C.-H., LEE, P.-Y. A. & CHANG, H.-F. G. 2017. Rapid and sensitive detection of *Mycoplasma synoviae* by an insulated isothermal polymerase chain reaction-based assay on a field-deployable device. 96**,** 35-41.

LAUERMAN, L. 1998. Mycoplasma PCR assays: nucleic acid amplification assays for diagnosis of animal diseases. *Department of Agriculture and Industries, Alabama. 150p*.

LAUERMAN, L. H., HOERR, F. J., SHARPTON, A. R., SHAH, S. M. & VAN SANTEN, V. L. 1993. Development and application of a polymerase chain reaction assay for *Mycoplasma synoviae*. *Avian Diseases***,** 829-834.

LEVISOHN, S. & KLEVEN, S. 2000. Avian mycoplasmosis (*Mycoplasma gallisepticum*). *Revue Scientifique et Technique-Office International des Epizooties,* 19**,** 425-434.

LORENZ, T. C. 2012. Polymerase chain reaction: basic protocol plus troubleshooting and optimization strategies. *JoVE (Journal of Visualized Experiments)***,** e3998.

LORIA, G., FERRANTELLI, E., GIARDINA, G., VECCHI, L. L., SPARACINO, L., OLIVERI, F., MCAULIFFE, L. & NICHOLAS, R. 2008. Isolation and characterization of unusual *Mycoplasma* spp. from captive Eurasian Griffon (*Gyps fulvus*) in Sicily. *Journal of Wildlife Diseases,* 44**,** 159-163.

LUSCOMBE, N. M., GREENBAUM, D. & GERSTEIN, M. 2001. What is bioinformatics? An introduction and overview. *Yearbook of medical informatics,* 1**,** 2.

MACKAY, I. M. 2004. Real-time PCR in the microbiology laboratory. *Clinical Microbiology and Infection,* 10**,** 190-212.

MARDASSI, B. B. A., MOHAMED, R. B., GUERIRI, I., BOUGHATTAS, S. & MLIK, B. 2005. Duplex PCR to differentiate between *Mycoplasma synoviae* and Mycoplasma gallisepticum on the basis of conserved species-specific

sequences of their hemagglutinin genes. *Journal of clinical microbiology,* 43**,** 948-958.

MARMION, B. & GOODBURN, G. M. 1961. Effect of an organic gold salt on Eaton's primary atypical pneumonia agent and other observations. *Nature,* 189**,** 247-248.

MAROIS, C., SAVOYE, C., KOBISCH, M. & KEMPF, I. 2002. A reverse transcription-PCR assay to detect viable *Mycoplasma synoviae* in poultry environmental samples. *Veterinary Microbiology,* 89**,** 17-28.

MCAULIFFE, L., ELLIS, R. J., LAWES, J. R., AYLING, R. D. & NICHOLAS, R. A. 2005. 16S rDNA PCR and denaturing gradient gel electrophoresis; a single generic test for detecting and differentiating *Mycoplasma* species. *Journal of Medical Microbiology,* 54**,** 731-739.

MEKKES, D. & FEBERWEE, A. 2005. Real-time polymerase chain reaction for the qualitative and quantitative detection of *Mycoplasma gallisepticum*. *Avian Pathology,* 34**,** 348-354.

MIYATA, M. & SETO, S. 1999. Cell reproduction cycle of mycoplasma. *Biochimie,* 81**,** 873-878.

MKHABELA, T. & NYHODO, B. 2011. Farm and retail prices in the South African poultry industry: do the twain meet?

MOALIC, P.-Y., GESBERT, F., LAIGRET, F. & KEMPF, I. 1997. Evaluation of polymerase chain reaction for detection of *Mycoplasma meleagridis* infection in turkeys. 58**,** 187-193.

MORETTI, S. A. 2012. A molecular study of *Mycoplasma gallisepticum* field isolates from poultry in southern Africa. *University of the Free State*.

MOSCOSO, H., THAYER, S. G., HOFACRE, C. L. & KLEVEN, S. H. 2004. Inactivation, storage, and PCR detection of mycoplasma on FTA® filter paper. *Avian Diseases,* 48**,** 841-850.

MTILENI, B. J., MUCHADEYI, F. C., MAIWASHE, A., CHIMONYO, M., MAPIYE, C. & DZAMA, K. 2012. Influence of socioeconomic factors on production constraints faced by indigenous chicken producers in South Africa. 45**,** 67-74.

NASCIMENTO, E. R., PEREIRA, V., NASCIMENTO, M. & BARRETO, M. 2005. Avian mycoplasmosis update. *Brazilian Journal of Poultry Science,* 7**,** 1-9.

NASCIMENTO, E. R., YAMAMOTO, R., HERRICK, K. R. & TAIT, R. C. 1991. Polymerase chain reaction for detection of *Mycoplasma gallisepticum*. *Avian Diseases***,** 62-69.

ONGOR, H., KALIN, R., KARAHAN, M., CETINKAYA, B., MCAULIFFE, L. & NICHOLAS, R. A. 2008. Isolation of *Mycoplasma bovis* from broiler chickens in Turkey. *Avian Pathology,* 37**,** 587-588.

PABLA, S. S. & PABLA, S. S. 2008. Real-time polymerase chain reaction : A revolution in diagnostics. *Resonance : Journal of Science Education,* 13**,** 369-377.

PAGE, A. J., CUMMINS, C. A., HUNT, M., WONG, V. K., REUTER, S., HOLDEN, M. T. G., FOOKES, M., FALUSH, D., KEANE, J. A. & PARKHILL, J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics,* 31**,** 3691-3693.

PATTISON, M., MCMULLIN, P., BRADBURY, J. M. & ALEXANDER, D. 2007. *Poultry diseases*, Elsevier Health Sciences.

PESTANA, E., BELAK, S., DIALLO, A., CROWTHER, J. R. & VILJOEN, G. J. 2010. *Early, rapid and sensitive veterinary molecular diagnostics-real time PCR applications*, Springer Science & Business Media.

PESTANA, E. A. 2010. *Early, rapid, and sensitive veterinary molecular diagnostics : real time PCR applications,* Dordrecht, Springer.

POLAK-VOGELZANG, A. A. 1977. Survival of *Mycoplasma gallisepticum* in mains water. *Avian Pathology,* 6**,** 93-95.

POVEDA, J. B., CARRANZA, J., MIRANDA, A., GARRIDO, A., HERMOSO, M., FERNANDEZ, A. & DOMENECH, J. 1990. An epizootiological study of avian mycoplasmas in Southern Spain. *Avian Pathology,* 19**,** 627-633.

RAGGI, E., THOMAS, K. & VAN VUGT, S. 2011. *Beginning Ubuntu Linux*, Springer.

RAMÍREZ, A. S., NAYLOR, C. J., HAMMOND, P. P. & BRADBURY, J. M. 2006. Development and evaluation of a diagnostic PCR for *Mycoplasma synoviae* using primers located in the intergenic spacer region and the 23S rRNA gene. 118**,** 76-82.

RAVIV, Z. & KLEVEN, S. H. 2009. The development of diagnostic real-time TaqMan PCRs for the four pathogenic avian mycoplasmas. *Avian diseases,* 53**,** 103-107.

RAZIN, S. 1985. Molecular biology and genetics of mycoplasmas (Mollicutes). *Microbiological reviews,* 49**,** 419.

RAZIN, S., BARILE, M., HARASAWA, R., AMIKAM, D. & GLASER, G. 1983. Characterization of the mycoplasma genome. *The Yale journal of biology and medicine,* 56**,** 357.

RAZIN, S. & HAYFLICK, L. 2010. Highlights of mycoplasma research—An historical perspective. *Biologicals,* 38**,** 183-190.

RAZIN, S., YOGEV, D. & NAOT, Y. 1998. Molecular biology and pathogenicity of mycoplasmas. *Microbiol. Mol. Biol. Rev.,* 62**,** 1094-1156.

REBRIKOV, D. & TROFIMOV, D. Y. 2006. Real-time PCR: a review of approaches to data analysis. *Applied biochemistry and microbiology,* 42**,** 455-463.

ROENIGK, W. 1999. Muscle growth and development. Keynote address: world poultry consumption. *Poultry Science,* 78**,** 722-728.

SALLING, H. & BANG-CHRISTENSEN, S. 2016. Multi-primer qPCR assay capable of highly efficient and specific detection of the vast majority of all known Mycoplasma. *Biologicals,* 44**,** 129-138.

SAPA. 2018. *South African Poultry Association 2018, Industry Profile* [Online]. Available: http://www.sapoultry.co.za [Accessed 2020].

SAPA. 2021. *South African Poultry Association, Annual Report* [Online]. Available: https://www.sapoultry.co.za/wp-content/uploads/2022/06/SAPA-ANNUAL-REPORT-2021.pdf [Accessed].

SARAYA, T. 2016. The History of *Mycoplasma pneumoniae* Pneumonia. *Front Microbiol,* 7**,** 364.

SEEMANN, T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics,* 30**,** 2068-2069.

SHERWOOD, L., WILLEY, J. M. & WOOLVERTON, C. 2011. *Prescott's microbiology*, McGraw-Hill.

SITTO, F. & BATTISTUZZI, F. U. 2020. Estimating Pangenomes with Roary. *Molecular Biology and Evolution,* 37**,** 933-939.

SPRYGIN, A., ANDREYCHUK, D., KOLOTILOV, A., VOLKOV, M., RUNINA, I., MUDRAK, N., BORISOV, A., IRZA, V., DRYGIN, V. & PEREVOZCHIKOVA, N. 2010. Development of a duplex real-time TaqMan PCR assay with an internal control for the detection of *Mycoplasma gallisepticum* and *Mycoplasma synoviae* in clinical samples from commercial and backyard poultry. *Avian pathology,* 39**,** 99-109.

SWAYNE, D. E. & GLISSON, J. R. 2013. *Diseases of poultry,* Ames, Iowa, Wiley-Blackwell.

TETTELIN, H., MASIGNANI, V., CIESLEWICZ, M. J., DONATI, C., MEDINI, D., WARD, N. L., ANGIUOLI, S. V., CRABTREE, J., JONES, A. L., DURKIN, A. S., DEBOY, R. T., DAVIDSEN, T. M., MORA, M., SCARSELLI, M., MARGARIT Y ROS, I., PETERSON, J. D., HAUSER, C. R., SUNDARAM, J. P., NELSON, W. C., MADUPU, R., BRINKAC, L. M., DODSON, R. J., ROSOVITZ, M. J., SULLIVAN, S. A., DAUGHERTY, S. C., HAFT, D. H., SELENGUT, J., GWINN, M. L., ZHOU, L., ZAFAR, N., KHOURI, H., RADUNE, D., DIMITROV, G., WATKINS, K., O'CONNOR, K. J. B., SMITH, S., UTTERBACK, T. R., WHITE, O., RUBENS, C. E., GRANDI, G., MADOFF, L. C., KASPER, D. L., TELFORD, J. L., WESSELS, M. R., RAPPUOLI, R. & FRASER, C. M. 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial "pan-genome". *Proceedings of the National Academy of Sciences,* 102**,** 13950-13955.

TULLY, J. G., WHITCOMB, R. F., CLARK, H. F. & WILLIAMSON, D. L. 1977. Pathogenic mycoplasmas: cultivation and vertebrate pathogenicity of a new spiroplasma. *Science,* 195**,** 892-894.

UMAR, S., MUNIR, M. T., UR-REHMAN, Z., SUBHAN, S., AZAM, T. & SHAH, M. A. A. 2017. Mycoplasmosis in poultry: update on diagnosis and preventive measures. *World's Poultry Science Journal,* 73**,** 17-28.

VALCESCHINI, E. Poultry meat trends and consumer attitudes. Proceedings of the XII European Poultry Conference, 2006. 1-10.

VAN KUPPEVELD, F., JOHANSSON, K., GALAMA, J., KISSING, J., BÖLSKE, G., VAN DER LOGT, J. & MELCHERS, W. 1994. Detection of mycoplasma contamination in cell cultures by a mycoplasma group-specific PCR. *Applied and environmental microbiology,* 60**,** 149-152.

VAN KUPPEVELD, F., VAN DER LOGT, J., ANGULO, A., VAN ZOEST, M., QUINT, W., NIESTERS, H., GALAMA, J. & MELCHERS, W. 1992. Genus-and species-specific identification of mycoplasmas by 16S rRNA amplification. *Applied and environmental microbiology,* 58**,** 2606-2615.

VAN RYSSEN, J. 2001. Poultry litter as a feedstuff for ruminants: A South African scene. *S. Afr. J. Anim. Sci,* 2**,** 1-8.

VILJOEN, G. J., NEL, L. H. & CROWTHER, J. R. 2005. Molecular diagnostic PCR handbook. Dordrecht: Springer.

WANG, H., FADL, A. & KHAN, M. 1997. Multiplex PCR for avian pathogenic mycoplasmas. *Molecular and Cellular Probes,* 11**,** 211-216.

WANG, Y., WHITHEAR, K. & GHIOCAS, E. 1990. Isolation of *Mycoplasma gallinarum* and *M. gallinaceurn* from the reproductive tract of hens. *Australian Veterinary Journal,* 67**,** 31-32.

WHITCOMB, R. & WILLIAMSON, D. 1975. Helical wall-free prokaryotes in insects: multiplication and pathogenicity. *Annals of the New York Academy of Sciences,* 266**,** 260-275.

WHITFORD, H. W., ROSENBUSCH, R. F. & LAUERMAN, L. H. 1994. Mycoplasmosis in animals: laboratory diagnosis, *Iowa State University Press*.

WHITHEAR, K. 1996. Control of avian mycoplasmoses by vaccination. *Revue scientifique et technique-Office international des épizooties,* 15**,** 1527-1554.

WHITNALL, T. & PITTS, N. 2019. Global trends in meat consumption. *Agricultural Commodities,* 9**,** 96-99.

WIESE, I. H. & BOT, J. 1971. Pesticide regulation in South Africa. Springer New York.

WILHELM, J. & PINGOUD, A. 2003. Real-Time Polymerase Chain Reaction. *ChemBioChem,* 4**,** 1120-1128.

WINDHORST, H.-W. 2006. Changes in poultry production and trade worldwide. *World's Poultry Science Journal,* 62**,** 585-602.

WITTWER, C. T., HERRMANN, M. G., GUNDRY, C. N. & ELENITOBA-JOHNSON, K. S. 2001. Real-time multiplex PCR assays. *Methods,* 25**,** 430-442.

YADAV, J. P., TOMAR, P., SINGH, Y. & KHURANA, S. K. 2021. Insights on *Mycoplasma gallisepticum* and *Mycoplasma synoviae* infection in poultry: A systematic review. *Animal Biotechnology***,** 1-10.

ZHAO, S. & YAMAMOTO, R. 1993a. Amplification of *Mycoplasma iowae* using polymerase chain reaction. *Avian diseases***,** 212-217.

ZHAO, S. & YAMAMOTO, R. 1993b. Detection of *Mycoplasma meleagridis* by polymerase chain reaction. 36**,** 91-97.

ZHU, L., KONSAK, B. M., OLAOGUN, O. M., AGNEW-CRUMPTONA, R., KANCI, A., MARENDA, M. S., BROWNING, G. F. & NOORMOHAMMADI, A. H. 2017. Identification of a new genetic marker in *Mycoplasma synoviae* vaccine strain MS-H and development of a strategy using polymerase chain reaction and high-resolution melting curve analysis for differentiating MS-H from field strains. *Veterinary Microbiology,* 210**,** 49-55.

