



Clouds and *Pasteurella multocida*

Final Project Report | OCTOBER 2020

A report for Australian Eggs Limited

by L. Omaleki, P. Blackall and C. Turni

© 2020 Australian Eggs Limited.

All rights reserved.

ISBN 978-1-920835-43-9

Project Title: Clouds and *Pasteurella multocida*

Australian Eggs Limited Project Number 1BS702UQ

The views expressed and the conclusions reached in this publication are those of the authors and not necessarily those of persons consulted. Australian Eggs Limited shall not be responsible in any way whatsoever to any person who relies in whole or in part on the contents of this report.

This publication is copyright. However, Australian Eggs Limited encourages wide dissemination of its research, providing that it is clearly acknowledged. For any other enquiries concerning reproduction, contact the Innovation Program Manager on 02 9409 6999.

Researcher/Author Contact Details

Name: Lida Omaleki, Patrick Blackall, Conny Turni

Address: QAAFI, The University of Queensland,
GPO Box 267, Brisbane, QLD 4001

Phone: 0406 411 712 (Lida Omaleki)
07 3255 4296 (Conny Turni)

Email: l.omaleki@uq.edu.au
p.blackall@uq.edu.au
c.turni1@uq.edu.au

In submitting this report, the researcher has agreed to Australian Eggs Limited publishing this material in its edited form.

Australian Eggs Limited Contact Details:

Australian Eggs Limited
A.B.N: 66 102 859 585
Suite 6.02, Level 6, 132 Arthur St
North Sydney NSW 2060

Phone: 02 9409 6999
Fax: 02 9954 3133
Email: research@australianeggs.org
Website: www.australianeggs.org.au

Published in October 2020

Foreword

This project was conducted to improve the ability of the industry both to prevent and control fowl cholera, an important disease that can cause significant mortalities, typically in laying birds. Importantly, fowl cholera, caused by the bacterium *Pasteurella multocida*, has emerged as a problem in free range and organic flocks in recent years.

The project developed, validated and applied genomic sequencing approaches and bioinformatic analysis to characterise *Pasteurella multocida* isolates. These tools allowed for an understanding of the infection dynamics of fowl cholera on the two free range layer farms. Importantly, the project generated knowledge that explains why the killed autogenous fowl cholera vaccines used on these farms could not provide protection against the disease. With the availability of this technology, far more specific and relevant genomic-based guidance can now be provided to the vaccine companies and veterinary professionals servicing the layer industry.

This project has produced a set of tools and techniques that provide significantly better information and guidance than that possible with the current typing technologies. Overall, with the support of the knowledge that flows from the genomic and bioinformatics tools developed in this project, improved prevention and control programs for fowl cholera are now possible.

This project was funded from industry revenue, which is matched by funds provided by the Australian Government.

This report is an addition to Australian Eggs Limited's range of peer reviewed research publications and an output of our R&D program, which aims to support improved efficiency, sustainability, product quality, education and technology transfer in the Australian egg industry.

Most of our publications are available for viewing or downloading through our website:

www.australianeggs.org.au

Printed copies of this report are available for a nominal postage and handling fee and can be requested by phoning (02) 9409 6999 or emailing research@australianeggs.org.au.

Acknowledgments

The authors would like to thank Associate professor Scott A. Beatson for providing support for the genomic analysis. During the project, full access to the Beatson laboratory and technical support was provided. We would also like to thank Dr Brian Forde and Mr Thom Cuddihy from Beatson Laboratory for their assistance during the project.

We thank the farm management, the veterinary consultants and the various veterinary diagnostic laboratories linked to the study farms for their assistance and expertise.

Australian Eggs Limited and AgriFutures Australia provided the funds that supported this project.

About the Authors

Dr Lida Omaleki is an early career researcher who has a veterinary science background that has been expanded by her deep interest in the use of genomics and bioinformatic to help understand disease outbreaks and to guide improved prevention and control programs. Dr Omaleki was the key researcher driving the technologies underpinning the project.

Dr Conny Turni is a mid-career researcher who has led projects on improved disease diagnostic tools and vaccines for both the pig and poultry industries. Dr Turni has been responsible for the development of a range of novel diagnostic tools that are now widely used in veterinary diagnostic laboratories.

Dr Blackall is an experienced researcher with a long history of working on bacterial diseases of meat chickens, layer chickens and pigs.

Table of Contents

Foreword.....	ii
Acknowledgments.....	iii
About the Authors	iii
List of Tables	vi
List of Figures	vi
Abbreviations.....	vii
Executive Summary.....	viii
Overall Conclusions.....	x
1 Establishment of a pipeline to allow effective rapid whole genome sequencing of <i>Pasteurella multocida</i>	1
1.1 Introduction.....	1
1.2 Methods	1
1.2.1 Bacterial isolates	1
1.2.2 WGS.....	1
1.3 Results	1
1.4 Conclusion	2
2 Use of whole genome sequencing and bioinformatics analysis to understand the genetic diversity of <i>P. multocida</i> within a fowl cholera outbreak on a layer farm (Farm A).....	3
2.1 Introduction.....	3
2.1.1 <i>P. multocida</i> LPS types	3
2.1.2 The problem	3
2.1.3 Our aim.....	4
2.2 Materials and methods	4
2.2.1 Layer Farm A.....	4
2.2.2 LPS typing and DNA fingerprinting.....	5
2.2.3 Whole genome sequencing.....	5
2.3 Results	6
2.3.1 Layer Farm A.....	6
2.4 Conclusions.....	10
3 Use of whole genome sequencing and bioinformatics analysis to understand the genetic diversity of <i>P. multocida</i> within fowl cholera outbreaks on a layer farm (Farm B)	12
3.1 Materials and methods	12
3.1.1 Farms and isolates.....	12
3.1.2 Layer Farm B.....	12

3.2	Results	12
3.2.1	Outbreaks history and vaccinations programs	12
3.2.2	Genomic results	14
3.3	Conclusions.....	17
4	Comparison of WGS/bioinformatics and gel electrophoresis with full chemical analysis to examine LPS structures in <i>Pasteurella multocida</i> isolates.....	18
4.1	Introduction.....	18
4.2	Methods	18
4.2.1	Isolates, WGS and bioinformatics analysis.....	18
4.2.2	LPS extraction.....	18
4.2.3	Carbohydrate separation and silver staining	19
4.3	Results	19
4.3.1	WGS and bioinformatics analysis.....	19
4.3.2	Gel electrophoresis patterns.....	20
4.3.3	Comparison of WGS/bioinformatics analysis and gel electrophoresis patterns with the results from full chemical analysis	24
4.4	Conclusions.....	26
5	Comparison of costs for different <i>P. multocida</i> typing methods and possible models for provision of typing services to industry	29
5.1	Introduction.....	29
5.2	Costs	29
5.3	Ongoing provision of the typing tools developed in this and previous projects.....	31
6	General conclusion and recommendations	32
7	References	34
8	Plain English Summary	36

List of Tables

Table 1-1 Assembly metrics for the <i>P. multocida</i> isolates used in the study	2
Table 2-1 Layer Farm A outbreaks and isolates	4
Table 3-1 Details of the outbreaks on layer B farms and the 32 isolates examined in this study	13
Table 4-1 Summary of the results of the genotypic and phenotypic tools developed in this project in terms of their ability to predict the presence of full, semi-truncated or fully truncated LPS, as previously determined by chemical analysis by Harper et al. (2013)	21
Table 4-2 Summarised results of the performance of the SDS-PAGE analysis for LPS truncation as compared with the gold standard chemical analysis of Harper et al. (2013)	26
Table 5-1 Costs and suitability of various typing options to guide fowl cholera vaccination programs*	30

List of Figures

Figure 2-1 rep-PCR analysis of the outbreak <i>P. multocida</i> isolates (black font) and the vaccine strains (red font) used at the time	7
Figure 2-2 Maximum likelihood tree of the core genome SNPs of the <i>P. multocida</i> isolates from the first investigated free range layer Farm A	8
Figure 2-3 Schematic diagram showing the major mutations identified in the LPS outer core biosynthesis loci of the isolates and the vaccine used at rearing	10
Figure 3-1 Genomics relatedness of the isolates from the free range layer Farm B outbreaks	16
Figure 4-1 Schematic diagram demonstrating how mutations that prevent gene function affect the LPS type L3 outer core structure	20
Figure 4-2 Nucleotide alignment of the LPS outer core biosynthesis loci and its flanking area of the <i>P. multocida</i> reference strain (Pm70), and those of the reference strains for Heddlestone serovars 3 (P1059) and 4 (P1662) and the 23 isolates of Harper et al. (2013), as well as PM1422 (red font) from the investigated chicken meat farm	22
Figure 4-3 Representative SDS-PAGE gels to show the typical carbohydrate separation and silver staining of the LPS from the field isolates included in this study	23
Figure 4-4 Comparisons of the carbohydrate silver stain pattern produced by 5 isolates that the mass spectrometry data suggested harboured a <i>gatG</i> mutation	23
Figure 4-5 Comparisons of the carbohydrate silver stain pattern produced by 2 of the Layer B isolates (PM842 and PM1541) with 2 of those previously analysed by Harper et al. (2013) (PM1120 and PM1258) as described in Chapter 4	24

Abbreviations

bp	Base pair
DNA	Deoxyribonucleic acid
Gal	Galactose
Glu	Glucose
Hep	Heptose
kDa	Kilodalton
LPS	Lipopolysaccharide
LptA	Lipopolysaccharide phosphoethanolamine Transferase A
MLST	Multi-locus sequence typing
MST	Minimum spanning tree
mPCR	Multiplex PCR
MRG	Microbiology Research Group
NCBI	National Center for Biotechnology Information
PCho	Phosphocoline
PCR	Polymerase chain reaction
QAAFI	The Queensland Alliance for Agriculture and Food Innovation, The University of Queensland
rep-PCR	<u>R</u> epetitive <u>e</u> xtragenic <u>p</u> alindromic polymerase chain reaction (PCR)
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SNP	Single nucleotide polymorphism
SRA	Sequence read archive
ST	Sequence type
WGS	Whole genome sequencing

Executive Summary

This project has used cutting-edge technology, whole genome sequencing and bioinformatic analysis to understand the infection dynamics of fowl cholera outbreaks on two free range layer farms.