Appendix A1: Summary of avian mycoplasma PCR-based assays

Paper & Reference	Mycoplasma species target	Gene / genome region targeted	Primer sequence (5′ – 3′)	Probe sequence	Type of assay	Advantages & Disadvantages or Use
PCR for detection of MG (Nascimento et al., 1991)	*M. gallisepticum*	DNA sequence data of fMG-2	Amp L: (GGATCCCATCTCG ACCACGAGAAAA) Amp R: (AGTAGTCAATGAGT GACTAACTTTC)	N/A	Conventional PCR	Highly sensitive to MG strains.
Development and Application of a Polymerase Chain Reaction Assay for Mycoplasma synoviae (Lauerman et al., 1993)	*M. synoviae*	16S rRNA	MS-1: (GAAGCAAAATAGTG ATATCA) MS-2 (GTCGTCTCCGAAG TTAACAA)	N/A	Conventional PCR	Used only for detection of MS.
Amplification of Mycoplasma iowae Using Polymerase Chain Reaction (Zhao and Yamamoto, 1993a)	*M. iowae*	pMI-12 (Left & Right) (Internal primers Li & Ri)	L (GAATTCTGAATCTT CATTTCTTAAA) R (CAGATTCTTTAATA ACTTATGTATC) Li (AATGGCAACTTTTG AGTCATCATCAA) Ri (CTTATGTATCAAAC AATAAAGAAGCAG)	N/A	Conventional PCR	Used only for detection of MI.

Paper & Reference	Mycoplasma species target	Gene / genome region targeted	Primer sequence (5′ – 3′)	Probe sequence	Type of assay	Advantages & Disadvantages or Use
Detection of Mycoplasma meleagridis by polymerase chain reaction (Zhao and Yamamoto, 1993b)	*M. meleagridis*	pMM-2 (Left & Right) (Internal primers Li & Ri)	L (GGATCCTTATATTAATTTAAACAAATTAATGA) R (GAATTCTTCTTTATTATTCAAAAGTAAAGTAC) Li (AGGACCAGATTTTCCTACTGGCGGCA) Ri (CTCAATTCAGCAATTGCTGTAGCTTG)	N/A	Conventional PCR	Used only for detection of MM.
Comparison of antigenic and pathogenic properties of MI strains & develop. of a PCR-based detection assay (Kempf et al., 1994)	*M. iowae*	16S rRNA (M24293)	$RNAI_0$ (TTGTTATCGCATGAGAGAATG) RNA_3 (ACGAGCTGACGACAACCATGCAC $RNAI_2$ (CTAATACCGCATAGGACATTG) $RNAI_1$ (GATACCGTCACACAGAAAGC)	N/A	Conventional PCR	Used for detection of MI strains.
Detection & ID of avian mycoplasmas by PCR & RFLP assay **(Kiss et al., 1997a)**	*M. gallisepticum* *M. iowae* *M. meleagridis* *M. synoviae*	16S rRNA	GPO-3 (GGGAGCAAACAGGATTAGATACCCT) MGSO (TGCACCATCTGTCACTCTGTTAACCTC) Species-specific (AACACCAGAGGCAAGGCGAGG) (ACGGATTTGCAACTGTTTGTATTGG)	N/A	Conventional PCR	Results are best obtained from samples swabbed withing a day. Can be applied for routine diagnostic techniques.

Paper & Reference	Mycoplasma species target	Gene / genome region targeted	Primer sequence (5′ – 3′)	Probe sequence	Type of assay	Advantages & Disadvantages or Use
Evaluation of PCR for detection of MM infection in turkeys (Moalic et al., 1997)	*M. meleagridis*	16S rDNA in V1 16S rDNA in V4	MM90f (CGAGCGAAGTTTTTCGGAAC) MM469r GGTACCGTCAGGATAAATGC	N/A	Conventional PCR	No cross-amplification with other avian mycoplasma. Modified from (Boyle et al., 1995)
Multiplex PCR for avian pathogenic mycoplasmas (Wang et al., 1997)	*M. gallisepticum*	DNA sequence data of fMG-2	MG 1 (GGATCCCATCTCGACCAGGAGAAAA) MG 2 (CTTTCAATCAGTGAGTAACTGATGA)	N/A	Conventional PCR	The multiplex PCR is specific, sensitive and cost effective. Used only for the detection of MG, MS, MM and MI.
	M. synoviae	16S rRNA sequences	MS 1 (GAAGCAAATAGTGATATCA) MS 2 (GTCGTCTCGAAGTTAACAA)			
	M. meleagridis	MM species-specific recomb. (pMM-2) DNA sequence	MM 1 (GGATCCTAATATTAATTTAAACAAATTAATGA) MM 2 (GAATTCTTCTTTATTATTCAAAAGTAAAGTAC)			
	M. iowae	MI recombinant DNA probe sequence data (pMI- 12)	MI 1 (GAATTCTGAATCTTCATTTCTTAAA) MI 2 (CAGATTCTTTAATAACTTATGTATC)			
A RT-PCR assay to detect viable MS in poultry environmental samples (Marois et al., 2002)	*M. synoviae*	16S rRNA to cDNA	MGSO (TGCACCATCTGTCACTCTGTTAACCTC) GPO3 (GGGAGCAAACAGGATTAGATACCCT)	NA	RT-PCR (conventional)	Used only for detection of MS. Detects live and recently dead bacteria.

Paper & Reference	Mycoplasma species target	Gene / genome region targeted	Primer sequence (5′ – 3′)	Probe sequence	Type of assay	Advantages & Disadvantages or Use
Real-time PCR for detection of MG in chicken trachea (Carli and Eyigor, 2003)	*M. gallisepticum*	Region within the sequence of MG lipoprotein gene partial codons	MG1 (GATTTCGAAGAATC AACTGT) MG2 (AAGGGATTAATATT CCCAAC)	Not stated	Real-time PCR	Rapid detection of MG from tracheal swabs of sub-clinically infected chickens.
Specific detection & typing of MS strains with PCR & DNA sequence analysis targeting the hemagglutinin encoding gene vlhA (Hong et al., 2004)	*M. synoviae*	vlhA	MSF: (GATGCGTAAAATAA AAGGAT) MSR: (GCTTCTGTTGTAGT TGCTTC)	N/A	Conventional PCR	Specific detection and sequencing analysis of MS.
Inactivation, Storage, and PCR Detection of Mycoplasma on FTA® Filter Paper (Moscoso et al., 2004)	*M. gallisepticum*	mgc2	F (CGCAATTTGGTCCT AATCCCCAACA) R (TAAACCCACCTCCA GCTTTATTTCC)	N/A	Conventional PCR	Used for inactivation of MG and MS on Flinders Technology Associates (FTA®) paper.
	M. synoviae	vlhA	F (GATGCGTAAAATAA AAGGAT) R (GCTTCTGTTGTAGT TGCTTC)			

Paper & Reference	Mycoplasma species target	Gene / genome region targeted	Primer sequence (5′ – 3′)	Probe sequence	Type of assay	Advantages & Disadvantages or Use
Duplex PCR to Differentiate between MS and MG based on conserved species-specific sequences of their Hemagglutinin Genes (Mardassi et al., 2005)	*M. gallisepticum*	MG pMGA1.2	$pMGAF_o$ (GTGAAGAAAAAAAACATATTAAAGTTT) $pMGAR_o$ (CTAAGATGGATTTGAAACATTAGT) $pMGAF1_i$ CTAGTTAATACTAGTGATCAAGTGAAACTA $pMGAR1_i$ TTGAACATTGTTCTTTGGAACCATCAT	N/A	Conventional PCR	Detection and differentiation of both MG and MS can be achieved in a single reaction.
	M. synoviae	MS2/12	$MS1.2F_o$ (AAACTACAAAACTTTGTAATGGCT) $MS1.2R_o$ (TTACAAGTACGGTGTTTAATCAAT) $MS1.2F1_i$ (ATTACCAAGCAGATGGTTACGACGT) $MS1.2R2_i$ (AGTTATAGTAACTCCGTTTGTTCCA)			
Real-time PCR for the qualitative and quantitative detection of MG (Mekkes and Feberwee, 2005)	*M. gallisepticum*	16S rRNA	Mg14F: (GAGCTAATCTGTAAAGTTGGTC) Mg13R (GCTTCCTTGCGGTTAGCAAC)	Not stated	Real-time PCR	Used for the qualitative and quantitative detection of MG.
16S rDNA and DGGE; a single generic test for detecting & differentiating *Mycoplasma* species. (McAuliffe et al., 2005)	Avian mycoplasmas	16S RNA	GC-341F (CGCCCGCCGCGCGCGGCGGGCGGGGCGGGGGCACGGGGGGCCTACGGGAGGCAGCAG) R543 (ACCTATGTATTACCGCG)	N/A	Conventional PCR & Denaturing gradient gel electrophoresis (DGGE)	Differentiation of 67 mycoplasma species of human and vet origin. Diagnosis of mycoplasma infection can be made directly from clinical samples in less than 24 h.

Paper & Reference	Mycoplasma species target	Gene / genome region targeted	Primer sequence (5′ – 3′)	Probe sequence	Type of assay	Advantages & Disadvantages or Use
Development and evaluation of a diagnostic PCR for MS (Ramírez et al., 2006)	*M. synoviae*	Intergenic spacer region and 23S rRNA gene	Ms1FF (TTAACTGAAAGCTTTTAG) Ms2FF (TAAAAGCGGTTGTGTATCGC) 23SR (CGCAGGTTTGCACGTCCTTCATCG)	N/A	Conventional PCR	Ms1F-23S failed & was not used. PCR based on primers Ms2F-23S is highly specific & sensitive for Ms.
Epidemiological survey on Mycoplasma gallisepticum and M. synoviae by multiplex PCR in commercial poultry (Buim et al., 2009)	*M. gallisepticum*	DNA sequence data of fMG-2	MG-f: (GGATCCCATCTCGACCACGAGAAAA) MG-r: (CTTTCAATCAGTGAGTAACTGATGA)	N/A	Conventional PCR	Used for simultaneous detection of MG and MS. Can be used as a diagnostic assay.
		Vaccine strain	MGF-f: (TAACCCTTCATCACCTCATCTAGAG) MGF-r: (CTGTTTGCTAAAGAACAAGTTGATC)			
	M. synoviae	MSLF MSLR	MS-f: (GAGAAGCAAAATAGTGATATCA) MS-r: (CAGTCGTCTCCGAAGTTAACAA)			
Development and evaluation of an improved diag. PCR for MS using primers in haemagglutinin encoding gene vlhA and its value for strain typing. (Hammond et al., 2009)	*M. synoviae*	vlhA	VlhAF (ATTAGCAGCTAGTGCAGTGGCC) VlhAR1 (CAGCGCTAGTTTTTGTTTTTTGG) VlhAR2 (AGTAACCGATCCGCTTAATGC)	N/A	Conventional PCR	Used for diagnosis and strain typing of MS.

Paper & Reference	Mycoplasma species target	Gene / genome region targeted	Primer sequence (5′ – 3′)	Probe sequence	Type of assay	Advantages & Disadvantages or Use
Development of a Real-Time PCR Assay for the Simultaneous detection of MG and MS under industry conditions (Jarquin et al., 2009)	*M. gallisepticum*	16S rRNA gene	MG-14F: (GAGCTAATCTGTAA AGTTGGTC) MG-13R: (GCTTCCTTGCGGTT AGCAAC)	SYBR green	SYBR green real-time PCR assay.	Used for the detection of MG and MS only.
	M. synoviae	MSLF MSLR	GAGAAGCAAAATAG TGATATCA CAGTCGTCTCCGAA GTTAACAA			
The Development of Diagnostic Real-Time TaqMan PCRs for the Four Pathogenic Avian Mycoplasmas (Raviv and Kleven, 2009)	*M. gallisepticum*	mgc2 (AE015450)	TTGGGTTTAGGGAT TGGGATT CCAAGGGATTCAAC CATCTT	5TexRd-XN-TGATGATCCAAGAA CGTGAAGAACACC-3BHQ_2	Real-time PCR	Used for the detection of MG, MS, MM, and MI. Assay has high specificity and sensitivity.
	M. synoviae	16S-23S rDNA ISR (AY768810)	CTAAATACAATAGC CCAAGGCAA CCTCCTTTCTTACG GAGTACA	56-FAM-AGCGATACACAACC GCTTTTAGAAT-3BHQ_1		
	M. meleagridis	16S-23S rDNA ISR (AY762641)	AACAAGGTATCCCT ACGAGAAC CTCAGAGCCTTAAA CCAAGTCA	56-FAM-CCTCCTTTCTACGG AGTACATTAGTT-3BHQ_1		
	M. iowae	Upstream to 16S rDNA (U29676)	ATGAGTCCATTATTT ATGCTTCC TCCATTTCTTTTGAA CGTGCATT	56-FAM-CTGTGTTGTGTGAT GTTCTTTTGTTTG-3BHQ_1		
Real-time PCR, culture and serology for MG diagnosis in breeder chickens (Kahya et al., 2010)	*M. gallisepticum*	Region within the sequence of MG lipoprotein gene partial codons	MG1 (GATTTCGAAGAATC AACTGT) MG2 (AAGGGATTAATATT CCCAAC)	Not stated	Real-time PCR	Used only for diagnosis of MG.

Paper & Reference	Mycoplasma species target	Gene / genome region targeted	Primer sequence (5′ – 3′)	Probe sequence	Type of assay	Advantages & Disadvantages or Use
Development of a duplex real-time TaqMan PCR assay with an internal control for the detection of MG and MS in clinical samples from commercial and backyard poultry. (Sprygin et al., 2010)	*M. gallisepticum*	*mgc2* fragment (AY556229)	*mgc2*-F (GCTGGGTTGATTGTTGTTTCTT) *mgc2*-R (TCTTCACGTTCTTGGATCATCAT)	*mgc2*-probe (Cy5: CTCTT(G/C)GGTTTAGGGATTGGGATTCCG)	Duplex real-time TaqMan PCR	Detects of both MG and MS in a single-tube reaction. Simultaneously reveal MG and MS-specific sequences from different strains in the background of DNA extracts from related and non-related bacteria capable of colonizing the avian respiratory tract.
	M. synoviae	*vlhA* fragment (AF035624)	*vlhA*-F (CCAGGAGGTGGTACAGTTGAC) *vlhA*-R (TTAATGCTTCTTTAACT(G/A)AATCTGA)	*vlhA*-probe (FAM: CTGCTAAAACAGAAGCTAAAAC(C/T)GCTAT)		
	Internal control	*S3* gene fragment-avian reovirus (U20642)	IC-F (5'-TTTCTCCGATGGTGGCTACACTA-3') IC-R (5'-CGAAGTGCGACGTCCAAAAT-3')	IC-probe (TAMRA: ATGCTGGCCCTGTAAAGCTTGCGAA)		
Detection of MG and MS by Real-Time PCRs and MG-antibody Detection by an ELISA in Chicken Breeder Flocks (Kahya Özge Yilmaz et al., 2015)	*M. gallisepticum*	MG- lipoprotein gene	GATTTCGAAGAATCAACTGT AAGGGATTAATATTCCCAAC	Not stated	Real-time PCR	Used to determine prevalence of MG and MS in chicken breeder flocks in Turkey. The MG-16S rRNA primers yielded better results than the MG-lipoprotein primers.
		MG-16S rRNA gene	GAGCTAATCTGTAAAGTTGGTC GCTTCCTTGCGGTTAGCAAC			
	M. synoviae	MS-16S rRNA gene	GAGAAGCAAAATAGTGATATC TCGTCTCCGAAGTTAACAA			

Paper & Reference	Mycoplasma species target	Gene / genome region targeted	Primer sequence (5′ – 3′)	Probe sequence	Type of assay	Advantages & Disadvantages or Use
Multi-primer qPCR assay capable of highly efficient and specific detection of the vast majority of all known *Mycoplasma* (Salling and Bang-Christensen, 2016)	All known *Mycoplasma*	Primer mix A	F (5'-GGATTAGATACCCTAGTAGTCCACA-3') R (5'-CGTGTACCGTCGAATTAAGCA-3')	Not stated	Multi-primer qPCR	The assay not only detects the majority of all *Mycoplasma* species and strains, but also *S. citri* and *A. laidlawii* with high sensitivity and with an extremely high degree of specificity. Allows unambiguous detection of Mollicute infections and can be used for rapid testing in the biopharmaceutical industry.
			F (5'-ACTAAGTGTTGGCCAAAAGGTC-3') R (5'-CCTCCGAATTTATTTCTAAGCCTTTG-3')			
		Primer mix B	F (5'-TCATCATGCCTCTTACGAGTG-3') R (5'-GCGGTGTGTACAAGACCCGA-3')			
			F (5'-CTCCGCCTGAGTAGTATGC-3') R (5'-CACCTGTCTCAATGTTAACCTC-3')			
Rapid and sensitive detection of MS by an insulated isothermal PCR-based assay on a field-deployable device (Kuo et al., 2017)	*M. synoviae*	vlhA	F (CCAGGAGGTGGTACAGTTGAC) R (TTAATGCTTCTTTAACT(G/A)AATCTGA)	FAM-5'-CTGCTAAAACAGAAGCTAAAAC(C/T)GCTAT-3'-BHQ1	Insulated isothermal qPCR	Excellent analytical sensitivity and specificity in detection of MS nucleic acids.

Paper & Reference	Mycoplasma species target	Gene / genome region targeted	Primer sequence (5′ – 3′)	Probe sequence	Type of assay	Advantages & Disadvantages or Use
ID of new genetic marker in MS-H and development of a strategy using PCR & high-resolution melting curve analysis for MS-H & field strains differentiation (Zhu et al., 2017)	*M. synoviae*	vlhA	oppF-F2 (5'-ACTTTGTATAAAACG GTTCAAAAAGTG-3') oppF-R2 (5'-TGGCGTTATCCAAG AAAAAGTTAAA-3')	N/A	Conventional PCR	Used only for differentiation of MS-H vaccine from field/isolates and detection of MS isolates directly from field samples.
*Comparative evaluation of the pathogenicity of *M. gallinaceum* in chickens (Adeyemi et al., 2018)	*M. gallinaceum*	MGC-Forward	CGTGCCCCCTTGAT TGGGATAACGCTG	MGC probe (FAM: TAGCGCTAATACCG GATACTTAA)	Quantitative real-time PCR	Used as part of an experiment - to assess mycoplasma shedding from the trachea and cloaca over the duration of the experiment.
		MGC-Reverse	TAGCTAATGTTACG CACCCCGATCCCCT TGT			
High-resolution melting-curve analysis on pvpA gene for detection and classification of MG strains (Hashemi et al., 2018)	*M. gallisepticum*	pvpA	PvpA-F (GCCAMTCCAACTC AACAAGCTGA) PvpA-R (GGACGTSGTCCTG GCTGGTTAGC)	N/A	Conventional PCR	Products used for sequencing.

*Only avian mycoplasma information listed.

Paper & Reference	Mycoplasma species target	Gene / genome region targeted	Primer sequence (5′ – 3′)	Probe sequence	Type of assay	Advantages & Disadvantages or Use
*Molecular ID of MS from breeder chicken flocks showing arthritis in Egypt (Amer et al., 2019)	*M. synoviae*	vlhA	F (GATGCGTAAAATAAAAGGAT) R (GCTTCTGTTGTAGTTGCTTC)	N/A	Conventional PCR	Used only for MS field and vaccine strains.
*Molecular detection and genetic characterization of MG, MS & IBV in poultry in Myanmar (Fujisawa et al., 2019)	*M. gallisepticum*	pMGA1.2	pMGAFo: (GTGAAGAAAAAAAACATATTAAAGTTT) pMGARo: (CTAAGATGGATTTGAAACATTAGT)	N/A	Conventional PCR	Products used for sequencing.
	M. synoviae	MS2/12	MS1.2Fo: (AAACTACAAAACTTTGTAATGGCT) MS1.2Ro: (TTACAAGTACGGTGTTTAATCAAT)			
Detection of Mycoplasma anatis, M. anseris, M. cloacale and Mycoplasma sp. 1220 in waterfowl using species-specific PCR assays (Grózner et al., 2019)	*M. anatis*	dnaX	CAGAGATCAGTCTGTTTTAGAATTACTTTTTTCTCAGATGCTTGTGAAATACAACTT	N/A	Species-specific conventional PCR	Cross-amplification among the species may occur.
	M. anseris	pcrA	CTAAAAACTCCTAAAGACTTAGAAGAATCATCCTCACCTTCATCATTTTCTGTATA			
	M. cloacale	dnaX	TTCATCCGATAAGTTAAAACCTTGTTAAAACTGCTTTTGTATTTTTAGAATATAGT			
	Mycoplasma. sp. 1220	rpoB	CCGTGATACTGCTCAATTCGAATAGAAGTATAAACATCATCCTTAACAAGCT			

*Only avian mycoplasma information listed.

Paper & Reference	Mycoplasma species target	Gene / genome region targeted	Primer sequence (5′ – 3′)	Probe sequence	Type of assay	Advantages & Disadvantages or Use
Molecular detection and characterization of Mycoplasma gallisepticum and Mycoplasma synoviae strains in backyard poultry in Italy (Felice et al., 2020)	*M. gallisepticum*	pvpA	pvpA3F (GCCAMTCCAACTC AACAAGCTGA) pvpA4R (GGACGTSGTCCTG GCTGGTTAGC)	N/A	Conventional PCR	Products used for sequencing and phylogenetic analysis.
	M. synoviae	vlhA	vlhA-F (GATGCGTAAAATAA AAGGAT) vlhAR2 (AGTAACCGATCCG CTTAATGC)			

Appendix A2: Summary of frequently used genes/genome regions

Reference	Mycoplasma species target	Gene/genome region targeted	Primer sequence (5′ – 3′)
(Kiss et al., 1997a, Marois et al., 2002)	MS	16S rRNA	MGSO (TGCACCATCTGTCACTCTGTTAACCTC) GPO-3 (GGGAGCAAACAGGATTAGATACCCT)
(Amer et al., 2019, Felice et al., 2020, Hong et al., 2004, Moscoso et al., 2004)	MS	vlhA (Felice only forward primer)	F (GATGCGTAAAATAAAAGGAT) R (GCTTCTGTTGTAGTTGCTTC)
(Kuo et al., 2017, Sprygin et al., 2010)	MS	vlhA	*vlhA*-F (CCAGGAGGTGGTACAGTTGAC) *vlhA*-R (TTAATGCTTCTTTAACT(G/A)AATCTGA)
(Felice et al., 2020, Hammond et al., 2009)	MS	vlhA (only reverse primer)	VlhAR2 (AGTAACCGATCCGCTTAATGC)
(Fujisawa et al., 2019, Mardassi et al., 2005)	MS	MS2/12	MS1.2Fo: (AAACTACAAAACTTTGTAATGGCT) MS1.2Ro: (TTACAAGTACGGTGTTTAATCAAT)
	MG	pMGA1.2	pMGAFo: (GTGAAGAAAAAAAACATATTAAAGTTT) pMGARo: (CTAAGATGGATTTGAAACATTAGT)
(Jarquin et al., 2009, Kahya Özge Yilmaz et al., 2015, Mekkes and Feberwee, 2005)	MG	MG-16S rRNA gene	F (GAGCTAATCTGTAAAGTTGGTC) R (GCTTCCTTGCGGTTAGCAAC)
(Carli and Eyigor, 2003, Kahya Özge Yilmaz et al., 2015, Kahya et al., 2010)	MG	Region within the sequence of MG lipoprotein gene partial codons	MG1 (GATTTCGAAGAATCAACTGT) MG2 (AAGGGATTAATATTCCCAAC)
(Felice et al., 2020, Hashemi et al., 2018)	MG	pvpA	PvpA-F (GCCAMTCCAACTCAACAAGCTGA) PvpA-R (GGACGTSGTCCTGGCTGGTTAGC)
(Buim et al., 2009, Nascimento et al., 1991)	MG	fMG-2 (forward primers highly similar, reverse primers are reversed)	Amp L (GGATCCCATCTCGACCACGAGAAAA) Amp R (AGTAGTCAATGAGTGACTAACTTTC)
(Wang et al., 1997)			MG 1 (GGATCCCATCTCGACCAGGAGAAAA) MG 2 (CTTTCAATCAGTGAGTAACTGATGA)
(Adeyemi et al., 2018)	*M. gallinaceum*	MGC	F (CGTGCCCCCTTGATTGGGATAACGCTG) R (TAGCTAATGTTACGCACCCCGATCCCCTTGT)

Note: Kiss et al., (1997) used the 16S rRNA for MG, *M. iowae* and *M. meleagridis* as well.

Appendix B1: Completed alignment sequence comparison - *M. gallisepticum*

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Hypoxanthine-guanine phosphoribosyltransferase (EC 2.4.2.8)	545	0	Yes	-	-	-
FIG002540: Haloacid dehalogenase-like hydrolase	866	0	Yes	-	-	-
3'-to-5' oligoribonuclease A, Bacillus type	986	0	Yes	-	-	-
LSU ribosomal protein L3p (L3e)	710	0	Yes	-	-	-
L-lactate dehydrogenase (EC 1.1.1.27)	971	0	Yes	-	-	-
GTP-binding protein EngB	560	1	-	Yes	-	-
LSU ribosomal protein L4p (L1e)	629	1	-	Yes	-	-
Flavodoxin 1	512	2	-	Yes	-	-
Uracil phosphoribosyltransferase (EC 2.4.2.9)	632	2	-	Yes	-	-
Purine nucleoside phosphorylase (EC 2.4.2.1)	710	2	-	Yes	-	-
HMP-PP hydrolase (pyridoxal phosphatase) Cof, detected in a genetic screen for thiamin metabolic genes (PMID:15292217)	869	2	-	Yes	-	-
Ribonucleotide reductase of class Ib (aerobic), alpha subunit (EC 1.17.4.1)	2168	2	-	Yes	-	-
Inorganic pyrophosphatase (EC 3.6.1.1)	557	3	-	Yes	-	-
HPr kinase/phosphorylase (EC 2.7.1.-) (EC 2.7.4.-)	932	3	-	Yes	-	-
Ribonucleotide reductase of class Ib (aerobic), beta subunit (EC 1.17.4.1)	1019	3	-	Yes	-	-
Putative periplasmic phosphate-binding protein PstS (Mycoplasma type)	1157	3	-	Yes	-	-
LSU ribosomal protein L2p (L8e)	851	3	-	Yes	-	-
Translation elongation factor P	566	4	-	Yes	-	-
Phosphate transport system regulatory protein PhoU	686	4	-	Yes	-	-
ABC transporter, permease protein 2 (cluster 11, riboflavin/purine nucleoside/unknown)	932	4	-	Yes	-	-

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
LSU ribosomal protein L5p (L11e)	557	5	-	-	Yes	-
ATPase component of general energizing module of ECF transporters	839	5	-	-	Yes	-
ABC transporter, permease protein	977	5	-	-	Yes	-
5-formyltetrahydrofolate cyclo-ligase (EC 6.3.3.2)	554	6	-	-	Yes	-
SSU ribosomal protein S4p (S9e) @ SSU ribosomal protein S4p (S9e), zinc-independent	623	6	-	-	Yes	-
SSU ribosomal protein S3p (S3e)	791	6	-	-	-	-
Thioredoxin reductase (EC 1.8.1.9)	986	6	-	-	Yes	-
Replication-associated recombination protein RarA	1241	6	-	-	-	-
NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)	1028	7	-	-	Yes	-
Prolipoprotein diacylglyceryl transferase	1109	7	-	-	Yes	-
ABC transporter, ATP-binding protein	1493	7	-	-	Yes	-
DNA polymerase III beta subunit (EC 2.7.7.7)	1196	8	-	-	Yes	-
Lysyl-tRNA synthetase (class II) (EC 6.1.1.6)	1478	8	-	-	Yes	-
ABC transporter, ATP-binding protein (cluster 11, riboflavin/purine nucleoside/unknown) / ABC transporter, ATP-binding protein (cluster 11, riboflavin/purine nucleoside/unknown)	1658	8	-	-	Yes	-
LSU ribosomal protein L6p (L9e)	551	9	-	-	Yes	-
FMN-dependent NADH-azoreductase (EC 1.7.1.6)	593	9	-	-	Yes	-
SSU ribosomal protein S2p (SAe)	971	9	-	-	Yes	-
Nicotinate-nucleotide adenylyltransferase (EC 2.7.7.18)	1094	9	-	-	Yes	-
Uridine monophosphate kinase (EC 2.7.4.22)	710	10	-	-	Yes	-
Inner membrane protein translocase and chaperone YidC, long form	1151	10	-	-	Yes	-
Excinuclease ABC subunit B	2033	10	-	-	Yes	-
Maltose phosphorylase (EC 2.4.1.8) / Trehalose phosphorylase (EC 2.4.1.64)	590	11	-	-	-	Yes
DegV family protein	869	11	-	-	-	Yes

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
(E)-4-hydroxy-3-methylbut-2-enyl-diphosphate synthase (flavodoxin) (EC 1.17.7.3)	1079	11	-	-	-	Yes
ABC transporter, permease protein 1 (cluster 11, riboflavin/purine nucleoside/unknown)	1652	11	-	-	-	Yes
TsaC protein (YrdC domain) required for threonylcarbamoyladenosine t(6)A37 modification in tRNA	548	12	-	-	-	Yes
Translation initiation factor 3	593	12	-	-	-	Yes
ATP synthase B chain (EC 3.6.3.14)	596	12	-	-	-	Yes
S-adenosylmethionine synthetase (EC 2.5.1.6)	1124	12	-	-	-	Yes
Chaperone protein DnaK	1781	12	-	-	-	Yes
Nonheme iron-containing ferritin	515	13	-	-	-	Yes
Trk system potassium uptake protein TrkA	671	13	-	-	-	Yes
oxidoreductase of aldo/keto reductase family, subgroup 1	839	13	-	-	-	Yes
Phosphate transport ATP-binding protein PstB (TC 3.A.1.7.1)	977	13	-	-	-	Yes
Enolase (EC 4.2.1.11)	1433	13	-	-	-	Yes
Spermidine/putrescine import ABC transporter ATP-binding protein PotA (TC 3.A.1.11.1)	1505	13	-	-	-	Yes
Transketolase (EC 2.2.1.1)	1988	13	-	-	-	Yes
Glycerol uptake facilitator protein	731	14	-	-	-	Yes
ATP synthase F0 sector subunit a	824	14	-	-	-	Yes
Pyruvate dehydrogenase E1 component beta subunit (EC 1.2.4.1)	977	14	-	-	-	Yes
6-phosphofructokinase (EC 2.7.1.11)	998	14	-	-		Yes
Iron-sulfur cluster assembly ATPase protein SufC	806	15	-	-	-	Yes
WcaA	998	15	-	-	-	Yes
Heat-inducible transcription repressor HrcA	1049	15	-	-	-	Yes
Oligopeptide transport system permease protein OppB (TC 3.A.1.5.1)	1052	15	-	-	-	Yes
NADH oxidase (EC 1.6.99.3)	1418	15	-	-	-	Yes
Glutamyl-tRNA synthetase (EC 6.1.1.17) @ Glutamyl-tRNA(Gln) synthetase (EC 6.1.1.24)	1496	15	-	-	-	Yes