Fowl cholera is one of the major threats to the fast-growing layer industry in Australia. Currently, killed *P. multocida* vaccines are specifically formulated and produced for individual farms using the bacteria that has been previously found on the farm in prior disease outbreaks. However, as the industry shifted to free range/organic production systems, the chickens suffered increasing exposure to wildlife and predators, both carrying different types of *P. multocida*. This has resulted in a significant increase of the disease, with more genetic diversity observed in *P. multocida* causing the disease. This genetic variation in the bacteria has caused major problems in the vaccine efficacy and the current problem of vaccination failure, which is impacting these fast growing industries.

For the purpose of this project, we have tried to capture of concept of genetic diversity by referring to 'clouds'. The key point we are highlighting is that an outbreak of fowl cholera is not necessarily caused by a single genetically identical isolate. Rather, an outbreak can consist of a 'cloud' of genetically diverse isolates.

A critical issue in the performance of killed fowl cholera vaccines is the carbohydrate structure that comprises the outer surface of *P. multocida*. This outer surface is called the lipopolysaccharide (LPS), and is a critical issue in the protection provided by killed fowl cholera vaccines. A killed vaccine strain provides protection only against field isolates that have the same carbohydrate structures – with even a single sugar change resulting in a loss of protection. A key issue of this project was to understand how the various typing methods developed in this project (as well as those from our previous industry funded projects) were able to predict the ability of a killed vaccine to protect against a field isolate. Essentially, this prediction requires the typing assay to provide an in depth understanding of the LPS structure.

This project had the following objectives:

1. Establish a pipeline to allow effective rapid whole genome sequencing and then full bioinformatic analysis to be able to understand and measure genetic diversity within a disease outbreak.
2. Use the above pipeline and analysis approach on a well characterised and extensive collection of *Pasteurella multocida* from two prolonged fowl cholera outbreaks – one involving an organic broiler operation and the other involving a free range layer operation (including the various killed vaccine strains used over the period of the two outbreaks).
3. Use the sequencing and bioinformatic analysis to understand the contribution of clouds (genetic diversity) to evasion of vaccine protection, and to thus assist in ensuring a match between vaccine protection and field challenge.
4. Transfer the established technologies to the user pays service provided by the Microbiology Research Group for ongoing use by the poultry industry beyond the life of the project.

This project firstly established and validated a flow of work that involved whole genome sequencing and bioinformatic analysis. This validated workflow was then used on 72 isolates of *P. multocida*, which had been previously obtained from cases of fowl cholera on two free range layer farms – Farm A (37 isolates) and Farm B (35 isolates). The isolates had been collected over 22 years (Farm A) and 10 years (Farm B). As well, the strains in the killed vaccines used on the farm were examined.

The sequence data was used to identify the multi-locus sequence type (MLST) of the isolates. MLST is a typing system based on seven conserved housekeeping genes. Five different MLST types (ST7, ST8, ST9, ST20 and ST274) were identified in association with the fowl cholera outbreaks (Farm A – ST9 and ST20; Farm B – ST7, ST8, ST9, ST20 and ST274). All tested isolates from the two farms were LPS type L3. Hence our results clearly demonstrate that different genetic lineages of *P. multocida* have the capacity to carry this LPS type. Our results also demonstrated a high degree of diversity in the LPS type L3.

Using genomic data and the phylogenetic study performed, enabled us to investigate whether an outbreak was as a result of the introduction of a new strain or the persistence of the previous strains on the farm.

Importantly, we identified that the currently in-use DNA fingerprinting techniques, rep-PCR and MLST, can fail to identify the variations in the carbohydrates in the LPS type L3 isolates. We showed that ST9 isolates with the same rep-PCR patterns carried different mutations in their LPS L3 region resulting in the production of a different LPS outer structure. As the protection afforded by killed vaccines is limited to isolates with identical LPS structures to the vaccine strains, this inability to detect LPS structure differences is a major failing of both rep-PCR and MLST. Interestingly, it was shown that *P. multocida* isolates associated with a single outbreak can have different outer LPS type L3 structures, which would result in a failure of a vaccine program based on just one of the structures. This finding of diversity within an outbreak is the first such report of this level of genetic diversity (in a key protective antigen) within an outbreak of fowl cholera. The finding emphasises the importance of investigating more than one isolate from each outbreak.

In the layer Farm B, a complex of farms, our results clearly demonstrated that each individual farm within a complex may have a different *P. multocida* population with completely different LPS outer structure. This enhances the need to treat each farm individually, or having a multi valent vaccine containing strains expressing all the different LPS structures found across the farms of the complex.

As well as the genomic work, we evaluated the use of sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) to determine the variation in the lipopolysaccharide (LPS) outer structure of *P. multocida* that carry LPS type L3. To do this, we used *P. multocida* isolates for which the LPS structure had already been fully determined by mass spectrometry. We then performed WGS and bioinformatic analysis to extract and examine the LPS outer core biosynthesis loci of those isolates. We then compared our genomic results and the patterns in the SDS-PAGE gels with those of the prior full characterisation as done by mass spectrometry. We found that our genomic analysis and our SDS-PAGE gel method are capable of differentiating between some forms of truncations in the LPS type L3 outer structure. However, neither of our methods could be used to confidently identify all of the various truncations of LPS structure of *P. multocida* isolates at this stage. However, the promising results demonstrated that further work on the Australian field isolates would enable us to better resolve the issue. We have then provided our recommended workflow for the use of the typing technologies developed in this project (and in previous industry funded projects).

In the final stage of our project, we compared the relative cost of the various typing assays developed in this project (as well as those we have developed in previous projects). This costing exercise was performed to demonstrate the comparative costs of the various typing tools. Both this costing exercise and the prior sections on the depth and quality of the information provided by each method will help guide industry in deciding which tools provide the relevant information and the relative cost of those tools.

Overall Conclusions

This project has important implications for the layer industry. Our work has shown for the first time that there is high inter- and intra-chicken genetic diversity in *P. multocida* isolates during an outbreak of fowl cholera. Importantly, it demonstrates the benefit of a genomics/bioinformatic analysis approach as a means of providing detailed information on the possible structure of the outer LPS. As killed fowl cholera vaccines only provide protection against strains with an outer core LPS that is identical or nearly identical, our approach is the optimal technology that can guide strain selection for killed autogenous fowl cholera vaccine. Importantly, our genomics/bioinformatics approach was able to explain why the outbreaks of fowl cholera on the study farms were not controlled by vaccination.

Our study clearly demonstrates the advantages of genomic analysis over the conventional PCR based approaches in providing clear insights in terms of the linkage of isolates within and between outbreaks. More importantly, it provides more detailed information than the multiplex LPS PCR (an assay that assigns isolates to an LPS genotype) on the possible structure of the outer LPS, which is very important in the case of strain selection for killed autogenous vaccines. As well, we have been able to show that our genomic approach can predict most (but not all) of the phenotypic variation in the LPS structure as detected by mass spectrometry. As mass spectrometry is not a feasible or practical routine method for confirmation of the detailed LPS structure, we believe that our genetic/bioinformatics approach is the best current option both to understand the epidemiology of fowl cholera outbreaks and to guide the selection of strains for killed fowl cholera vaccines.

Recent failures to attract ongoing research funding have prevented us from providing a firm plan to support the ongoing provision of these *P. multocida* typing tools to industry via a 'user pays' service provided by our research group. In the absence of industry funding, we have constructed a temporary arrangement that will allow the services to be provided for the rest of 2020. Access to the technologies from 2021 onwards is doubtful. It is possible that other Australian laboratories will use the formal publications describing our work to provide access to the assays.

1 Establishment of a pipeline to allow effective rapid whole genome sequencing of *Pasteurella multocida*

1.1 Introduction

This chapter describes the establishment of a pipeline to allow effective rapid whole genome sequencing and then full bioinformatic analysis to be able to understand and measure genetic diversity within a disease outbreak. To do that, the Microbiology Research Group established a close collaboration with the laboratory of Associate Professor Scott Beatson from the School of Chemistry and Molecular Biosciences at the University of Queensland. The results from this collaboration are reflected in the WGS and bioinformatic related section of the methods set out in each Chapter throughout this report.