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Uncharacterized amino acid permease, GabP family	1760	15	-	-	-	-
Thymidylate kinase (EC 2.7.4.9)	632	16	-	-	-	Yes
Probable transcriptional regulatory protein YebC	704	16	-	-	-	Yes
DNA recombination and repair protein RecO	725	16	-	-	-	Yes
predicted protease	815	16	-	-	-	Yes
Ribonuclease HIII (EC 3.1.26.4)	995	16	-	-	-	Yes
tRNA pseudouridine(38-40) synthase (EC 5.4.99.12)	740	17	-	-	-	Yes
SSU rRNA (adenine(1518)-N(6)/adenine(1519)-N(6))-dimethyltransferase (EC 2.1.1.182)	806	17	-	-	-	Yes
Methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9) / Methylenetetrahydrofolate dehydrogenase (NADP+) (EC 1.5.1.5)	851	17	-	-	-	Yes
Translation elongation factor Ts	878	17	-	-	-	Yes
Pyruvate dehydrogenase E1 component alpha subunit (EC 1.2.4.1)	1082	17	-	-	-	Yes
Seryl-tRNA synthetase (EC 6.1.1.11)	1262	17	-	-	-	Yes
Phosphate transport system permease protein PstC (TC 3.A.1.7.1)	1958	17	-	-	-	Yes
Protein-N(5)-glutamine methyltransferase PrmC, methylates polypeptide chain release factors RF1 and RF2	854	18	-	-	-	Yes
Thymidylate synthase (EC 2.1.1.45)	869	18	-	-	-	Yes
N-Acetyl-D-glucosamine ABC transport system, permease protein	989	18	-	-	-	Yes
DEAD-box ATP-dependent RNA helicase CshB (EC 3.6.4.13)	1370	18	-	-	-	Yes
DNA gyrase subunit B (EC 5.99.1.3)	1946	18	-	-	-	Yes
Lipoate-protein ligase A	995	18	-	-	-	Yes
Glycerol-3-phosphate dehydrogenase (EC 1.1.5.3)	1190	19	-	-	-	Yes
Protein translocase subunit SecA	2672	19	-	-	-	Yes
Caffeoyl-CoA O-methyltransferase (EC 2.1.1.104)	569	20	-	-	-	Yes
Ribose-phosphate pyrophosphokinase (EC 2.7.6.1)	1022	20	-	-	-	Yes
Transmembrane component of general energizing module of ECF transporters	1169	20	-	-	-	Yes
Phosphoglycerate kinase (EC 2.7.2.3)	1241	20	-	-	-	Yes

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Cell division protein FtsH	2297	20	-	-	-	Yes
Methionine aminopeptidase (EC 3.4.11.18)	752	21	-	-	-	Yes
1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51)	797	21	-	-	-	Yes
GTP-binding protein Era	902	21	-	-	-	Yes
tRNA 4-thiouridine synthase (EC 2.8.1.4)	1190	21	-	-	-	Yes
Dihydrolipoamide dehydrogenase of pyruvate dehydrogenase complex (EC 1.8.1.4)	1391	21	-	-	-	Yes
Valyl-tRNA synthetase (EC 6.1.1.9)	2582	21	-	-	-	Yes
Adenylate kinase (EC 2.7.4.3)	644	22	-	-	-	Yes
Cell division trigger factor (EC 5.2.1.8)	1331	22	-	-	-	Yes
ATP synthase alpha chain (EC 3.6.3.14)	1556	22	-	-	-	Yes
SSU ribosomal protein S5p (S2e)	680	23	-	-	-	Yes
DNA polymerase III delta prime subunit (EC 2.7.7.7)	782	23	-	-	-	Yes
Uncharacterized metal-dependent hydrolase YcfH	800	23	-	-	-	Yes
dNTP triphosphohydrolase, broad substrate specificity	1343	23	-	-	-	Yes
Cell division protein FtsZ	1385	23	-	-	-	Yes
ATP synthase beta chain (EC 3.6.3.14)	1415	23	-	-	-	Yes
D-Ribose 1,5-phosphomutase (EC 5.4.2.7)	1631	23	-	-	-	Yes
Translation elongation factor G	2081	23	-	-	-	Yes
Uncharacterized protein YmdB	854	24	-	-	-	Yes
dnaJ-like protein	968	24	-	-	-	Yes
Glycerol kinase (EC 2.7.1.30)	1526	24	-	-	-	Yes
Maltodextrin ABC transporter, ATP-binding protein MsmX	1733	24	-	-	-	Yes
tRNA (guanine(46)-N(7))-methyltransferase (EC 2.1.1.33)	641	25	-	-	-	Yes
Endonuclease IV (EC 3.1.21.2)	911	25	-	-	-	Yes
Cytosol aminopeptidase PepA (EC 3.4.11.1)	1337	25	-	-	-	Yes
N(6)-L-threonylcarbamoyladenine synthase (EC 2.3.1.234)	962	26	-	-	-	Yes
Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)	1016	26	-	-	-	Yes

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Acetate kinase (EC 2.7.2.1)	1166	26	-	-	-	Yes
16S rRNA (cytidine(1402)-2'-O)-methyltransferase (EC 2.1.1.198) - rsmI	821	28	-	-	-	Yes
Heterodimeric efflux ABC transporter, permease/ATP-binding subunit 2	1901	28	-	-	-	Yes
1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.267)	1106	29	-	-	-	Yes
GTP-binding protein Obg	1304	31	-	-	-	Yes
Glyceraldehyde-3-phosphate dehydrogenase (NADP(+)) (EC 1.2.1.9)	1487	32	-	-	-	Yes
Ribonuclease J1 (endonuclease and 5' exonuclease)	1700	32	-	-	-	Yes
RecA protein	1055	33	-	-	-	Yes
Phosphatidate cytidylyltransferase (EC 2.7.7.41)	1127	33	-	-	-	Yes
Protein translocase subunit SecY	1496	33	-	-	-	Yes
Methionyl-tRNA synthetase (EC 6.1.1.10)	1763	33				Yes
3'-to-5' exoribonuclease RNase R	2174	33	-	-	-	Yes
dCMP deaminase (EC 3.5.4.12)	953	34	-	-	-	Yes
Arginine deiminase (EC 3.5.3.6)	1211	34	-	-	-	Yes
Protease2	1229	34	-	-	-	Yes
4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (EC 2.7.1.148)	725	35	-	-	-	Yes
Segregation and condensation protein B	875	35	-	-	-	Yes
Prolipoprotein	2375	35	-	-	-	Yes
Isoleucyl-tRNA synthetase (EC 6.1.1.5)	2693	35	-	-	-	Yes
tRNA-5-carboxymethylaminomethyl-2-thiouridine(34) synthesis protein MnmE	1346	36				Yes
Ribonuclease J2 (endoribonuclease in RNA processing)	1730	37	-	-	-	Yes
Glycyl-tRNA synthetase (EC 6.1.1.14)	1391	39	-	-	-	Yes
DNA polymerase IV (EC 2.7.7.7)	1196	40	-	-	-	Yes
KtrCD potassium uptake system, integral membrane component KtrD	1766	40	-	-	-	Yes
Oligopeptide transport ATP-binding protein OppD (TC 3.A.1.5.1)	1331	41	-	-	-	Yes
Pyruvate kinase (EC 2.7.1.40)	1526	42	-	-	-	Yes

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Acyl-phosphate:glycerol-3-phosphate O-acyltransferase PlsY (EC 2.3.1.n3)	728	43	-	-	-	Yes
16S rRNA (cytosine(1402)-N(4))-methyltransferase (EC 2.1.1.199)	938	43	-	-	-	Yes
2,3-bisphosphoglycerate-independent phosphoglycerate mutase (EC 5.4.2.12)	1511	43	-	-	-	Yes
Heat shock protein 60 family chaperone GroEL	1604	44	-	-	-	Yes
1-deoxy-D-xylulose 5-phosphate synthase (EC 2.2.1.7)	1745	45	-	-	-	Yes
Arginyl-tRNA synthetase (EC 6.1.1.19)	1649	46	-	-	-	Yes
ClpB protein	2153	46	-	-	-	Yes
Aldehyde dehydrogenase ALDH (EC 1.2.1.3)	1418	48	-	-	-	Yes
Phosphoenolpyruvate-protein phosphotransferase of PTS system (EC 2.7.3.9)	1736	48	-	-	-	Yes
Methionyl-tRNA formyltransferase (EC 2.1.2.9)	947	49	-	-	-	Yes
Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94)	989	49	-	-	-	Yes
Excinuclease ABC subunit C	1778	49	-	-	-	Yes
Aminopeptidase C (EC 3.4.22.40)	1640	50	-	-	-	Yes
Translation initiation factor 2	1847	50	-	-	-	Yes
Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.12)	1331	53	-	-	-	Yes
Topoisomerase IV subunit A (EC 5.99.1.-)	2390	54	-	-	-	Yes
DNA primase (EC 2.7.7.-)	1988	58	-	-	-	Yes
DNA gyrase subunit A (EC 5.99.1.3)	2540	58	-	-	-	Yes
Transcription termination protein NusA	1874	59	-	-	-	Yes
predicted integral membrane protein	2030	59	-	-	-	Yes
ATP-dependent protease La (EC 3.4.21.53) Type I	2438	60	-	-	-	Yes
RNA polymerase sigma factor RpoD	1910	62	-	-	-	Yes
ATP-dependent DNA helicase UvrD/PcrA (EC 3.6.4.12)	2255	62	-	-	-	Yes
4-hydroxybutyrate:acetyl-CoA CoA transferase (EC 2.3.1.-)	1316	63	-	-	-	Yes
Ribonuclease III (EC 3.1.26.3)	2132	64	-	-	-	Yes

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Excinuclease ABC subunit A	2858	65	-	-	-	Yes
DNA ligase (NAD(+)) (EC 6.5.1.2)67	2147	67	-	-	-	Yes
Chaperone protein DnaJ	1133	70	-	-	-	Yes
Siderophore-mediated iron transport protein	2027	73	-	-	-	Yes
Oligopeptide transport ATP-binding protein OppF (TC 3.A.1.5.1)	2495	79	-	-	-	Yes
Leucyl-tRNA synthetase (EC 6.1.1.4)	2417	79	-	-	-	Yes
Alanyl-tRNA synthetase (EC 6.1.1.7)	2708	82	-	-	-	Yes
Putative tRNA-m1A22 methylase	680	87	-	-	-	Yes
Deoxyribose-phosphate aldolase (EC 4.1.2.4)	671	89	-	-	-	Yes
tRNA-specific 2-thiouridylase MnmA (EC 2.8.1.13)	1973	89	-	-	-	Yes
Phosphate:acyl-ACP acyltransferase PlsX (EC 2.3.1.n2)	962	107	-	-	-	Yes
proline rich protein 2	1157	344	-	-	-	Yes
Lipid A export ATP-binding/permease protein MsbA	1787	200+	-	-	-	Yes
Oligopeptide transport system permease protein OppC (TC 3.A.1.5.1)	1076	400+	-	-	-	Yes
Deoxyadenosine kinase (EC 2.7.1.76) / Deoxyguanosine kinase (EC 2.7.1.113)	626	Incomplete data, possible sequencing error				
PTS system, mannitol-specific IIB component / PTS system, mannitol-specific IIC component	764	Incomplete data, possible sequencing error				
protease1	953	Incomplete data, possible sequencing error				
Mobile element protein	1271	Incomplete data, possible sequencing error				
Heterodimeric efflux ABC transporter, permease/ATP-binding subunit 1	1760	Incomplete data, possible sequencing error				

Appendix B2: Completed alignment and sequence comparison - *M. synoviae*

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Inorganic pyrophosphatase (EC 3.6.1.1)	545		Yes	-	-	-
Guanylate kinase (EC 2.7.4.8)	596	0	Yes	-	-	-
Phenylalanyl-tRNA synthetase domain protein (Bsu YtpR)	599	0	Yes	-	-	-
ATP synthase F0 sector subunit a (EC 3.6.3.14)	803	0	Yes	-	-	-
3'-to-5' oligoribonuclease A, Bacillus type	968	0	Yes	-	-	-
NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)	1004	0	Yes	-	-	-
Cell division protein FtsZ	1700	0	Yes	-	-	-
Adenine phosphoribosyltransferase (EC 2.4.2.7)	509	1	-	Yes	-	-
ATP synthase delta chain (EC 3.6.3.14)	551	1	-	Yes	-	-
Hypoxanthine-guanine phosphoribosyltransferase (EC 2.4.2.8)	554	1	-	Yes	-	-
Ribosome recycling factor	554	1	-	Yes	-	-
Single-stranded DNA-binding protein	566	1	-	Yes	-	-
Transcription antitermination protein NusG	608	1	-	Yes	-	-
Uracil phosphoribosyltransferase (EC 2.4.2.9)	620	1	-	Yes	-	-
Thymidylate kinase (EC 2.7.4.9)	644	1	-	Yes	-	-
Cytidylate kinase (EC 2.7.4.25)	656	1	-	Yes	-	-
SSU ribosomal protein S3p (S3e)	680	1	-	Yes	-	-
ABC transporter ATP-binding protein, associated with thiamin (pyrophosphate?) binding lipoprotein p37	770	1	-	Yes	-	-
Methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9) / Methylenetetrahydrofolate dehydrogenase (NADP+) (EC 1.5.1.5)	836	1	-	Yes	-	-
DegV family protein	893	1	-	Yes	-	-
dCMP deaminase (EC 3.5.4.12)	506	2	-	Yes	-	-
tRNA (cytidine(34)-2'-O)-methyltransferase (EC 2.1.1.207)	536	2	-	Yes	-	-

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Translation elongation factor P	560	2	-	Yes	-	-
Deoxyadenosine kinase (EC 2.7.1.76) @ Deoxyguanosine kinase (EC 2.7.1.113)	674	2	-	Yes	-	-
16S rRNA (guanine(527)-N(7))-methyltransferase (EC 2.1.1.170)	677	2	-	Yes	-	-
tRNA (guanine(37)-N(1))-methyltransferase (EC 2.1.1.228)	683	2	-	Yes	-	-
LSU ribosomal protein L1p (L10Ae)	695	2	-	Yes	-	-
SMF family protein, DNA processing chain A (DprA)	740	2	-	Yes	-	-
tRNA(1)(Val) (adenine(37)-N(6))-methyltransferase (EC 2.1.1.223)	791	2	-	Yes	-	-
ATPase component of general energizing module of ECF transporters	800	2	-	Yes	-	-
Acetyltransferase, GNAT family	821	2	-	Yes	-	-
Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.12)	884	2	-	Yes	-	-
Heat shock protein GrpE	890	2	-	Yes	-	-
Thioredoxin reductase (EC 1.8.1.9)	929	2	-	Yes	-	-
Predicted sialic acid transporter	1847	2	-	Yes	-	-
Dephospho-CoA kinase (EC 2.7.1.24)	506	3	-	Yes	-	-
LSU ribosomal protein L5p (L11e)	560	3	-	Yes	-	-
GTP-binding protein EngB	563	3	-	Yes	-	-
Segregation and condensation protein B	569	3	-	Yes	-	-
ATP synthase F0 sector subunit b (EC 3.6.3.14)	581	3	-	Yes	-	-
Holliday junction ATP-dependent DNA helicase RuvA (EC 3.6.4.12)	596	3	-	Yes	-	-
tRNA (guanine(46)-N(7))-methyltransferase (EC 2.1.1.33)	617	3	-	Yes	-	-
Adenylate kinase (EC 2.7.4.3)	635	3	-	Yes	-	-
LemA PROTEIN	653	3	-	Yes	-	-
Transcriptional regulator, GntR family	662	3	-	Yes	-	-
DNA recombination and repair protein RecO	686	3	-	Yes	-	-
KtrCD potassium uptake system, peripheral membrane component KtrC	695	3	-	Yes	-	-