1.2 Methods

1.2.1 Bacterial isolates

Six *P. multocida* isolates (PM1330, PM1331, PM1332, PM1333, PM1334, PM1335) from the culture collection of the MRG were selected for initial examination and establishment of the required bioinformatics methods. The bacterial isolates were revived and cultured aerobically overnight on 5% sheep blood agar. DNeasy UltraClean Microbial Kit (Qiagen GmbH, Hilden, Germany) was used for DNA extractions. The DNA samples were then sent for sequencing on the Illumina NextSeq 500 platform (150 bp paired ends).

1.2.2 WGS

Raw illumina reads were quality controlled using FastQC v0.11.5 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and possible read contamination was accessed using Kraken v2.0.8 (Wood and Salzberg, 2014). Raw Illumina sequencing data was quality filtered using Trimmomatic v0.36 (Bolger et al., 2014) or Neson v0.132 (<https://github.com/Victorian-Bioinformatics-Consortium/neson>), quality assessed using FastQC and *de novo* assembled using SPAdes v3.10.1 (Bankevich et al., 2012). QUAST v2.3 was used to check the final assemblies' metrics (Gurevich et al., 2013). *In silico* multi-locus sequence typing (MLST) profiling was performed using MLST v2.8 (<https://github.com/tseemann/mlst>) and the *P. multocida* RIRDC-MLST scheme (Jolley et al., 2004; Subaaharan et al., 2010). A local database consisting of the gene sequences representing the 16 Heddlestone serovar LPS outer core biosynthesis loci was prepared and Blastn v2.2.31+ was used to perform *in silico* LPS typing. The relevant contig containing the outer core biosynthesis locus was extracted from the draft assemblies. Geneious R11 (<https://www.geneious.com>) was used for annotation of the gene in the LPS region and ClustalW v2.1 (Larkin et al., 2007) for alignments.

1.3 Results

We assessed two different techniques for trimming the raw Illumina reads. The trimmed reads from

each method were then used for *de novo* assemblies and extracting the LPS outer core biosynthesis loci from the relevant contig(s) from the draft genome assemblies. Those genomes generated following the Trimmomatic filtered reads gave better assembly metrics with the region of interest mostly in one contig as compared with Neson-based filtering. The assembly metrics from the six tested isolates using the Trimmomatic filtering are summarised in Table 1-1.

Table 1-1 Assembly metrics for the *P. multocida* isolates used in the study

Isolates	Total length ¹	GC (%)	No. contigs ²	Largest contig	N50 ³
PM1330	2243550	40.31	24	486,243	329,385
PM1331	2253286	40.33	23	504,096	300,260
PM1332	2256570	40.32	23	596,262	300,260
PM1333	2242306	40.31	25	504,197	169,364
PM1334	2301114	40.53	25	504,141	300,260
PM1335	2252801	40.33	24	504,096	183,760

¹ The total length of the genome in base pairs in the draft assemblies.

² Number of contigs in the final draft assemblies that are more than 100bp.

³ N50 is the minimum contig length needed to cover 50% of the genome. It means half of the genome sequence is in contigs larger than or equal to the N50 contig size.

We have then performed *in silico* MLST and LPS typing using the WGS data. These results were then compared to those obtained by conventional single gene amplification and sequencing for MLST and the LPS multiplex-PCR. These conventional assays were performed in 2009 as a routine diagnostic service provided by MRG.

All six tested isolates belong to ST20 and carried LPS type L3 when the WGS data was analysed. This matched the conventional assay results (all ST20 and LPS type L3). As well, in other studies, we have shown that our WGS approach is valid for a range of the other *P. multocida* LPS types – type L1 in meat chickens (Omaleki et al., 2020b), types L3, L4 and L6 in squirrel gliders (Omaleki et al., 2020a), and type L3 in fur seals (Crawford et al. 2019). Overall, we now have considerable evidence of the feasibility of our procedure.

1.4 Conclusion

The aim of the study was to establish the capacity of the MRG to perform WGS and bioinformatic analysis. The MRG established a successful collaboration with the laboratory of Associate Professor Scott Beatson, which enabled us to have full access to its servers and computational support. This has been reflected in the recently published manuscript that resulted from the meat chicken aspect of this study. A similar formal publication based on the layer chicken studies described in the Chapters 2 and 3 is under preparation.

2 Use of whole genome sequencing and bioinformatic analysis to understand the genetic diversity of *P. multocida* within a fowl cholera outbreak on a layer farm (Farm A)

Chapter 2 and 3 are focused on investigations into fowl cholera outbreaks on two different free range layer farms. Chapter 2 will cover Farm A and Chapter 3 will cover Farm B. The diversity of *P. multocida* isolates associated with fowl cholera, the concept of 'clouds' of isolates, have been investigated within each farm. As well, WGS and bioinformatics analysis were used to investigate the LPS outer core biosynthesis loci of *P. multocida* isolates. The WGS results were then compared to those obtained from rep-PCR DNA fingerprinting, MLST and LPS multiplex PCR, the techniques currently in use to guide vaccination.

2.1 Introduction

2.1.1 *P. multocida* LPS types

Fowl cholera, caused by *P. multocida*, has become a re-emerging problem in the free range layer industry in Australia (Singh et al., 2013).

Lipopolysaccharides are one of the most important immunogenic virulence factors of *P. multocida*. Based on their lipopolysaccharide (LPS) antigens, *P. multocida* strains are classified into 16 somatic (Heddlestone) serovars (1 to 16) (Brogden et al., 1978; Heddlestone et al., 1972).

Eight LPS genotypes (L1 to L8) have been recognised between the 16 Heddlestone serovars in *P. multocida*, based on their LPS outer core biosynthesis locus (Harper et al., 2015). LPS type L3 (represented by Heddlestone serovars 3 and 4) has been recognised as the most prevalent LPS type associated with fowl cholera in Australia (Harper et al., 2013; Turni et al., 2018). The outer core biosynthesis loci of LPS type L3 is a 6,170 bp region consisting of six glycosyltransferase genes required for the assembly of the outer core. These transferase genes are *htpE* (required for the addition of heptose, Hep IV, the first outer core sugar), *gctC* (required for the addition of the 2 glucose molecules, Glc III and Glc IV, to the first heptose), *gatF* (required for the addition of the first galactose, Gal I, to Glc IV), *gatG* (required for the addition of the second galactose molecules Gal II, to Gal I), and *natB* and *natC* (required for the addition of the two *N*-acetyl-galactosamine, GalNAcI and GalNAcII, respectively (Harper et al., 2013).

A total of six different LPS outer core structures have been recognised in field strains carrying LPS outer core biosynthesis loci L3 (Harper et al., 2013). This could be as a result of inactivating mutations, such as those resulting in a frame shift or introducing a stop codon, or a result of some yet unknown mechanisms (Harper et al., 2013).

2.1.2 The problem

Vaccination with killed whole cells has long been believed to provide protection only against isolates of the same Heddlestone serovar as present in the vaccine (Blackall and Hofacre, 2020). Recently, it has

been shown that killed whole cell vaccines give protection only against strains with identical or nearly identical LPS structures (Harper et al., 2016). In Australia, multi-locus sequence typing (MLST), Heddlestone serotyping, and repetitive element PCR finger-printing (rep-PCR) are used for studies into the epidemiology of disease outbreaks and also to guide autogenous vaccinations (Turni et al., 2018). However, none of these techniques is discriminatory enough to provide the required insight into the LPS structure of the strains of the LPS type L3 genotype.

2.1.3 Our aim

Our aim was to use whole genome sequencing and bioinformatic analysis to examine strain relatedness of *P. multocida* isolates associated with fowl cholera outbreaks as well as comparisons between their LPS outer core biosynthesis loci. We wanted to assess the feasibility of WGS data to compare the vaccine strains and the outbreak strains in terms of their capacity to produce diverse LPS structures, and investigate if the repeating outbreaks in the investigated farms were as the result of the changes in the LPS.

2.2 Materials and methods

2.2.1 Layer Farm A

A total of 37 isolates obtained during a 22-year period (1994 to 2016) were included in this study on Farm A. The first two outbreaks (1994 and 2002) have been previously described (Zhang et al. 2004). The farm was a multi-age farm located in coastal south-east Queensland, which continuously had birds with no down time or depopulation period. In 1994, the farm had 5 sheds and a total of 8,000 layers, which expanded to 15 sheds with a total number of 50,000 layers in 2002 (Zhang et al., 2004). The 1994 outbreak happened in one of the 5 sheds only, following a dingo attack (Zhang et al., 2004).