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
16S rRNA (cytidine(1402)-2'-O)-methyltransferase (EC 2.1.1.198)	698	3	-	Yes	-	-
23S rRNA (guanosine(2251)-2'-O)-methyltransferase (EC 2.1.1.185)	701	3	-	Yes	-	-
Peptide chain release factor N(5)-glutamine methyltransferase (EC 2.1.1.297)	725	3	-	Yes	-	-
Segregation and condensation protein A	743	3	-	Yes	-	-
Protein serine/threonine phosphatase PrpC, regulation of stationary phase	746	3	-	Yes	-	-
Ribosome small subunit biogenesis RbfA-release protein RsgA	830	3	-	Yes	-	-
Endonuclease IV (EC 3.1.21.2)	833	3	-	Yes	-	-
FMN adenylyltransferase (EC 2.7.7.2) / Riboflavin kinase (EC 2.7.1.26)	836	3	-	Yes	-	-
Methionyl-tRNA formyltransferase (EC 2.1.2.9)	842	3	-	Yes	-	-
LSU ribosomal protein L2p (L8e)	845	3	-	Yes	-	-
Multiple sugar ABC transporter, membrane-spanning permease protein MsmG	857	3	-	Yes	-	-
Fructose-bisphosphate aldolase class II (EC 4.1.2.13)	866	3	-	Yes	-	-
SSU ribosomal protein S2p (SAe)	890	3	-	Yes	-	-
Efflux ABC transporter, ATP-binding protein	938	3	-	Yes	-	-
Glycosyltransferase	1019	3	-	Yes	-	-
NADH oxidase (EC 1.6.99.3)	1376	3	-	Yes	-	-
Ribonuclease J1 (endonuclease and 5' exonuclease)	1889	3	-	Yes	-	-
Dimeric dUTPase (EC 3.6.1.23)	500	4	-	Yes	-	-
LSU ribosomal protein L6p (L9e)	542	4	-	Yes	-	-
Thymidine kinase (EC 2.7.1.21)	551	4	-	Yes	-	-
Uracil-DNA glycosylase, family 1 (EC 3.2.2.27)	674	4	-	Yes	-	-
LSU ribosomal protein L4p (L1e)	689	4	-	Yes	-	-
Ribosomal large subunit pseudouridine synthase B (EC 5.4.99.22)	707	4	-	Yes	-	-
Probable transcriptional regulatory protein YebC	722	4	-	Yes	-	-

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Acyl-CoA:1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51)	758	4	-	Yes	-	-
Transmembrane component of general energizing module of ECF transporters	881	4	-	Yes	-	-
N-acetylneuraminate lyase (EC 4.1.3.3)	890	4	-	Yes	-	-
DNA polymerase III delta prime subunit (EC 2.7.7.7)	917	4	-	Yes	-	-
no hits	917	4	-	Yes	-	-
Phosphate acetyltransferase (EC 2.3.1.8)	959	4	-	Yes	-	-
Peptide chain release factor 1	1079	4	-	Yes	-	-
tRNA-specific 2-thiouridylase MnmA (EC 2.8.1.13)	1118	4	-	Yes	-	-
Oligopeptide transport system permease protein OppB (TC 3.A.1.5.1)	1124	4	-	Yes	-	-
Cysteinyl-tRNA synthetase (EC 6.1.1.16)	1220	4	-	Yes	-	-
Oligopeptide transport system permease protein OppC (TC 3.A.1.5.1)	1343	4	-	Yes	-	-
Glutamyl-tRNA synthetase (EC 6.1.1.17) @ Glutamyl-tRNA(Gln) synthetase (EC 6.1.1.24)	1400	4	-	Yes	-	-
Arginyl-tRNA synthetase (EC 6.1.1.19)	1655	4	-	Yes	-	-
LSU ribosomal protein L11p (L12e)	581	5	-	-	Yes	-
Substrate-specific component PanT of predicted pantothenate ECF transporter	584	5	-	-	Yes	-
Cysteinyl-tRNA synthetase related protein	647	5	-	-	Yes	-
Ribulose-phosphate 3-epimerase (EC 5.1.3.1)	671	5	-	-	Yes	-
Ribonuclease HIII (EC 3.1.26.4)	728	5	-	-	Yes	-
LSU ribosomal protein L3p (L3e)	806	5	-	-	Yes	-
tRNA pseudouridine(55) synthase (EC 5.4.99.25)	863	5	-	-	Yes	-
N-acetylmannosamine kinase (EC 2.7.1.60)	872	5	-	-	Yes	-
Uncharacterized lipoprotein MYPU_1930	905	5	-	-	Yes	-
Hydrolase, HAD superfamily	929	5	-	-	Yes	-
Serine/threonine protein kinase PrkC, regulator of stationary phase, short form	998	5	-	-	Yes	-

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
DNA-directed RNA polymerase alpha subunit (EC 2.7.7.6)	1001	5	-	-	Yes	-
Heat-inducible transcription repressor HrcA	1022	5	-	-	Yes	-
DNA polymerase III beta subunit (EC 2.7.7.7)	1133	5	-	-	Yes	-
Translation elongation factor Tu	1184	5	-	-	Yes	-
Acetate kinase (EC 2.7.2.1)	1193	5	-	-	Yes	-
Asparaginyl-tRNA synthetase (EC 6.1.1.22)	1343	5	-	-	Yes	-
Putative Dihydrolipoamide dehydrogenase (EC 1.8.1.4); Mercuric ion reductase (EC 1.16.1.1); PF00070 family, FAD-dependent NAD(P)-disulphide oxidoreductase	1379	5	-	-	Yes	-
Cytosol aminopeptidase PepA (EC 3.4.11.1)	1382	5	-	-	Yes	-
ATP synthase beta chain-like protein (EC 3.6.3.14)	1412	5	-	-	Yes	-
Transcription accessory protein (S1 RNA-binding domain)	2117	5	-	-	Yes	-
Large-conductance mechanosensitive channel	521	6	-	-	Yes	-
Peptidyl-tRNA hydrolase (EC 3.1.1.29)	563	6	-	-	Yes	-
Recombination protein RecR	584	6	-	-	Yes	-
Lipoprotein signal peptidase (EC 3.4.23.36)	719	6	-	-	Yes	-
SSU ribosomal protein S5p (S2e)	746	6	-	-	Yes	-
Methionine aminopeptidase (EC 3.4.11.18)	746	6	-	-	Yes	-
Uncharacterized protein YmdB	818	6	-	-	Yes	-
Ribosomal large subunit pseudouridine synthase C (EC 5.4.99.24)	845	6	-	-	Yes	-
LSU ribosomal maturation GTPase RbgA (B. subtilis YlqF)	857	6	-	-	Yes	-
ATP synthase gamma chain (EC 3.6.3.14)	869	6	-	-	Yes	-
N(6)-L-threonylcarbamoyladenine synthase (EC 2.3.1.234)	923	6	-	-	Yes	-
6-phosphofructokinase (EC 2.7.1.11)	977	6	-	-	Yes	-
Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94)	992	6	-	-	Yes	-
Prolipoprotein diacylglyceryl transferase	1007	6	-	-	Yes	-
Aminopeptidase YpdF (MP-, MA-, MS-, AP-, NP- specific)	1070	6	-	-	Yes	-

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Nicotinate-nucleotide adenylyltransferase (EC 2.7.7.18) / Hydrolase (HAD superfamily), YqeK	1085	6	-	-	Yes	-
GTP-binding and nucleic acid-binding protein YchF	1097	6	-	-	Yes	-
Phosphopentomutase (EC 5.4.2.7)	1163	6	-	-	Yes	-
2,3-bisphosphoglycerate-independent phosphoglycerate mutase (EC 5.4.2.12)	1505	6	-	-	Yes	-
ATP synthase alpha chain-like protein (EC 3.6.3.14)	1544	6	-	-	Yes	-
Phosphoenolpyruvate-protein phosphotransferase of PTS system (EC 2.7.3.9)	1706	6	-	-	Yes	-
Transport system permease protein, associated with thiamin (pyrophosphate?) binding lipoprotein p37	1706	6	-	-	Yes	-
ABC transporter, ATP-binding protein (cluster 11, riboflavin/purine nucleoside/unknown) / ABC transporter, ATP-binding protein (cluster 11, riboflavin/purine nucleoside/unknown)	1811	6	-	-	Yes	-
Ribonucleotide reductase of class Ib (aerobic), alpha subunit (EC 1.17.4.1)	2174	6	-	-	Yes	-
Phospholipid-binding protein	539	7	-	-	Yes	-
16S rRNA (guanine(966)-N(2))-methyltransferase (EC 2.1.1.171)	542	7	-	-	Yes	-
Pantothenate kinase type III, CoaX-like (EC 2.7.1.33)	752	7	-	-	Yes	-
Similar to citrate lyase beta chain, 5	833	7	-	-	Yes	-
Sporulation transcription regulator WhiA	854	7	-	-	Yes	-
Multiple sugar ABC transporter, membrane-spanning permease protein MsmF	854	7	-	-	Yes	-
Lipoate-protein ligase A	980	7	-	-	Yes	-
Tryptophanyl-tRNA synthetase (EC 6.1.1.2)	992	7	-	-	Yes	-
Phosphoglycerate kinase (EC 2.7.2.3)	1196	7	-	-	Yes	-
Chromosomal replication initiator protein DnaA	1370	7	-	-	Yes	-
Chaperone protein DnaK	1790	7	-	-	Yes	-
Oligoendopeptidase F (EC 3.4.24.-)	1835	7	-	-	Yes	-
Heterodimeric efflux ABC transporter, permease/ATP-binding subunit 2	1946	7	-	-	Yes	-

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
DNA ligase (NAD(+)) (EC 6.5.1.2)	2093	7	-	-	Yes	-
tRNA threonylcarbamoyladenosine biosynthesis protein TsaB	569	8	-	-	Yes	-
Peptide deformylase (EC 3.5.1.88)	572	8	-	-	Yes	-
Acyl-phosphate:glycerol-3-phosphate O-acyltransferase PlsY (EC 2.3.1.n3)	674	8	-	-	Yes	-
N-acetylmannosamine-6-phosphate 2-epimerase (EC 5.1.3.9)	695	8	-	-	Yes	-
COF family HAD hydrolase protein	818	8	-	-	Yes	-
Thymidylate synthase (EC 2.1.1.45)	869	8	-	-	Yes	-
GTP-binding protein Era	887	8	-	-	Yes	-
DNA polymerase III delta subunit (EC 2.7.7.7)	932	8	-	-	Yes	-
Ribose-phosphate pyrophosphokinase (EC 2.7.6.1)	1004	8	-	-	Yes	-
Chaperone protein DnaJ	1118	8	-	-	Yes	-
Serine hydroxymethyltransferase (EC 2.1.2.1)	1265	8	-	-	Yes	-
Aspartyl-tRNA synthetase (EC 6.1.1.12)	1700	8	-	-	Yes	-
DNA polymerase III subunits gamma and tau (EC 2.7.7.7)	1829	8	-	-	Yes	-
Cell division protein FtsH	1994	8	-	-	Yes	-
Ribosomal small subunit pseudouridine synthase A (EC 5.4.99.19)	710	9	-	-	Yes	-
Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)	725	9	-	-	Yes	-
ABC transporter, permease protein 2 (cluster 11, riboflavin/purine nucleoside/unknown)	953	9	-	-	Yes	-
RecA protein	977	9	-	-	Yes	-
S-adenosylmethionine synthetase (EC 2.5.1.6)	1148	9	-	-	Yes	-
Extracytoplasmic thiamin (pyrophosphate?) binding lipoprotein p37, specific for Mycoplasma	1214	9	-	-	Yes	-
GTP-binding protein Obg	1274	9	-	-	Yes	-
ATP synthase beta chain (EC 3.6.3.14)	1394	9	-	-	Yes	-
ATP synthase alpha chain (EC 3.6.3.14)	1592	9	-	-	Yes	-
ABC transporter, permease component (Na+?)	1802	9	-	-	Yes	-
Transcription termination protein NusA	1817	9	-	-	Yes	-

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
ATP-dependent DNA helicase UvrD/PcrA (EC 3.6.4.12)	2204	9	-	-	Yes	-
16S rRNA (uracil(1498)-N(3))-methyltransferase (EC 2.1.1.193)	2294	9	-	-	Yes	-
Chromate transport protein	638	10	-	-	Yes	-
Thermonuclease family protein	719	10	-	-	Yes	-
Uncharacterized metal-dependent hydrolase YcfH	806	10	-	-	Yes	-
Signal recognition particle receptor FtsY	1052	10	-	-	Yes	-
Tyrosyl-tRNA synthetase (EC 6.1.1.1)	1235	10	-	-	Yes	-
Aspartyl-tRNA(Asn) amidotransferase subunit A (EC 6.3.5.6) @ Glutamyl-tRNA(Gln) amidotransferase subunit A (EC 6.3.5.7)	1319	10	-	-	Yes	-
Protein translocase subunit SecY	1433	10	-	-	Yes	-
Pyruvate kinase (EC 2.7.1.40)	1433	10	-	-	Yes	-
Cyclic-di-AMP phosphodiesterase GdpP	2045	10	-	-	Yes	-
Phosphatidate cytidylyltransferase (EC 2.7.7.41)	1064	11	-	-	-	Yes
DNA recombination protein RmuC	1499	11	-	-	-	Yes
Ribonuclease J2 (endoribonuclease in RNA processing)	1673	11	-	-	-	Yes
DNA gyrase subunit B (EC 5.99.1.3)	1952	11	-	-	-	Yes
Translation elongation factor G	2090	11	-	-	-	Yes
Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)	941	12	-	-	-	Yes
POTASSIUM CHANNEL PROTEIN	1073	12	-	-	-	Yes
tRNA-5-carboxymethylaminomethyl-2-thiouridine(34) synthesis protein MnmE	1391	12	-	-	-	Yes
Aspartyl-tRNA(Asn) amidotransferase subunit B (EC 6.3.5.6) @ Glutamyl-tRNA(Gln) amidotransferase subunit B (EC 6.3.5.7)	1421	12	-	-	-	Yes
Lysyl-tRNA synthetase (class II) (EC 6.1.1.6)	1475	12	-	-	-	Yes
KtrCD potassium uptake system, integral membrane component KtrD	1643	12	-	-	-	Yes
Threonyl-tRNA synthetase (EC 6.1.1.3)	1742	12	-	-	-	Yes
RecD-like DNA helicase YrrC	2222	12	-	-	-	Yes
Phosphopantothenoylcysteine decarboxylase (EC 4.1.1.36) / Phosphopantothenoylcysteine synthetase (EC 6.3.2.5)	1154	13	-	-	-	Yes