Table 2-1 Layer Farm A outbreaks and isolates

Year	Outbreak Isolates	No. of sheds affected	No. of chickens	ST	Serovar ^c	rep-PCR	LPS L3 genomic features
1994	PM204 to PM214	1	4	20	Not done	Not done	<i>gatG</i> 14 bp deletion
2002	PM979 to PM982	3	4	20	4	3	<i>gatG</i> 14 bp deletion
2011	PM1533 & PM1542 ^a	2	2	9	3	2	<i>natC</i> 7 bp insertion
2013	PM1775 to PM1777	2	3	9	3	1	No missense mutation
2014	PM1778 ^b to PM1947	na ^d			3	2	<i>natC</i> 7 bp insertion
	PM1949	2		9	3	2	<i>gatG</i> 1 bp insertion
2015	PM2142	na	1	9	Non-typable	1	No missense mutation
2016	PM2232 to PM2240	na	6	9	Not done	1	No missense mutation

^a PM1542 was isolated from different aged hen in a different shed than PM1533.

^b PM1778 harboured a *gatF* E63K change compared to the 2011 strains.

^c Heddlestone serovar.

^d na: not available.

The number of isolates per outbreak received by the laboratory was different across the years, ranging from only one isolate to multiple isolates from the same chicken and also isolates from different chickens within an outbreak. In 2003, strain PM979, isolated during the 2002 outbreak, was used as a killed vaccine and administered as two doses during rearing. However, due to financial constraints the vaccination schedule was changed to once at 12 weeks of age and fowl cholera re-emerged on the farm in 2011. The re-emergence of fowl cholera in 2011 prompted a return to the original vaccination routine (twice per bird) but the disease persisted from the years 2013 to 2016, despite several changes in the vaccine strains based on MLST profiling and DNA fingerprinting. A total number of 19 isolates were available from the 2011, 2013, 2014, 2015 and 2016 outbreaks, and were included in the current study (Table 2-1).

2.2.2 LPS typing and DNA fingerprinting

Both the LPS typing and DNA fingerprinting assays were performed on the isolates as part of the initial investigations of the MRG performed at the time of the investigations. Genomic DNA of all the isolates was prepared using the PrepMan Ultra Sample Preparation reagent (Life Technology, Applied Biosystems, Warrington, UK). The isolates were subjected to the LPS multiplex-PCR (Harper et al., 2015). As well, the repetitive extragenic palindromic polymerase chain reaction (rep-PCR) was performed using a modification of a previously described method (Gunawardana et al., 2000). The modified rep-PCR used Go Taq and the Go Taq reaction buffer (Promega, Alexandria, NSW).

2.2.3 Whole genome sequencing

The bacterial isolates were revived and cultured aerobically overnight on 5% sheep blood agar. The DNeasy UltraClean Microbial Kit (Qiagen GmbH, Hilden, Germany) was used for DNA extractions according to the manufacturer's instructions, and the DNA sequenced on the Illumina NextSeq 500 platform (150 bp paired ends) by the Australian Centre for Ecogenomics at the University of Queensland (St Lucia QLD 4072, Australia). Raw Illumina sequencing data was quality filtered using Trimmomatic v0.36 (Bolger et al., 2014), quality assessed using FastQC v0.11.5 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and *de novo* assembled using SPAdes v3.10.1 (Bankevich et al., 2012). QUAST v2.3 was used to check the final assemblies' metrics (Gurevich et al., 2013). Bandage (Wick et al., 2015) was used for visualisation of the assemblies. *In silico* multi-locus sequence typing (MLST) profiling was performed using MLST v2.8 (<https://github.com/tseemann/mlst>) and the *P. multocida* RIRDC-MLST scheme (Jolley et al., 2004; Subaaharan et al., 2010). A local database consisting of the gene sequences representing the 16 Heddlestone serovar LPS outer core biosynthesis loci was prepared and Blastn v2.2.31+ was used to perform *in silico* LPS typing and extraction of the corresponding region from the carrying contigs, where the matching LPS type was selected based on the e-value. Geneious R10 (<https://www.geneious.com>) was used for annotation of the gene in the LPS region and ClustalW v2.1 (Larkin et al., 2007) for alignments and comparisons to the corresponding loci in *P. multocida* strain Pm70 (NC_002663.1).

Nesoni v0.132 (<https://github.com/Victorian-Bioinformatics-Consortium/nesoni>) was used for read mapping of the trimmed paired end reads against Pm70 and X73 for LPS type 3 and LPS type 1 carrying isolates, respectively. Artemis v16.0.0 (Carver et al., 2011) was used to check the read pile ups.

Parsnp v1.2 (Treangen et al., 2014) was used for core genome alignment and SNP calling in the draft genome assemblies. Phylogenetic trees were then constructed with RAxML x8.2.9 using the core genome SNP alignment after removal of predicted recombination sites by Gubbins v2.1.0 (Stamatakis,

2014). A general-time reversible nucleotide substitution model with a GAMMA correction for site variation was used for tree construction (bootstrap 1000 with Lewis ascertainment correction). Phylogenetic trees were visualised using FigTree v1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree/>) and Phandango (Hadfield et al., 2017). PHYLOViZ v2.0 (Nascimento et al., 2017) was used to build a minimum spanning tree from core genome SNPs of the isolates obtained from Farm B1.

2.3 Results

2.3.1 Layer Farm A

2.3.1.1 *Outbreak history, rep-PCR typing and vaccination programs*

Since 1994, seven outbreaks of fowl cholera (1994, 2002, 2011, 2013, 2014, 2015 and 2016) have been reported on Farm A (see Figure 2-2). The 1994 outbreak occurred in an unvaccinated flock. Vaccination was introduced after this outbreak but was discontinued in 2000 after a 6 year absence of any disease. Following the 2002 outbreak, an autogenous killed fowl cholera vaccine based on isolate PM979 from the 2002 outbreak was used twice at rearing. Subsequently, for cost reasons, the program was reduced to a single dose. After the re-emergence of the disease in 2011, isolate PM1533 was included together with PM979 as a bivalent autogenous killed vaccine. This decision was based on the serotyping assay done at the time which identified PM979 as Heddlestone serovar 4 and PM1533 as Heddlestone serovar 3.

In October 2013, at least two of the sheds, with 37 and 51-week old birds, were affected by fowl cholera despite the routine vaccination (Table 2-1). The pullets were vaccinated with a commercial tetra-valent (non-autogenous) vaccine at 10 weeks of age and the bi-valent killed autogenous vaccine (PM979 and PM1533) at 16 weeks of age. Comparison of the rep-PCR profiles of the four 2013 outbreak isolates with the vaccine strains indicated that three of the isolates (PM1775 to PM1777) produced a different, albeit closely related, fingerprint to that of vaccine strain PM1533 (rep-PCR type 1 in Figure 2-1 A). The fourth 2013 field isolate (PM1778) had the same rep-PCR profile as vaccine strain PM1533 (rep-PCR type 2). The rep-PCR fingerprint of vaccine strain PM979 was not seen in the field isolates. However, all four 2013 isolates were identified as Heddlestone serovar 3 by the gel immunoprecipitation assay and LPS type L3 via the LPS multiplex-PCR (Table 2-1). Therefore, a decision was made to not make any changes in the vaccine.

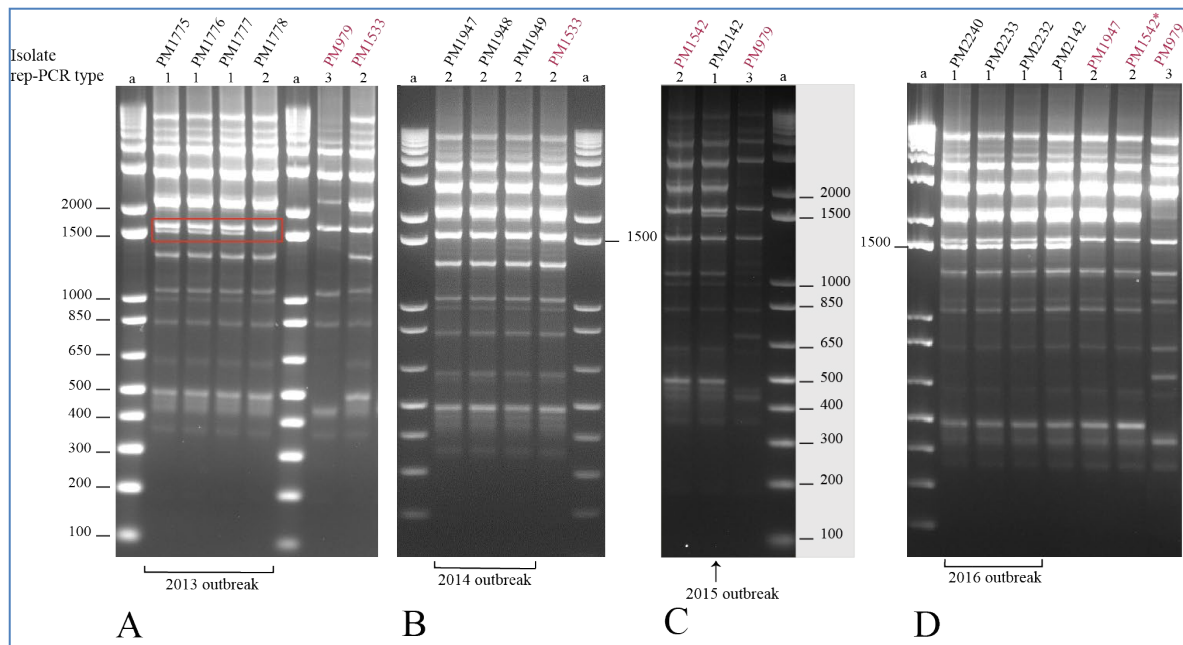


Figure 2-1 rep-PCR analysis of the outbreak *P. multocida* isolates (black font) and the vaccine strains (red font) used at the time

Panel A: A new genotype emerged in 2013 (rep-PCR genotype 1) with only one band difference with rep-PCR genotype 2 covered by vaccine strain PM1533. As well, the 2013 outbreak isolates showed 2 different patterns (red box) with isolate PM1778 having the same pattern as the vaccine strain PM1533.