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Enolase (EC 4.2.1.11)	1358	13	-	-	-	Yes
tRNA-5-carboxymethylaminomethyl-2-thiouridine(34) synthesis protein MnmG	1838	13	-	-	-	Yes
Dihydrolipoamide dehydrogenase of pyruvate dehydrogenase complex (EC 1.8.1.4)	1868	13	-	-	-	Yes
Replicative DNA helicase (DnaB) (EC 3.6.4.12)	1526	14	-	-	-	Yes
Excinuclease ABC subunit C	1748	14	-	-	-	Yes
Excinuclease ABC subunit B	2006	14	-	-	-	Yes
Cell division trigger factor (EC 5.2.1.8)	1388	15	-	-	-	Yes
Histidyl-tRNA synthetase (EC 6.1.1.21)	1496	15	-	-	-	Yes
Maltose phosphorylase (EC 2.4.1.8) / Trehalose phosphorylase (EC 2.4.1.64)	2372	15	-	-	-	Yes
Maltodextrin ABC transporter, ATP-binding protein MsmX	1742	17	-	-	-	Yes
Maltose maltodextrin transport ATP-binding protein malK	2075	17	-	-	-	Yes
Leucyl-tRNA synthetase (EC 6.1.1.4)	2297	17	-	-	-	-
Ribonucleotide reductase of class Ib (aerobic), beta subunit (EC 1.17.4.1)	1019	18	-	-	-	Yes
N-acetylglucosamine-6-phosphate deacetylase (EC 3.5.1.25)	1145	18	-	-	-	Yes
CRISPR-associated protein Cas1	887	19	-	-	-	Yes
Heterodimeric efflux ABC transporter, permease/ATP-binding subunit 1	1778	19	-	-	-	Yes
Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)	2180	19	-	-	-	Yes
Uncharacterized amino acid permease, GabP family	1718	20	-	-	-	Yes
Topoisomerase IV subunit B (EC 5.99.1.-)	1925	20	-	-	-	Yes
tRNA(Ile)-lysidine synthetase (EC 6.3.4.19)	875	23	-	-	-	Yes
Ribosomal large subunit pseudouridine synthase D (EC 5.4.99.23)	911	25	-	-	-	Yes
Triacylglycerol lipase (EC 3.1.1.3)	788	27	-	-	-	Yes
Oligopeptide transport ATP-binding protein OppD (TC 3.A.1.5.1)	1322	29	-	-	-	Yes
Beta-phosphoglucomutase (EC 5.4.2.6)	677	33	-	-	-	Yes

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Uridine monophosphate kinase (EC 2.7.4.22)	719	33	-	-	-	Yes
ABC transporter, substrate-binding protein (cluster 11, riboflavin/purine nucleoside/unknown)	1343	36	-	-	-	Yes
Signal recognition particle protein Ffh	1391	37	-	-	-	Yes
Endo-1,4-beta-glucanase	1082	46	-	-	-	Yes
Hemolysin C	1319	49	-	-	-	Yes
Oligopeptide transport ATP-binding protein OppF (TC 3.A.1.5.1)	1757	51	-	-	-	Yes
M. homonis p80-related protein	2201	52	-	-	-	Yes
Holliday junction ATP-dependent DNA helicase RuvB (EC 3.6.4.12)	992	55	-	-	-	Yes
DNA polymerase I 5'-3' exonuclease domain	872	58	-	-	-	Yes
Ribonuclease III (EC 3.1.26.3)	695	59	-	-	-	Yes
Glucosamine-6-phosphate deaminase (EC 3.5.99.6)	758	62	-	-	-	Yes
Dihydroxyacetone kinase-like protein, phosphatase domain / Dihydroxyacetone kinase-like protein, kinase domain	1652	65	-	-	-	Yes
Mobile element protein	557	70	-	-	-	Yes
Phosphate:acyl-ACP acyltransferase PlsX (EC 2.3.1.n2)	977	70	-	-	-	Yes
Pyruvate dehydrogenase E1 component alpha subunit (EC 1.2.4.1)	1124	72	-	-	-	Yes
Translation initiation factor 2	1808	74	-	-	-	Yes
Trehalose-6-phosphate hydrolase (EC 3.2.1.93)	1643	92	-	-	-	Yes
FIG011178: rRNA methylase	719	100+	-	-	-	Yes
ABC transporter, permease protein	809	100+	-	-	-	Yes
Cysteine desulfurase (EC 2.8.1.7) > SufS	1163	100+	-	-	-	Yes
DNA polymerase IV (EC 2.7.7.7)	1241	100+	-	-	-	Yes
4-hydroxybutyrate:acetyl-CoA CoA transferase (EC 2.3.1.-)	1334	100+	-	-	-	Yes
23S rRNA (uracil(1939)-C(5))-methyltransferase (EC 2.1.1.190)	1340	100+	-	-	-	Yes
Aldehyde dehydrogenase	1442	100+	-	-	-	Yes
Cardiolipin synthetase (EC 2.7.8.-)	1514	100+	-	-	-	Yes
Methionyl-tRNA synthetase (EC 6.1.1.10)	1556	100+	-	-	-	Yes

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
ABC transporter, permease protein 1 (cluster 11, riboflavin/purine nucleoside/unknown)	1610	100+	-	-	-	Yes
predicted coding region	1646	100+	-	-	-	Yes
Type III restriction-modification system methylation subunit (EC 2.1.1.72)	1712	100+	-	-	-	Yes
DNA topoisomerase I (EC 5.99.1.2)	1877	100+	-	-	-	Yes
Transketolase (EC 2.2.1.1)	1940	100+	-	-	-	Yes
Inner membrane protein translocase and chaperone YidC, long form	1958	100+	-	-	-	Yes
Valyl-tRNA synthetase (EC 6.1.1.9)	2498	100+	-	-	-	Yes
SSU rRNA (adenine(1518)-N(6)/adenine(1519)-N(6))-dimethyltransferase (EC 2.1.1.182)	779	200+	-	-	-	Yes
Pyruvate dehydrogenase E1 component beta subunit (EC 1.2.4.1)	992	200+	-	-	-	Yes
ABC transporter, ATP-binding protein	1178	200+	-	-	-	Yes
GTP-binding protein EngA	1307	200+	-	-	-	Yes
3'-to-5' exoribonuclease RNase R	2207	200+	-	-	-	Yes
ABC transporter, ATP-binding and permease protein	2216	200+	-	-	-	Yes

Appendix B3: Completed alignment and sequence comparison - *M. gallinaceum*

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Translation initiation factor 3	536		-	Yes	-	-
LSU ribosomal protein L11p (L12e)	584	15	-	-	-	Yes
16S rRNA (cytosine(1402)-N(4))-methyltransferase (EC 2.1.1.199)	902	18	-	-	-	Yes
tRNA (guanine(37)-N(1))-methyltransferase (EC 2.1.1.228)	683	22	-	-	-	Yes
DNA-directed RNA polymerase alpha subunit (EC 2.7.7.6)	1010	22	-	-	-	Yes
LSU ribosomal protein L1p (L10Ae)	695	23	-	-	-	Yes
Alcohol dehydrogenase (EC 1.1.1.1)	1049	26	-	-	-	Yes
Uracil phosphoribosyltransferase (EC 2.4.2.9)	644	27	-	-	-	Yes
Adenine phosphoribosyltransferase (EC 2.4.2.7)	512	31	-	-	-	Yes
Methionine aminopeptidase (EC 3.4.11.18)	749	31	-	-	-	Yes
KtrCD potassium uptake system, peripheral membrane component KtrC	671	36	-	-	-	Yes
Oligopeptide transport system permease protein OppB (TC 3.A.1.5.1)	1055	38	-	-	-	Yes
Nicotinamidase (EC 3.5.1.19)	539	40	-	-	-	Yes
Glycyl-tRNA synthetase (EC 6.1.1.14)	1355	40	-	-	-	Yes
N-acetylmuramic acid 6-phosphate etherase (EC 4.2.1.126)	875	41	-	-	-	Yes
Spermidine/putrescine import ABC transporter permease protein PotC (TC 3.A.1.11.1)	791	42	-	-	-	Yes
Spermidine/putrescine import ABC transporter permease protein PotB (TC 3.A.1.11.1)	833	43	-	-	-	Yes
Thymidine kinase (EC 2.7.1.21)	551	44	-	-	-	Yes
Ribonuclease HIII (EC 3.1.26.4)	698	44	-	-	-	Yes
16S rRNA (uracil(1498)-N(3))-methyltransferase (EC 2.1.1.193)	683	45	-	-	-	Yes
Ribosomal large subunit pseudouridine synthase B (EC 5.4.99.22)	734	45	-	-	-	Yes
Fructose-bisphosphate aldolase class II (EC 4.1.2.13)	866	47	-	-	-	Yes

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Oligopeptide transport system permease protein OppC (TC 3.A.1.5.1)	1061	48	-	-	-	Yes
Oligopeptide transport ATP-binding protein OppD (TC 3.A.1.5.1)	1307	48	-	-	-	Yes
Tryptophanyl-tRNA synthetase (EC 6.1.1.2)	2459	48	-	-	-	Yes
3'-to-5' oligoribonuclease A, Bacillus type1	971	50	-	-	-	Yes
Peptide chain release factor 1	1064	50	-	-	-	Yes
Spermidine/putrescine import ABC transporter ATP-binding protein PotA (TC 3.A.1.11.1)	1391	50	-	-	-	Yes
PTS system, cellobiose-specific IIC component	1472	52	-	-	-	Yes
tRNA(1)(Val) (adenine(37)-N(6))-methyltransferase (EC 2.1.1.223)	773	53	-	-	-	Yes
Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23)	854	53	-	-	-	Yes
Hydrolase, HAD superfamily	905	54	-	-	-	Yes
LSU ribosomal maturation GTPase RbgA (B. subtilis YlqF)	857	56	-	-	-	Yes
Bis-ABC ATPase YbiT	1616	56	-	-	-	Yes
FIG011178: rRNA methylase	725	57	-	-	-	Yes
Transcriptional regulator	872	57	-	-	-	Yes
Aspartate--ammonia ligase (EC 6.3.1.1)	977	57	-	-	-	Yes
Outer surface protein of unknown function, cellobiose operon	1088	59	-	-	-	Yes
tRNA 4-thiouridine synthase (EC 2.8.1.4)	1142	60	-	-	-	Yes
Peptide chain release factor N(5)-glutamine methyltransferase (EC 2.1.1.297)	719	62	-	-	-	Yes
Endo-1,4-beta-glucanase	1079	65	-	-	-	Yes
Nicotinate-nucleotide adenylyltransferase (EC 2.7.7.18)	1100	65	-	-	-	Yes
S-adenosylmethionine synthetase (EC 2.5.1.6)	1136	66	-	-	-	Yes
Chaperone protein DnaK	1793	66	-	-	-	Yes
Replicative DNA helicase (DnaB) (EC 3.6.4.12)	1466	73	-	-	-	Yes
Anhydro-N-acetylmuramic acid kinase (EC 2.7.1.170)	1139	76	-	-	-	Yes
Phosphoglycerate kinase (EC 2.7.2.3)	1184	77	-	-	-	Yes
Translation initiation factor 2	1799	83	-	-	-	Yes

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
GTP-binding and nucleic acid-binding protein YchF	1100	84	-	-	-	Yes
Adenylate kinase (EC 2.7.4.3)	698	88	-	-	-	Yes
Methionyl-tRNA synthetase (EC 6.1.1.10)	1556	88	-	-	-	Yes
Ribonuclease J1 (endonuclease and 5' exonuclease)	1931	88	-	-	-	Yes
2,3-bisphosphoglycerate-independent phosphoglycerate mutase (EC 5.4.2.12)	1517	89	-	-	-	Yes
Cyclic-di-AMP phosphodiesterase GdpP	2000	89	-	-	-	Yes
Protein translocase subunit SecY	1418	91	-	-	-	Yes
Transcription termination protein NusA	1679	97	-	-	-	Yes
Cation-transporting ATPase, E1-E2 family	2744	131	-	-	-	Yes
Transcription antitermination protein NusG	599	100+	-	-	-	Yes
tRNA (cytidine(34)-2'-O)-methyltransferase (EC 2.1.1.207)	650	100+	-	-	-	Yes
Cysteinyl-tRNA synthetase (EC 6.1.1.16)	1217	100+	-	-	-	Yes
Arginyl-tRNA synthetase (EC 6.1.1.19)	1658	100+	-	-	-	Yes
Efflux ABC transporter, permease/ATP-binding protein YfiC	1841	100+	-	-	-	Yes
tRNA-5-carboxymethylaminomethyl-2-thiouridine(34) synthesis protein MnmG	1841	100+	-	-	-	Yes
Oligoendopeptidase F (EC 3.4.24.-)	1862	100+	-	-	-	Yes
Transketolase (EC 2.2.1.1)	1946	100+	-	-	-	Yes
DNA gyrase subunit B (EC 5.99.1.3)	1967	100+	-	-	-	Yes
DNA ligase (NAD(+)) (EC 6.5.1.2)	1979	100+	-	-	-	Yes
3'-to-5' exoribonuclease RNase R	2234	100+	-	-	-	Yes
Leucyl-tRNA synthetase (EC 6.1.1.4)	2330	100+	-	-	-	Yes
Protein translocase subunit SecA	2558	100+	-	-	-	Yes
ATP-dependent protease La (EC 3.4.21.53) Type I	2741	100+	-	-	-	Yes
Chromate transport protein	659	200+	-	-	-	Yes
amino acid permease	1502	200+	-	-	-	Yes
Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72)	1562	200+	-	-	-	Yes

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
DNA polymerase IV (EC 2.7.7.7)	1655	200+	-	-	-	Yes
Heterodimeric efflux ABC transporter, permease/ATP-binding subunit 1	1808	200+	-	-	-	Yes
PTS system, cellobiose-specific IIA component	872	300+	-	-	-	Yes
Cysteine desulfurase (EC 2.8.1.7) > SufS	1160	300+	-	-	-	Yes
Hemolysin C	1235	300+	-	-	-	Yes
Oligopeptide transport ATP-binding protein OppF (TC 3.A.1.5.1)	1361	300+	-	-	-	Yes
Probable spermidine/putrescine substrate binding protein in Mollicutes	1976	300+	-	-	-	Yes

Appendix B4: Completed alignment and sequence comparison - *M. pullorum*

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
2-dehydro-3-deoxy-D-gluconate 5-dehydrogenase (EC 1.1.1.127) @ 2-deoxy-D-gluconate 3-dehydrogenase (EC 1.1.1.125)	764		Yes	-	-	-
Deoxyadenosine kinase (EC 2.7.1.76) @ Deoxyguanosine kinase (EC 2.7.1.113)	659	0	Yes	-	-	-
DNA polymerase III delta prime subunit (EC 2.7.7.7)	935	0	Yes	-	-	-
putative TYPE II DNA MODIFICATION ENZYME (METHYLTRANSFERASE)	899	0	Yes	-	-	-
rRNA small subunit 7-methylguanosine (m7G) methyltransferase GidB	692	0	Yes	-	-	-
Single-stranded DNA-binding protein	662	0	Yes	-	-	-
SSU ribosomal protein S4p (S9e) @ SSU ribosomal protein S4p (S9e), zinc-independent	599	0	Yes	-	-	-
Translation elongation factor Tu	1187	0	Yes	-	-	-
CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase (EC 2.7.8.5)	722	1	-	Yes	-	-
Inorganic pyrophosphatase (EC 3.6.1.1)	551	1	-	Yes	-	-
phosphosugar-binding transcriptional regulator, RpiR family	872	1	-	Yes	-	-
RNA methyltransferase, TrmH family	728	1	-	Yes	-	-
Thymidine kinase (EC 2.7.1.21)	575	1	-	Yes	-	-
LSU ribosomal protein L4p (L1e)	632	2	-	Yes		-
N(6)-L-threonylcarbamoyladenine synthase (EC 2.3.1.234)	935	2	-	Yes	-	-
Nicotinamidase (EC 3.5.1.19)	542	2	-	Yes	-	-
SSU ribosomal protein S3p (S3e)	650	2	-	-	-	-
tRNA (cytidine(34)-2'-O)-methyltransferase (EC 2.1.1.207)	536	2	-	Yes	-	-
tRNA 4-thiouridine synthase (EC 2.8.1.4)	1139	2	-	Yes	-	-
Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23)	851	3	-	Yes	-	-

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Phosphoenolpyruvate-dihydroxyacetone phosphotransferase (EC 2.7.1.121), ADP-binding subunit DhaL	587	3	-	Yes	-	-
Ribonuclease III (EC 3.1.26.3)	710	3	-	Yes	-	-
4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase (EC 5.3.1.17)	830	4	-	Yes	-	-
LSU ribosomal protein L6p (L9e)	539	4	-	Yes	-	-
Peptide deformylase (EC 3.5.1.88)	548	4	-	Yes	-	-
PTS system, N-acetylglucosamine-specific IIA component (EC 2.7.1.69) / PTS system, N-acetylglucosamine-specific IIB component (EC 2.7.1.69) / PTS system, N-acetylglucosamine-specific IIC component (EC 2.7.1.69)	524	4	-	Yes	-	-
RNA polymerase sigma factor RpoD	1490	4	-	Yes	-	-
Serine/threonine protein kinase PrkC, regulator of stationary phase	998	4	-	Yes	-	-
Transcriptional regulator, GntR family	662	4	-	Yes	-	-
23S rRNA (guanosine(2251)-2'-O)-methyltransferase (EC 2.1.1.185)	704	5	-	-	Yes	-
Adenylate kinase (EC 2.7.4.3)	647	5	-	-	Yes	-
GTP-binding protein EngB	578	5	-	-	Yes	-
LSU ribosomal protein L2p (L8e)	845	5	-	-	Yes	-
LSU ribosomal protein L3p (L3e)	824	5	-	-	Yes	-
Prolyl-tRNA synthetase (EC 6.1.1.15), archaeal/eukaryal type	1436	5	-	-	Yes	-
Ribonuclease HII (EC 3.1.26.4)	617	5	-	Yes	-	-
SSU ribosomal protein S5p (S2e)	632	5	-	-	Yes	-
Thiol peroxidase, Tpx-type (EC 1.11.1.15)	500	5	-	-	Yes	-
Translation initiation factor 3	1841	5	-	-	Yes	-
Triosephosphate isomerase (EC 5.3.1.1)	734	5	-	-	Yes	-
tRNA (guanine(46)-N(7))-methyltransferase (EC 2.1.1.33)	611	5	-	-	-	-
tRNA threonylcarbamoyladenosine biosynthesis protein TsaB	560	5	-	-	Yes	-
Uracil phosphoribosyltransferase (EC 2.4.2.9)	626	5	-		Yes	-
Alcohol dehydrogenase (EC 1.1.1.1)	1046	6	-	-	Yes	-
D-lactate dehydrogenase (EC 1.1.1.28)	1037	6	-	-	Yes	-

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Guanylate kinase (EC 2.7.4.8)	599	6	-	-	Yes	-
LemA protein	665	6	-	-	Yes	-
LSU ribosomal protein L1p (L10Ae)	695	6	-	-	Yes	-
Thymidylate kinase (EC 2.7.4.9)	656	6	-	-	Yes	-
Asparaginyl-tRNA synthetase (EC 6.1.1.22)	1352	7	-	-	Yes	-
Chromosomal replication initiator protein DnaA	1484	7	-	-	Yes	-
Endo-1,4-beta-glucanase	1085	7	-	-	Yes	-
NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)	1001	7	-	-	Yes	-
Protein QmcA (possibly involved in integral membrane quality control)	920	7	-	-	Yes	-
Ribosome small subunit biogenesis RbfA-release protein RsgA	839	7	-	-	Yes	-
4-Hydroxy-2-oxoglutarate aldolase (EC 4.1.3.16) @ 2-dehydro-3-deoxyphosphogluconate aldolase (EC 4.1.2.14)	629	8	-	-	Yes	-
Aminopeptidase YpdF (MP-, MA-, MS-, AP-, NP- specific)	1046	8	-	-	Yes	-
Dihydroxyacetone kinase-like protein, phosphatase domain / Dihydroxyacetone kinase-like protein, kinase domain	1634	8	-	-	Yes	-
Fructose-bisphosphate aldolase class II (EC 4.1.2.13)	872	8	-	-	Yes	-
Histidinol-phosphatase (EC 3.1.3.15)	833	8	-	-	Yes	-
Recombination protein RecR	587	8	-	-	Yes	-
Ribosomal large subunit pseudouridine synthase B (EC 5.4.99.22)	725	8	-	-	Yes	-
16S rRNA (cytosine(1402)-N(4))-methyltransferase (EC 2.1.1.199)	905	9	-	-	Yes	-
Cysteinyl-tRNA synthetase (EC 6.1.1.16)	1220	9	-	-	Yes	-
Glycerol-3-phosphate dehydrogenase (EC 1.1.5.3)	1151	9	-	-	Yes	-
GLYCOSYLTRANSFERASE	1019	9	-	-	Yes	-
LSU ribosomal protein L5p (L11e)	554	9	-	-	Yes	-
NADH oxidase (EC 1.6.99.3)	1373	9	-	-	Yes	-
Phosphate acetyltransferase (EC 2.3.1.8)	962	9	-	-	Yes	-
Protein translocase subunit SecY	1409	9	-	-	Yes	-

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Segregation and condensation protein A	731	9	-	-	Yes	-
tRNA-specific 2-thiouridylase MnmA (EC 2.8.1.13)	1118	9	-	-	Yes	-
ABC transporter, ATP-binding protein (cluster 11, riboflavin/purine nucleoside/unknown) / ABC transporter, ATP-binding protein (cluster 11, riboflavin/purine nucleoside/unknown)	1817	10	-	-	-	Yes
DNA-directed RNA polymerase alpha subunit (EC 2.7.7.6)	1004	10	-	-	Yes	-
Mannose-6-phosphate isomerase (EC 5.3.1.8)	851	10	-	-	Yes	-
Methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9) / Methylenetetrahydrofolate dehydrogenase (NADP+) (EC 1.5.1.5)	836	10	-	-	Yes	-
Methionine aminopeptidase (EC 3.4.11.18)	776	10	-	-	-	Yes
Phosphoenolpyruvate-dihydroxyacetone phosphotransferase (EC 2.7.1.121), dihydroxyacetone binding subunit DhaK	977	10	-	-	Yes	-
Probable transcriptional regulatory protein YebC	731	10	-	-	Yes	-
PTS system, hyaluronate-oligosaccharide-specific IID component	869	10	-	-	Yes	-
Ribonuclease J2 (endoribonuclease in RNA processing)	1649	10	-	-	Yes	-
Uncharacterized amino acid permease, GabP family	1661	10	-	-	Yes	-
Cysteine desulfurase (EC 2.8.1.7) > SufS	1154	11	-	-	-	Yes
Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94)	1004	11	-	-	-	Yes
LSU ribosomal protein L11p (L12e)	587	11	-	-	-	Yes
Pyruvate dehydrogenase E1 component alpha subunit (EC 1.2.4.1)	1127	11	-	-	-	Yes
Pyruvate dehydrogenase E1 component beta subunit (EC 1.2.4.1)	1001	11	-	-	-	Yes
Signal recognition particle receptor FtsY	998	11	-	-	-	Yes
Transmembrane component of general energizing module of ECF transporters	866	11	-	-	-	Yes
Aspartate aminotransferase (EC 2.6.1.1)	851	12	-	-	-	Yes
GTP cyclohydrolase 1 type 2 homolog YbgI	773	12	-	-	-	Yes
GTP-binding protein Era	872	12	-	-	-	Yes
Hemolysins and related proteins containing CBS domains	1232	12	-	-	-	Yes
Segregation and condensation protein B	587	12	-	-	-	Yes

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Spermidine/putrescine import ABC transporter ATP-binding protein PotA (TC 3.A.1.11.1)	1373	12	-	-	-	Yes
Transcription antitermination protein NusG	599	12	-	-	-	Yes
Acetate kinase (EC 2.7.2.1)	1187	13	-	-	-	Yes
Argininosuccinate synthase (EC 6.3.4.5)	1238	13	-	-	-	Yes
Cell division protein FtsZ	1271	13	-	-	-	-
chromate ion transporter (CHR) family, putative	674	13	-	-	-	Yes
Inner membrane protein translocase and chaperone YidC, long form	1997	13	-	-	-	Yes
RecA protein	992	13	-	-	-	Yes
S-adenosylmethionine synthetase (EC 2.5.1.6)	1127	13	-	-	-	Yes
Spermidine/putrescine import ABC transporter permease protein PotC (TC 3.A.1.11.1)	1706	13	-	-	-	Yes
Translation elongation factor LepA	1799	13	-	-	-	Yes
DNA recombination and repair protein RecO	680	14	-	-	-	Yes
Maltose/maltodextrin ABC transporter, permease protein MalG	959	14	-	-	-	Yes
Pyruvate kinase (EC 2.7.1.40)	1433	14	-	-	-	Yes
Transport protein SgaT, putative	1547	14	-	-	-	Yes
16S rRNA (cytidine(1402)-2'-O)-methyltransferase (EC 2.1.1.198)	698	15	-	-	-	Yes
16S rRNA (guanine(966)-N(2))-methyltransferase (EC 2.1.1.171)	557	15	-	-	-	Yes
3'-to-5' oligoribonuclease A, Bacillus type	971	15	-	-	-	Yes
Beta-phosphoglucomutase (EC 5.4.2.6)	677	15	-	-	-	Yes
FMN-dependent NADH-azoreductase (EC 1.7.1.6)	599	15	-	-	-	Yes
Glycerol kinase (EC 2.7.1.30)	1529	15	-	-	-	Yes
Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)	2174	15	-	-	-	Yes
tRNA pseudouridine(55) synthase (EC 5.4.99.25)	845	15	-	-	-	Yes
ABC transporter ATP-binding-Pr1	1541	16	-	-	-	Yes
ABC transporter, ATP-binding protein	704	16	-	-	-	Yes

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Aspartate--ammonia ligase (EC 6.3.1.1)	980	16	-	-	-	Yes
ATP-dependent DNA helicase UvrD/PcrA (EC 3.6.4.12)	2147	16	-	-	-	Yes
Cell division trigger factor (EC 5.2.1.8)	1295	16	-	-	-	Yes
Cytidylate kinase (EC 2.7.4.25)	674	16	-	-	-	Yes
Cytosol aminopeptidase PepA (EC 3.4.11.1)	1409	16	-	-	-	Yes
Heterodimeric efflux ABC transporter, permease/ATP-binding subunit 1	1790	16	-	-	-	Yes
Histidyl-tRNA synthetase (EC 6.1.1.21)	1391	16	-	-	-	Yes
Unsaturated chondroitin disaccharide hydrolase (EC 3.2.1.180)	1211	16	-	-	-	Yes
Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.12)	920	17	-	-	-	Yes
DNA polymerase I (EC 2.7.7.7)	863	17	-	-	-	Yes
DNA primase (EC 2.7.7.-)	2075	17	-	-	-	Yes
Phosphate:acyl-ACP acyltransferase PlsX (EC 2.3.1.n2)	995	17	-	-	-	Yes
predicted MutT-like hydrolases	560	17	-	-	-	Yes
Seryl-tRNA synthetase (EC 6.1.1.11)	1268	17	-	-	-	Yes
ATPase component of general energizing module of ECF transporters	797	18	-	-	-	Yes
GTP-binding protein Obg	1265	18	-	-	-	Yes
tRNA(1)(Val) (adenine(37)-N(6))-methyltransferase (EC 2.1.1.223)	782	18	-	-	-	Yes
Cardiolipin synthetase (EC 2.7.8.-)	1514	19	-	-	-	Yes
Putative Dihydrolipoamide dehydrogenase (EC 1.8.1.4); Mercuric ion reductase (EC 1.16.1.1); PF00070 family, FAD-dependent NAD(P)-disulphide oxidoreductase	1373	19	-	-	-	Yes
Signal recognition particle protein Ffh	1352	19	-	-	-	Yes
Threonine dehydrogenase and related Zn-dependent dehydrogenases	1070	19	-	-	-	Yes
Aspartyl-tRNA synthetase (EC 6.1.1.12)	1712	20	-	-	-	Yes
DNA polymerase III subunits gamma and tau (EC 2.7.7.7)	2132	20	-	-	-	Yes
NAD synthetase (EC 6.3.1.5)	800	20	-	-	-	Yes

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
Peptide chain release factor 1	1070	20	-	-	-	Yes
Phosphopentomutase (EC 5.4.2.7)	1187	20	-	-	-	Yes
2,3-bisphosphoglycerate-independent phosphoglycerate mutase (EC 5.4.2.12)	1505	21	-	-	-	Yes
Aspartyl-tRNA(Asn) amidotransferase subunit B (EC 6.3.5.6) @ Glutamyl-tRNA(Gln) amidotransferase subunit B (EC 6.3.5.7)	1430	21	-	-	-	Yes
Glucose-6-phosphate isomerase (EC 5.3.1.9)	1298	21	-	-	-	Yes
Glycyl-tRNA synthetase (EC 6.1.1.14)	1337	21	-	-	-	Yes
GTP-binding protein EngA	1304	21	-	-	-	Yes
Heterodimeric efflux ABC transporter, permease/ATP-binding subunit 2	1910	21	-	-	-	Yes
Ribose-phosphate pyrophosphokinase (EC 2.7.6.1)	977	22	-	-	-	Yes
Thioredoxin reductase (EC 1.8.1.9)	923	22	-	-	-	Yes
Uracil-DNA glycosylase, family 1 (EC 3.2.2.27)	671	22	-	-	-	Yes
Argininosuccinate lyase (EC 4.3.2.1)	1394	23	-	-	-	Yes
Peptide chain release factor N(5)-glutamine methyltransferase (EC 2.1.1.297)	725	23	-	-	-	Yes
fic family protein	1046	24	-	-	-	Yes
Endonuclease IV (EC 3.1.21.2)	827	25	-	-	-	Yes
Excinuclease ABC subunit B	1988	25	-	-	-	Yes
Replicative DNA helicase (DnaB) (EC 3.6.4.12)	1481	25	-	-	-	Yes
Serine hydroxymethyltransferase (EC 2.1.2.1)	1268	25	-	-	-	Yes
6-phosphofructokinase (EC 2.7.1.11)	977	27	-	-	-	Yes
DNA topoisomerase I (EC 5.99.1.2)	1964	27	-	-	-	Yes
Lipoprotein signal peptidase (EC 3.4.23.36)	668	27	-	-	-	Yes
Methionyl-tRNA synthetase (EC 6.1.1.10)	1535	27	-	-	-	Yes
Uncharacterized metal-dependent hydrolase YcfH	800	27	-	-	-	Yes
FIG007079: UPF0348 protein family	914	28	-	-	-	Yes
ATP synthase alpha chain (EC 3.6.3.14)	1529	29	-	-	-	Yes