Panel B: The 2014 isolates showed the same rep-PCR pattern as the vaccine strain PM1533. However, the genomic results showed that they had the capacity to produce a shorter LPS structure compared to PM1533.

Panels C and D: The rep-PCR 1 pattern reappeared in the 2016 outbreak (panel D) while the 2015 outbreak isolates had the rep-PCR 2 pattern (panel C).

a: 1kb plus DNA ladder (Invitrogen).

In October 2014, another outbreak occurred in 2 flocks that were 47 (isolates PM1947 and PM1948) and 54 (isolate PM1949) weeks of age. Older flocks also present on the farm did not show any signs of the disease. The rep-PCR fingerprinting showed that these 2014 isolates shared the same fingerprint (rep-PCR type 2) as PM1778 from the 2013 outbreak and PM1533 (the autogenous vaccine strain adopted after the 2011 outbreak). A decision was made to substitute PM1533 with PM1542, a 2011 isolate known to have the same rep-PCR fingerprint as PM1533, in the autogenous vaccine (Figure 2-1B). Following the 2015 outbreak from which the laboratory received only one isolate (PM2142) (Table 2-1), the vaccine was changed again to contain PM979 and PM1947. However, the 2016 outbreaks following this change yielded isolates that all shared the same rep-PCR as the 2013 isolates, i.e. rep-PCR type 2 (Figure 2-1 C and D). This necessitated a change in the autogenous vaccine in 2017, removing PM1947 and substituting in PM979 and PM2240 (the latter being a 2016 outbreak isolate).

2.3.1.2 Genomic results

The *in silico* MLST, generated from the WGS data, recognised that the isolates from the earlier outbreaks (1994 and 2002) all belonged to ST20 while the isolates from the 2011 outbreak onwards belonged to ST9.

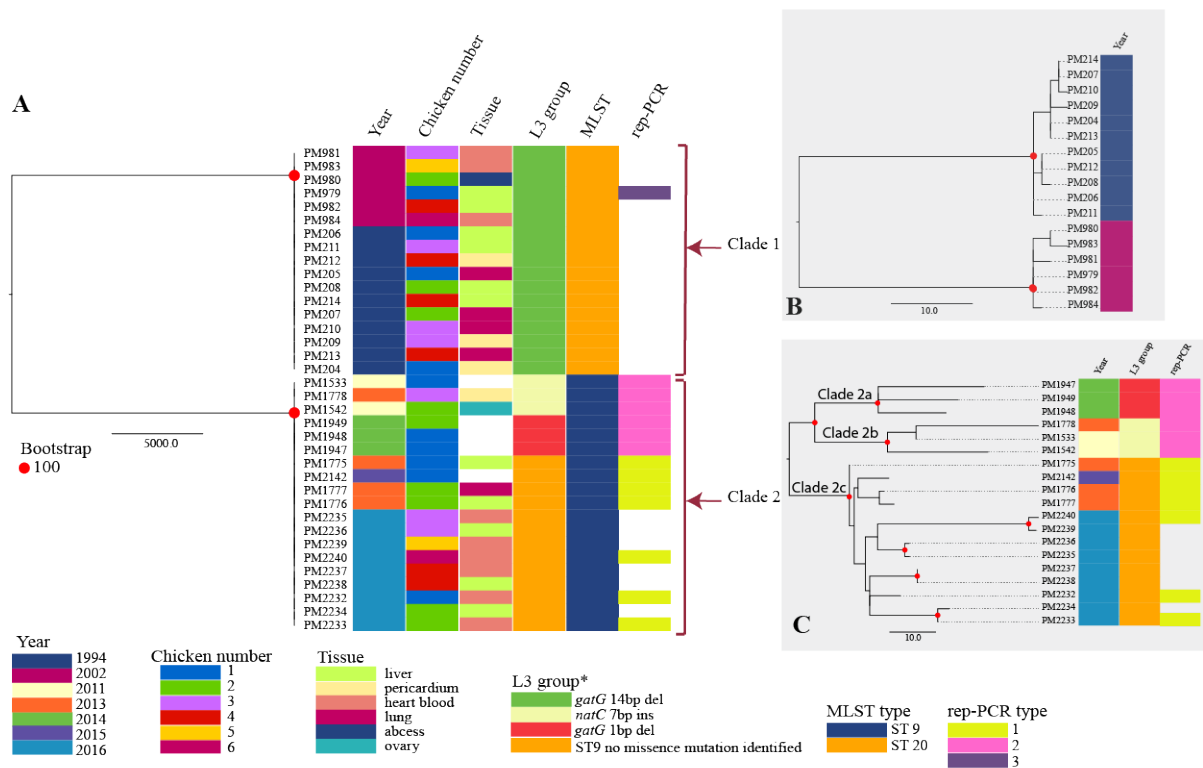


Figure 2-2 Maximum likelihood tree of the core genome SNPs of the *P. multocida* isolates from the first investigated free range layer Farm A

Panel A: Phylogenetic analysis separated the isolates into two major clades in agreement with the ST typing data. Between 0 to 15,467 core genome SNPs were identified between the core genomes of all the isolates.

Panel B: Maximum likelihood analysis of the core genome of the Clade 1/ST20 isolates.

Panel C: Maximum likelihood tree of the core genome SNP of the Clade 2/ST9 isolates.

Trees have been rooted from midpoint. Scale bar represents number of SNPs.

Maximum likelihood analysis of core genome SNPs identified two major clades in agreement with the *in silico* MLST results (Figure 2-2 A). The 1994 and 2002 outbreaks were more closely related to each other with a mean of 15.6 ± 13.7 non-recombinant SNPs (Clade 1), with isolates from each of the two years in a separate sub-clade (Figure 2-2 B). The subsequent outbreak isolates formed Clade 2 and were ST9 (mean of 34.7 ± 16 non-recombinant SNPs). These Clade 2 isolates were obtained over 6 years with multiple isolates only being available from the 2016 outbreak (Figure 2-2).

The LPS *in silico* typing revealed that all of the ST20 and ST9 isolates from this farm were carrying LPS type L3. However, the isolates from the two different clades were different in the nucleotide sequence of their LPS type L3. When compared to the sequence of *P. multocida* strain Pm70, the Clade 1/ST20 isolates all carried a 14 bp deletion in their *gatG* gene from position 491 to 505, as well as a one bp insertion in their *natC* gene (313_314insA).

More diversity was observed in the sequences of the LPS type L3 carried by Clade 2/ST9 isolates. These isolates could be separated into 3 subclades 2a, 2b and 2c based on their core genome SNPs, which was in agreement with their LPS type L3 subtype (Figure 2-2). Clade 2a consisted of the three isolates from the 2014 outbreak, PM1947 to PM1949. These isolates shared a one bp deletion at position 722 of their *gatG* gene (721_723delA), resulting in a frame shift and the likely early termination of transcription (Figure 2-3A).

Clade 2b consisted of isolates PM1533 and PM1542 from 2011 and isolate PM1778 from 2013. These isolates shared a 7 bp insertion (TTATTAT) in their *natC* gene at position 730, as well as two nucleotide changes in their *natB* gene compared to that from Pm70. In addition, PM1778 carried a non-synonymous change (E63K) in its *gatF* that was absent from other isolates within Clade 2b.

Clade 2c consisted of the remaining three 2013 isolates, a single isolate from 2015 (PM2142) and nine isolates obtained in 2016. All Clade 2c isolates shared 100% similarity in their LPS outer core biosynthesis loci, while carrying only two nucleotide changes compared to that from Pm70. This included a non-synonymous change (G487A) resulting in an amino acid change R163C and a synonymous change at position A807G in their *natB* gene. These two missense mutations were, however, shared by all the Clade 2 isolates.

A schematic diagram showing the major mutations in the LPS in the outbreak strains as well as those in the vaccine strains used at the time, is provided in Figure 2-3. It is worth noting that for the 2011, 2013, 2014, 2015 and 2016 outbreaks, the LPS L3 genotype of the challenge isolates was not represented in the vaccines at use at the time. In particular, for the 2011, 2013, 2015 and 2016 outbreaks, at least one or more of the outbreak isolates were, based on our bioinformatics analysis, producing an LPS structure that was longer than any produced by the vaccine strains.

The technique of rep-PCR has been extensively used in our laboratory to recognise genetic diversity within sets of *P. multocida* isolates. A series of studies have provided some evidence of the typing capacity of rep-PCR. As an example, rep-PCR has been shown to produce the same allocation to groups as multi-locus enzyme electrophoresis (the phenotypic precursor to MLST) when applied to a well characterised set of fowl cholera isolates (Blackall and Mifflin, 2000). In contrast, in a more recent study involving both avian and porcine isolates of *P. multocida*, rep-PCR sometimes kept isolates of the same MLST type together and sometimes subdivided isolates of the same MLST type (Turni

et al. 2018). The current study is the first study to examine whether rep-PCR patterns correlate with LPS L3 sub-genotype. The rep-PCR technique sometimes kept isolates of the same L3 genotype together, e.g. all isolates of sub-clade 2C belonged to the same rep-PCR type (type 1) and the same L3 genotype. However, all the isolates of sub-clades 2A and 2B shared the same rep-PCR type (type 2) even though the sub-clades differed in L3 genotypes. This finding clearly shows that rep-PCR typing can reveal genetic diversity but that genetic diversity is not necessarily reflective of the LPS genetic diversity, and therefore rep-PCR cannot be relied upon to predict whether a particular set of strains in a killed fowl cholera vaccine can provide protection against a field challenge strain.



Figure 2-3 Schematic diagram showing the major mutations identified in the LPS outer core biosynthesis loci of the isolates and the vaccine used at rearing

Panel A: Comparisons of the 5 different LPS 3 genotypes with that from *P. multocida* type strain Pm70 (on right). The nucleotide sequence of LPS from Pm70 is set as reference. Top grey bar next to each isolate name depicts the nucleotide sequence and the lower one depicts the amino acid translation. Black vertical lines within the grey bars show area of difference between each isolate and that of Pm70. The green arrows show insertion sites and the red arrows show the deletion sites. The proposed outer structure is depicted on left and adapted from Harper et al. (2015). The dashed line separates inner core from outer core LPS. The white circle represents heptose, the blue circle represents glucose, the yellow circle represents galactose and the yellow box represents GalNAc.

Panel B: Comparison between the outbreak strain (coloured circles) and vaccine strains (coloured rectangles) since the 2011 outbreak in layer A. Each circle represents one bacterial sample as described in Table 2-1.

2.4 Conclusions

Two MLST types, ST9 and ST20, were identified on Farm A over a 22 year period. Our data confirmed the persistence of a single strain of *P. multocida* on Farm A for over 10 years by demonstrating the isolates from both the 1994 and 2002 outbreaks were closely related at the core genome level, with isolates from each of those years allocated to a separate subclade.

The Clade 2 isolates on Farm A, first detected in 2011 (strain PM1533), belonged to ST9 and were the result of the introduction of a new population, as they were only distantly related to the Clade 1 isolates at the core genome level. This clade included all the isolates obtained during the more recent outbreaks, which emerged after almost a ten-year period where the farm was able to limit any impact of fowl cholera. However, diversity was observed in this clade of isolates including variation in the

sequence of their LPS type L3 loci. Two different genotypes in the LPS outer core biosynthesis loci of the 2013 isolates were observed. While one of the isolates shared the same genetic mutation in the LPS type L3 with the 2011 isolates, the other three isolates obtained from 2 other pullets (PM1775, PM1776 and PM1777) were 100% identical to each other and those from isolates obtained in 2015 and 2016 in their LPS type L3 sequences. Hence our results showed the diversity in the LPS outer core biosynthesis loci of *P. multocida* isolates obtained within a single outbreak for the first time.

Importantly, our genomic analysis clearly shows why the outbreaks of 2012, 2013, 2015 and 2016 occurred despite vaccination. For all four of these outbreaks, the bioinformatics analysis of the WGS data indicates that at least one of the field challenge isolates was producing an LPS structure that was longer than any structure produced by vaccine strains then in use. As Harper et al. (2016) have shown that even a single sugar change in the LPS structure can result in no cross-protection between a killed vaccine and a challenge isolate, the differences that our WGS/bioinformatics analysis have identified would explain the vaccine failures.

Our genomic/bioinformatics analysis suggests that the field isolates from the 2014 outbreak had the same LPS structure as the vaccine strains. This suggests that the failure of protection in this case was not due to inappropriate vaccine strains. There are a many reasons why an apparent 'vaccine break' occurs. Our analysis suggests that, for the 2014 outbreak, other factors such as stress, vaccination protocols and timing of the vaccination appear to underpin this vaccine break. This highlights the ability of the WGS/bioinformatics approach as a means of eliminating strain variation and allowing management and veterinary consultants to focus on the other issues related to vaccine efficacy.

Finally, these results also show that while rep-PCR typing does detect genetic diversity, the detected diversity is not necessarily reflective of the diversity present in the LPS outer core biosynthesis loci of *P. multocida* isolates.

3 Use of whole genome sequencing and bioinformatic analysis to understand the genetic diversity of *P. multocida* within fowl cholera outbreaks on a layer farm (Farm B)

3.1 Materials and methods

The details of methods used for WGS and bioinformatics analysis, rep-PCR and LPS multi-plex PCR are described in Chapter 2.

3.1.1 Farms and isolates

3.1.2 Layer Farm B

Layer Farm B was a large complex with 4 farms, named farms B1 to B4, located in south-east Queensland close to the border with New South Wales. A total number of 35 isolates from the outbreaks of fowl cholera between 2010 and 2019 as well as two of the three original vaccine strains (PM842 and PM844) associated with outbreaks in 2001, were included in this study (Table 3-1). The third vaccine strain (PM843) was an isolate from B3 that was known to be ST8 (Singh et al., 2013), which was not available for WGS at the time of the project. Each of the farms had several layer sheds with the hens having easy access to the open area and being inside the sheds during the night. The complex has been constantly suffering fowl cholera.

3.2 Results

3.2.1 Outbreaks history and vaccinations programs

The original vaccine strains used across the farms of the complex were obtained from outbreaks of fowl cholera in 2001, with PM842 obtained from the breeder operation (B5) previously owned by the complex, and PM843 and PM844 from Farm B3. In 2001, Farm B3 contained breeders in sheds and B5 was a free range layer farm. There was transfer of nest boxes from B3 to B5 and eggs from both farms were sent to the same hatchery. The two farms were 100 kms apart, with Farm B3 later populated with layers in sheds. Initially, fowl cholera vaccination was performed twice at rearing. However, since 2003, the program became a single dose at 10–12 weeks for cost reasons. In mid 2010, Farm B3 became a free range layer operation, with three production sheds and one rearing shed.

The first series of recent outbreaks in B3 was diagnosed in November 2010 with *P. multocida* isolated from the heart sac and peritoneal cavity (PM1447). The outbreak happened in a shed containing 17,400 45-week old hens. The total mortality from 16 to 76 weeks of age was 16% in this shed. The second confirmed outbreak happened in another shed on the B3 Farm, in September 2011, with isolate PM1541 obtained from the subcutaneous tissue of a hen with a swollen face and swollen wattle. The third confirmed outbreak was the reappearance of the disease in this second shed at 70 weeks of age in January 2012, with again only one isolate sent for analysis (PM1582). This shed had a total mortality of 10% up to processing.

The first outbreak in Farms B1 and B2 happened in 2016, and in Farm B4 in 2017. Farm B1 continued to suffer reoccurring fowl cholera in one of their sheds, however, cholera also affected an adjusting shed in July 2019 at 43 weeks of age.

Table 3-1 Details of the outbreaks on layer B farms and the 32 isolates examined in this study

Year	Outbreak isolate(s)	Farm	Age in weeks	Clade/ST	LPS 3 specification	Vaccine used at rearing
2001	PM842	B5		4/20	<i>gatG</i> 14 bp deletion	
2001	PM844	B3		2/8	no stop codon or frame shift	
2011	PM1447	B3		1/274	<i>htpE</i> insertion position 126	
2011	PM1541	B3		1/274	<i>htpE</i> insertion position 126	
2012	PM1582	B3		1/274	<i>htpE</i> insertion position 126	
2014	PM1845 to PM1853 and PM1855	B3		2/8	no stop codon or frame shift	
2016	PM2248	B2	62 w	4/20	<i>gatG</i> 14 bp deletion	PM842, 843, 1855
2016	PM2269 ^a and PM2305	B1	50 w	5/9	<i>natC</i> 7 bp insertion	PM842, 843, 1855
2016	PM2290	B4	50 w	3/7	<i>htpE</i> del position 695	PM842, 843, 1855
2017	PM2356	B2	50 w	4/20	<i>gatG</i> 14 bp deletion	PM842, 843, 1855
2017	PM2369	B1		5/9	<i>natC</i> 7 bp insertion	
2018 Apr	PM2383 to PM2391	B1 (I) ^b	47 w	5/9	<i>natC</i> 7 bp insertion	PM842, 843, 1855
2018 May	PM2398 and PM2401	B1		5/9	<i>natC</i> 7 bp insertion	
2018 Oct	PM2445	B1 (I) ^b		5/9	<i>natC</i> 7 bp insertion	
2018 Nov	PM2470	B1 (I) ^b		5/9	<i>natC</i> 7 bp insertion	
2019	PM2585 to PM2587	B1 (II and III) ^b		5/9	<i>natC</i> 7 bp insertion	PM1844, <u>2269</u> , 2290

^a All isolates with the *natC* 7 bp insertion, except PM2269, also carried an amino acid change in their *gatG* (C128Y) gene.