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
ATPase	1259	29	-	-	-	Yes
Cyclic-di-AMP phosphodiesterase GdpP	1991	29	-	-	-	Yes
Phage protein	1211	29	-	-	-	Yes
Sporulation transcription regulator WhiA	842	31	-	-	-	Yes
ATP synthase beta chain-like protein (EC 3.6.3.14)	1364	32	-	-	-	Yes
Methionyl-tRNA formyltransferase (EC 2.1.2.9)	833	32	-	-	-	Yes
RecD-like DNA helicase YrrC	2156	32	-	-	-	Yes
SSU rRNA (adenine(1518)-N(6)/adenine(1519)-N(6))-dimethyltransferase (EC 2.1.1.182)	782	34	-	-	-	Yes
Valyl-tRNA synthetase (EC 6.1.1.9)	2510	34	-	-	-	Yes
Arginyl-tRNA synthetase (EC 6.1.1.19)	1649	35	-	-	-	Yes
Maltose phosphorylase (EC 2.4.1.8) / Trehalose phosphorylase (EC 2.4.1.64)	2366	36	-	-	-	Yes
Transketolase (EC 2.2.1.1)	1955	38	-	-	-	Yes
DNA polymerase IV (EC 2.7.7.7)	1253	39	-	-	-	Yes
Prolipoprotein diacylglyceryl transferase	992	40	-	-	-	Yes
Aspartyl-tRNA(Asn) amidotransferase subunit A (EC 6.3.5.6) @ Glutamyl-tRNA(Gln) amidotransferase subunit A (EC 6.3.5.7)	1328	42	-	-	-	Yes
Excinuclease ABC subunit C	1739	42	-	-	-	Yes
Transcription accessory protein (S1 RNA-binding domain)	2111	42	-	-	-	Yes
Chromate transport protein	584	47	-	-	-	Yes
Neutral endopeptidase O (EC 3.4.24.-)	1904	49	-	-	-	Yes
Dihydrolipoamide dehydrogenase of pyruvate dehydrogenase complex (EC 1.8.1.4)	2186	50	-	-	-	Yes
Xylulose-5-phosphate phosphoketolase (EC 4.1.2.9) @ Fructose-6-phosphate phosphoketolase (EC 4.1.2.22)	2378	51	-	-	-	Yes
LSU ribosomal maturation GTPase RbgA (B. subtilis YlqF)	902	52	-	-	-	Yes
Glutamyl-tRNA synthetase (EC 6.1.1.17) @ Glutamyl-tRNA(Gln) synthetase (EC 6.1.1.24)	1412	55	-	-	-	Yes
DNA ligase (NAD(+)) (EC 6.5.1.2)	1997	56	-	-	-	Yes

Gene	Max group size nucleus	Number of differences	Identical	< 5 differences	5 - 10 differences	> 10 differences
HPr kinase/phosphorylase (EC 2.7.1.-) (EC 2.7.4.-)	932	63	-	-	-	Yes
Uncharacterized protein YmdB	899	77	-	-	-	Yes
Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)	2180	102	-	-	-	Yes
3'-to-5' exoribonuclease RNase R	2027	198	-	-	-	Yes
Alpha-amylase (EC 3.2.1.1)	1838	200+	-	-	-	Yes
Ribonuclease J1 (endonuclease and 5' exonuclease)	1919	200+	-	-	-	Yes
Ribulose-phosphate 3-epimerase (EC 5.1.3.1)	671	200+	-	-	-	Yes
tRNA-5-carboxymethylaminomethyl-2-thiouridine(34) synthesis protein MnmG	2258	200+	-	-	-	Yes

Appendix C: Research Approvals - REC, AEC, and Section 20

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Faculty of Veterinary Science

Research Ethics Committee 10 July 2020

CONDITIONALLY APPROVAL

Ethics Reference No	**REC075-20**
Protocol Title	**Development of a multiplex real-time polymerase chain reaction to distinguish between Mycoplasma species found in South African poultry**
Principal Investigator	**Miss PP Wambulawaye**
Supervisors	**Dr A Beylefeld**

Dear Miss PP Wambulawaye,

We are pleased to inform you that your submission has been conditionally approved by the Faculty of Veterinary Sciences Research Ethics committee, subject to other relevant approvals.

Please note the following about your ethics approval:

1. Please use your reference number (REC075-20) on any documents or correspondence with the Research Ethics Committee regarding your research.
2. Please note that the Research Ethics Committee may ask further questions, seek additional information, require further modification, monitor the conduct of your research, or suspend or withdraw ethics approval.
3. Please note that ethical approval is granted for the duration of the research as stipulated in the original application for post graduate studies (e.g. Honours studies: 1 year, Masters studies: two years, and PhD studies: three years) and should be extended when the approval period lapses.
4. The digital archiving of data is a requirement of the University of Pretoria. The data should be accessible in the event of an enquiry or further analysis of the data.

Ethics approval is subject to the following:

1. The ethics approval is conditional on the research being conducted as stipulated by the details of all documents submitted to the Committee. In the event that a further need arises to change who the investigators are, the methods or any other aspect, such changes must be submitted as an Amendment for approval by the Committee.
2. **Applications using Animals:** FVS ethics recommendation does not imply that AEC approval is granted. The application has been pre-screened and recommended for review by the AEC. Research may not proceed until AEC approval is granted.

NOTES: Conditionally approved pending the following (and to ensure that rerouting to AEC is not delayed):
(i) Obtaining all other relevant approvals.
(ii) Upload of all permission letters to make use of samples collected: (a) UP Poultry Mycoplasma biobank and (b) use of samples received through clients of the Poultry Section in the Department of Production Animal Studies.

We wish you the best with your research.

Yours sincerely

PROF M. OOSTHUIZEN
Chairperson: Research Ethics Committee

100 YEARS OF VETERINARY EDUCATION

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Faculty of Veterinary Science
Animal Ethics Committee

6 October 2020

Approval Certificate
New Application

AEC Reference No.:	REC075-20
Title:	Development of a multiplex real-time polymerase chain reaction to distinguish between Mycoplasma species found in South African poultry
Researcher:	Miss PP Wambulawaye
Student's Supervisor:	Prof C Abolnik

Dear Miss PP Wambulawaye,

The **New Application** as supported by documents received between 2020-06-11 and 2020-10-02 for your research, was approved by the Animal Ethics Committee on its quorate meeting of 2020-10-02.

Please note the following about your ethics approval:

1. The use of species is approved:

Species	**Number**
Poultry (various sources)	204 (only swabs taken)
Samples Tracheal swabs	 204

2. Ethics Approval is valid for 1 year and needs to be renewed annually by 2021-10-06.

3. Please remember to use your protocol number (REC075-20) on any documents or correspondence with the AEC regarding your research.

4. Please note that the AEC may ask further questions, seek additional information, require further modification, monitor the conduct of your research, or suspend or withdraw ethics approval.

5. **All incidents** must be reported by the PI by email to Ms Marleze Rheeder (AEC Coordinator) within 3 days, and must be subsequently submitted electronically on the application system within 14 days.

6. As part of your approval, the committee requires that you record a **short video footage** of major animal procedures approved in your study. The committee may request them for monitoring purposes at any later point.

Ethics approval is subject to the following:

- The ethics approval is conditional on the research being conducted as stipulated by the details of all documents submitted to the Committee. In the event that a further need arises to change who the investigators are, the methods or any other aspect, such changes must be submitted as an Amendment for approval by the Committee.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Faculty of Veterinary Science
Animal Ethics Committee

11 October 2021

Approval Certificate
Annual Renewal
(EXT1)

AEC Reference No.:	REC075-20
Title:	Development of a multiplex real-time polymerase chain reaction to distinguish between Mycoplasma species found in South African poultry
Researcher:	Miss PP Wambulawaye
Student's Supervisor:	Prof C Abolnik

Dear Miss PP Wambulawaye,

The **Annual Renewal** as supported by documents received between 2021-08-27 and 2021-10-01 for your research, was approved by the Animal Ethics Committee on its quorate meeting of 2021-10-01.

Please note the following about your ethics approval:

1. The use of species is approved:

Species	Number Available
Poultry (Layer hens) various sources	204 (only swabs taken)
Samples Tracheal swabs	204

2. Ethics Approval is valid for 1 year and needs to be renewed annually by 2022-10-11.
3. Please remember to use your protocol number (REC075-20) on any documents or correspondence with the AEC regarding your research.
4. Please note that the AEC may ask further questions, seek additional information, require further modification, monitor the conduct of your research, or suspend or withdraw ethics approval.
5. **All incidents** must be reported by the PI by email to Ms Marleze Rheeder (AEC Coordinator) within 3 days, and must be subsequently submitted electronically on the application system within 14 days.
6. The committee also requests that you record major procedures undertaken during your study for own-archiving, using any available digital recording system that captures in adequate quality, as it may be required if the committee needs to evaluate a complaint. However, if the committee has monitored the procedure previously or if it is generally can be considered routine, such recording will not be required.

Ethics approval is subject to the following:

- The ethics approval is conditional on the research being conducted as stipulated by the details of all documents submitted to the Committee. In the event that a further need arises to change who the investigators are, the methods or any other aspect, such changes must be submitted as an Amendment for approval by the Committee.

agriculture, land reform
& rural development

Department:
Agriculture, Land Reform and Rural Development
REPUBLIC OF SOUTH AFRICA

Directorate Animal Health, Department of Agriculture, Land Reform and Rural Development Private Bag X138, Pretoria 0001
Enquiries: Mr Herry Gololo • Tel: +27 12 319 7532 • Fax: +27 12 319 7470 • E-mail: HerryG@dalrrd.gov.za
Reference: 12/11/1/1/8 (1546KL)

Ms. Pamela Wambulawaye
Senior Technologist
Faculty of Veterinary Science
University of Pretoria
Tel +27 (0)12 529 8529
Email: pamela.wambulawaye@up.ac.za

Dear Miss Wambulawaye

RE: PERMISSION TO DO RESEARCH IN TERMS OF SECTION 20 OF THE ANIMAL DISEASES ACT, 1984 (ACT NO 35 OF 1984)

Your application dated 8 June 2020 requesting permission under Section 20 of the Animal Disease Act, 1984 (Act No. 35 of 1984) to perform a research project or study, refers.

I am pleased to inform you that permission is hereby granted to perform the following study, with the following conditions:

Conditions:

1. This permission does not relieve the researcher of any responsibility which may be placed on him by any other act of the Republic of South Africa;
2. This permission is given upon finding the biosecurity of the research project as described to be acceptable to DAFF. This permission does not serve as any approval or endorsement by DAFF for the commercial use or registration of any diagnostic test for any purpose in South Africa;
3. The research project is approved as per the application form dated 8 June 2020 and the correspondence thereafter. Written permission from the Director: Animal Health must be obtained prior to any deviation from the conditions approved for this

research project under this Section 20 permit. Please apply in writing to HerryG@daff.gov.za;

4. If required, an application for an extension must be made by the responsible researcher at least one month prior to the expiry of this Section 20 permit. Please apply in writing to HerryG@dalrrd.gov.za;
5. Only the following may be used in this research project:
   5.1. The *Mycoplasma.spp* isolates as referred to in the sample list that was attached to the abovementioned Section 20 application, stored at the Bacteriology Laboratory at the Department of Veterinary Tropical Diseases
   5.2. Chicken tracheal swabs obtained from the Poultry Unit at the Department of Production Animal Sciences;
   5.3. Written permission from the Director: Animal Health must be obtained prior to any additional isolates or samples being used in the research project. Please apply in writing to HerryG@dalrrd.gov.za;
6. No live animals may be used in this research project under this Section 20 permit;
7. Any incidence or suspected incidence of a controlled or notifiable disease in terms of the Animal Diseases Act 1984 (Act no 35 of 84) must be reported immediately to the state veterinarian of the area;
8. All potentially infectious material utilised or generated during or by the research project is to be destroyed at completion of the research project with the exception of the following:
   8.1. *Mycoplasma spp* cultures may be stored under access control at the Poultry Research Building, Faculty of Veterinary Sciences, Onderstepoort after completion of the research project;
9. Stored samples may not be outsourced or used for further research without prior written approval from the Director: Animal Health
10. Only a waste disposal company registered to remove biohazardous waste may be used for the removal of all potentially infectious waste from the research project;
11. It is the responsibility of the researcher and relevant laboratory or facility managers to ensure that the human safety aspects of this research project are adequately addressed
12. Records must be kept for five years for auditing purposes;

Title of research/study Development of a multi-primer multiplex real-time polymerase chain reaction to distinguish between *Mycoplasma* species found in South African poultry

Researcher: Ms. Pamela Wambulawaye
Institution: Poultry Research Laboratory, Faculty of Veterinary Sciences, Onderstepoort;
Permit Expiry date: 31 December 2021
Our ref Number: 12/11/1/1/8 (1546KL)
Your ref: REC075-20

Kind regards,

DR. MPHO MAJA
DIRECTOR: ANIMAL HEALTH
Date:
2020 -07- 17

SUBJECT: *PERMISSION TO DO RESEARCH IN TERMS OF SECTION 20 OF THE ANIMAL DISEASES ACT, 1984 (ACT NO. 35 OF 1984) 12/11/1/1/8 (1546KL)*

agriculture, land reform & rural development
Department:
Agriculture, Land Reform and Rural Development
REPUBLIC OF SOUTH AFRICA

Directorate Animal Health, Department of Agriculture, Land Reform and Rural Development
Private Bag X138, Pretoria 0001

Enquiries: Ms Marna Laing • Tel: +27 12 319 7442 • Fax: +27 12 319 7470 • E-mail: MarnaL@dalrrd.gov.za
Reference: 12/11/1/1/8A (1546KL) (JD)

Ms Pamela Wambulawaye
Faculty of Veterinary Science, University of Pretoria, Onderstepoort, 0110
Email: nushib28@gmail.com; pamela.wambulawaye@up.ac.za

Dear Ms Wambulawaye,

RE: AMENDMENT OF SECTION 20 APPROVAL IN TERMS OF THE ANIMAL DISEASES ACT, 1984 (ACT NO 35 OF 1984)

Title of research project / study: "Development of a multiplex real-time PCR to distinguish between Mycoplasma species found in South African poultry"

Your application requesting an amendment of the Section 20 permit issued by the Director of Animal Health on 2020-07-07 for the study mentioned above refers. I am pleased to inform you that the amendment is hereby granted with the following conditions:

Conditions:

1. The validity of the section 20 approval is extended to 31 December 2023;
2. The following *Mycoplasma spp.* strains may be added to the research project: *Mycoplasma gallisepticum* strain NCTC 10115 and *Mycoplasma synovaie* strain NTCT 25204;
3. This amendment does not relieve the researcher of any of the other conditions as contained in the Section 20 permit issued on 2020-07-07 for this study;

4. Written permission from the Director of Animal Health must be obtained prior to any deviation from the conditions approved for this study under the Section 20 permit. Please apply in writing to MarnaL@dalrrd.gov.za

Kind regards,

pp.

Dr Mpho Maja
DIRECTOR: ANIMAL HEALTH
Date:
2022 -02- 2 5

SUBJECT: Amendment of Section 20 approval in terms of Section 20 of the Animal Diseases Act, 1984 (Act No 35 of 1984)