^b Roman numbers in parenthesis are the names of affected shed(s).

Vaccine strain in red has the 14 bp deletion in the *gatG* gene, so should have theoretically protected against the 2016 and 2017 outbreaks in B2. These birds were vaccinated twice at 9 and 13 weeks of age.

The vaccine strain in blue, underlined font had the 7 bp insertion in the *natC* gene.

3.2.2 Genomic results

A total of 37 isolates were included in the core genome alignments. Between 0 and 61,190 SNPs were identified between the core genomes of the isolates investigated from the farms of the complex across the years. *In silico* MLST typing identified five distinct sequence types, STs 7, 8, 9, 20 and 274, between 2001 and 2018.

Core genome alignment grouped the isolates into five clades, named Clades 1 to 5, with Clades 3/ST 7 and 4/ST 20 clustering close to each other. Comparisons of the allelic profiles of ST 7 and ST 20 revealed that they share identical alleles in four of the seven housekeeping gene sequences (adk 2, est 6, zwf 4, mdh 6). The *gdh* and *pgi* alleles have 5 and 10 SNPs differences respectively. However, the *pmi* alleles have only the 2 base differences between the two STs. These relatively small differences support the close clustering revealed in the core genome comparison.

Each clade also carried distinct genetic mutations in their LPS type L3 outer core biosynthesis loci (Figure 3-1A).

P. multocida isolates in Clade 1 were all collected from chickens at Farm B3 during outbreaks of fowl cholera in 2010 (PM1447), 2011 (PM1541) and 2012 (PM1582). All three isolates were ST 274 and harboured an insertion at position 126 of their *htpE* gene, resulting in a frame shift and early termination of transcription.

Clade 2 consisted of the original vaccine strain PM844 obtained in 2001 from Farm B3, as well as the 2014 outbreak isolates obtained from the same farm after 13 years (PM1845 to PM1855). They were all ST 8 and carried no mutation resulting in any frame shift in their LPS outer core biosynthesis loci.

Clade 3 consisted of the single ST 7 isolate, PM2290 (the only isolate from Farm B4). This isolate carried an insertion at position 695 of its *htpE* gene, presumably resulting in producing a fully truncated LPS with no outer core as previously described (Harper et al., 2013), a characteristic shared with the Clade 1 isolates from Farm B3.

Clade 4 isolates consisted of the three ST 20 isolates (PM842, PM2248 and PM2356). All three isolates, carried an identical 14 bp deletion in their *gatG* gene as well as a one bp insertion in the *natC* gene compared to *P. multocida* strain Pm70, a finding shared by isolates from Farm A. PM2248 and PM2356 were both obtained from Farm B2 during the 2016 and 2017 outbreaks, respectively.

Clade 5 consisted of 19 isolates from the outbreaks on Farm B1 from 2016 to 2019. This farm had recurring outbreaks of fowl cholera since 2016, despite routine vaccinations. These isolates belong to ST 9, carried 80 SNPs in their LPS type L3 compared to that of Pm70, and all carried a 7 bp insertion in their *natC* gene starting at position 730. However, all Clade 5 isolates except PM2269 also carried an amino acid change in their *gatG* gene, compared to the strain Pm70 (C128Y).

The draft genome assemblies of the 19 isolates from this farm were further compared to each other using parsnp, and a minimum spanning tree (MST) was constructed. These isolates shared an average 14.7 ± 5.2 SNPs non-recombinant core genome SNPs with each other, with isolates obtained from same outbreak (April 2018) sharing between 4 and 22 non-recombinant SNPs with each other (Figure 3-1C).

While many molecular epidemiological studies use an MST as a means of visualising a chain of pathogen transmission, there are important underlying constraints. Originally, MSTs were applied to situations such as connecting, by the shortest possible means, physical objects, e.g. in a telephone

network. In these physical applications, the shortest possible connection is clear and no other possible networks exist (Salipante and Hall, 2011). However, it is now understood that, in epidemiological applications (such as the current study), multiple alternative MSTs are possible (Salipante and Hall, 2011). Hence, the MST presented here as Figure 3-1C is simply one of many possible MSTs. The presented MST should simply be viewed as one (of many) possible mathematical visualisations of how the isolates of *P. multocida* isolates from Farm B1 over a three year span could be connected.

While recognising this limitation of multiple possible MSTs, the visualisation shown in Figure 3-1C is useful. Firstly, the MST clearly shows the diversity that occurs within an outbreak. The April 2018 outbreak was the outbreak with the most isolates, and the MST clearly demonstrates that the outbreak consists of a 'cloud' of related but different isolates (the purple circles, PM2383 to PM2391). Similar diversity is demonstrated with the 2019 isolates (the red circles, PM2585 to PM2587). Secondly, in general, the MST also shows that the isolates from a specific outbreak in time were typically more closely linked to each other than to the isolates from the other outbreaks, despite all these outbreaks occurring on the one farm. The fact that the presented MST is only one of many possible MSTs that could be drawn from this dataset means that a conclusion that the 2019 outbreak (the red circle isolates, PM2585 to PM2587) directly evolved from the 2016 outbreak isolate (blue circle, PM2305) should not be drawn.

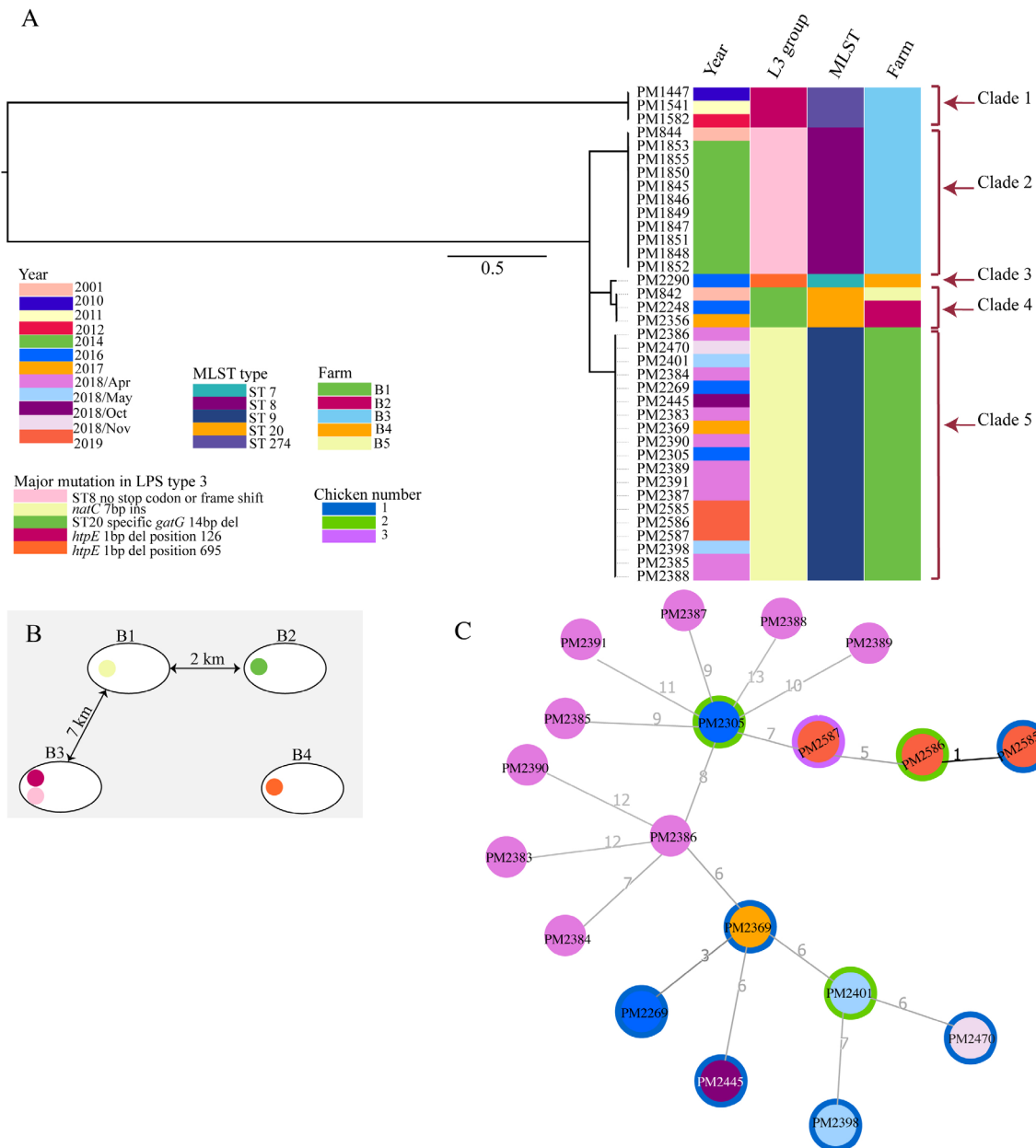


Figure 3-1 Genomics relatedness of the isolates from the free range layer Farm B outbreaks

Panel A: Phylogenetic tree of the isolates obtained from all 4 farms across the complex between 2010 and 2019. Tree is rooted from midpoint, and scale bar represents the number of substitution per site.

Panel B: Schematic diagram representing the distance between the 4 farms. Coloured circles represent the major mutations in the LPS type L3 carried by the isolates obtained from each farm.

Panel C: Minimum spanning tree of the core genome SNPs of isolates obtained from B1 across the years. The internal circle colour represents year of isolation, the external circle colour represents the chicken (i.e. all circles with an outer ring of green are from the same chicken, and so on). The nine isolates from April 2018 (dark pink circles) were obtained from the liver of nine different chickens. The numbers on the branches represent the number of different SNPs between the relevant two isolates.

3.3 Conclusions

Five different clades of *P. multocida*, each associated with a unique ST, have been associated with fowl cholera across the farms of the Layer complex B, with two of them (Clade 3/ST 7 and Clade 4/ST 20) being closely related to each other. Within a clade, all isolates had the same sub-type of the LPS type L3 loci. The isolates could also be divided into five different LPS type L3 sub-types based on the presence or absence of major genetic mutations in their LPS outer core biosynthesis loci. Isolates of *P. multocida* of LPS type L3 and belonging to ST 7 and ST 20 were previously identified in association with both pig and poultry hosts (Omaleki et al., 2020b; Turni et al., 2018) with ST 7 isolates also reported from fatal pasteurellosis in captive pinnipeds (Crawford et al., 2019). However, the isolates from ST 7 and ST 20 described in the current study appear to carry different missense mutations in their glycosyltransferase genes as compared with these previously reported isolates. These differences were predicted to produce different LPS outer structure with the ST 7 isolate on this farm having the capacity to produce a fully truncated LPS due to a newly identified missense mutation in the *htpE* gene, while the ST 20 isolates were predicted to produce an LPS that would end after the addition of the first galactose molecule, i.e. Hep.[Glc].Glc.Gal. We have previously reported a frame shift in the *gatG* gene of ST 7 *P. multocida* isolates obtained from diseased pinnipeds in south-east Queensland (Crawford et al., 2019). These results indicate that ST 7 *P. multocida* isolates have the capacity to produce different LPS structures within an LPS PCR type, a capacity also shared with the ST 9 *P. multocida* isolates described in the current study. The observed diversity between the strains from different farms of the Layer complex B would make choosing a single autogenous vaccine to be used across all the farms of the complex very challenging.

4 Comparison of WGS/bioinformatic and gel electrophoresis with full chemical analysis to examine LPS structures in *Pasteurella multocida* isolates

4.1 Introduction

Advances in biochemical, analytical and bioinformatics approaches have resulted in major progress in determining the structures of complex carbohydrates (Imperiali, 2019). However, it has been suggested that the current knowledge of bacterial carbohydrates diversity, their modifications and the glycans into which they are embedded may be just the ‘tip of the iceberg’ (Imperiali, 2019). Certainly, *P. multocida* is not an exception to this situation.

In this Chapter we describe using Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) to determine the possible variation in the LPS outer structure of *P. multocida* that carry LPS type L3. To do this, we used the isolates for which the LPS structure has been previously fully analysed by mass-spectrometry (Harper et al., 2013). We performed WGS and bioinformatics analysis to extract and examine the LPS outer core biosynthesis loci of those isolates and compared our results with those of Harper et al. (2013). We also compared the known structure data of Harper et al. (2013) with the patterns on the carbohydrate gels. The aim of this work was to determine if our tools (WGS and bioinformatics analysis and SDS-PAGE) could be used to predict when isolates of *P. multocida* had full or truncated LPS structures.

4.2 Methods

4.2.1 Isolates, WGS and bioinformatic analysis

The 23 *P. multocida* isolates whose LPS structure was previously analysed by Harper et al. (2013) were subjected to WGS and bioinformatics analysis as detailed in the previous Chapters. The extracted LPS region and the flanking area were then annotated using Geneious v10 inbuilt annotation function. The region was then aligned to that from *P. multocida* isolate Pm7 in Geneious using ClustalW. Manual curation was undertaken to ensure start codons were aligned correctly.

4.2.2 LPS extraction

The 23 *P. multocida* isolates mentioned above, as well as two isolates from Layer B (PM842 and PM1541), were subjected to LPS extraction and carbohydrate analysis. The LPS was extracted following the protocol on the Hancock Lab website (<http://cmdr.ubc.ca/bobh/>) according to Hitchcock and Brown (Hitchcock and Brown, 1983). The *P. multocida* isolates were grown overnight at 37°C in air on 5% sheep blood agar. The cells were harvested using a sterile loop and suspended in 0.9% NaCl with an optimised absorption density 600 of 0.4.

A 1.5 ml aliquot of the suspension was then centrifuged ($8,000 \times g$ for 3 minutes). The pellet was then resuspended in 100 µl of a lysis buffer containing 2% SDS, 4% 2-mercaptoethanol, 10% glycerol and 0.1 M Tris-HCl, pH 6.8. The lysate was then heated using a heating block at 100°C for 10 minutes. For protein digestion, 10 µl of the lysing buffer containing 25 µg of proteinase K was added to each sample and incubated in a water bath at 60°C for 1 hour.

4.2.3 Carbohydrate separation and silver staining

The extracted LPS was then mixed with equal amounts of Tris-Tricine SDS sample buffer (2×), heated for 5 minutes and a 10 µl aliquot was run on each of a 10%, 16% or a gradient (10-20%) Tris-Tricine polyacrylamide gel using Tris-Tricine SDS buffer (Novex™ ThermoFisher Scientific). SeeBlue Plus2 pre-stained protein standard (ThermoFisher Scientific Australia) was used as a marker for running the gel. The gels were run at a constant 165 V for an hour followed by staining using a Pierce Silver Stain kit (ThermoFisher Scientific Australia), using the protocol provided by the manufacturer.

4.3 Results

The list of the isolates, their LPS outer structure as determined by mass spectrometry and the detected gene mutations, all as previously determined by Harper et al. (2013) are shown in Table 4-1.

It is important to note that Harper et al. (2013) only examined the sequence of those genes whose disruption was predicted by the LPS structure analysis. In short, Harper et al. (2013) performed amplification and sequencing of only the predicted gene expected to explain the chemical structure.

Our WGS analysis, however, confirmed the results obtained by Harper et al. (2013) while also providing more information on the complete LPS outer core biosynthesis loci type L3, with details described in the following section.

4.3.1 WGS and bioinformatic analysis

Unlike the Harper et al. (2013) study, our use of WGS allowed a full examination of the whole LPS outer core biosynthesis loci and the flanking area (a 6,500 bp region) via bioinformatics analysis. The full LPS loci were then compared with the originally sequenced *P. multocida* strain Pm70, carrying LPS type L3 and known to produce the longest outer structure. A visual overview of the data is presented in Figure 4-1.

The LPS region from 20 of the 23 isolates was in one contig. Three isolates (PM1099, PM1103 and PM1205) had their LPS region present in more than one contig. In order to investigate the LPS region of these three isolates, the assembly graphs were investigated in Bandage where the nodes containing the LPS region were extracted again from the graphs, and annotated in Geneious in order to look for any possible large insertion or deletion. The results confirmed the presence of a 42 bp nucleotide insertion at position 353 of *gatG* in isolate PM1205 and a 183 bp nucleotide deletion at position 184 of *natC* in PM1103. For PM1099, both assembly-based and read mapping-based methodologies indicated two major deletions in the beginning of *gatG* and *natC* genes. No signs of insertion sequences were found when the assembly was checked in Bandage. However, the results from the mass spectrometry of the LPS of PM1099 were indicative of the presence of an intact *gatG* region. Therefore a final resolution of the occurrence of LPS loci of PM1099 was not achieved.

Each isolate was then assigned to an LPS type L3 genotype, in which a genotype was defined as consisting of isolates that were identical in the sequence across the entire region of the LPS outer core biosynthesis loci and the flanking area (see Table 4-1). Isolates that differed by no more than two base pair changes were included in the same genotype and differentiated by being marked with a 'prime' added to the genotype name. Three pairs of isolates were 100% identical in their LPS region – PM3 and PM8 (LPS type L3 genotype D), PM48 and PM1434 (LPS type L3 genotype L), and PM72 and PM1441 (LPS type L3 genotype O). In addition, two other pair of isolates and one triple set of isolates showed only a maximum of two base pairs difference between each other – PM18 (LPS type L3

genotype E) and PM1369 (LPS type L3 genotype E'); PM1470 (LPS type L3 genotype J') and PM1474 (LPS type L3 genotype J); PM146 (LPS type L3 genotype F), PM1268 and PM1320 (LPS type L3 genotype F').

Next, the WGS data of the LPS loci was analysed to determine the presence of stop codons or missense mutations, both of which would result in an inactivation of the relevant gene. If the gene in which these significant changes were detected was *hptE*, then the result (as this is the key initial gene in biosynthesis) would result in a full truncated LPS. Significant changes in any of the other genes would result in a shortened or semi-truncated LPS (Figure 4-1). Changes, in comparison with the Pm70 sequence, which did not result in stop codons or missense mutations were recorded as indeterminate (i.e. not clear if the LPS would be semi-truncated or full). The results of this bioinformatics analysis are presented in Table 4-1.

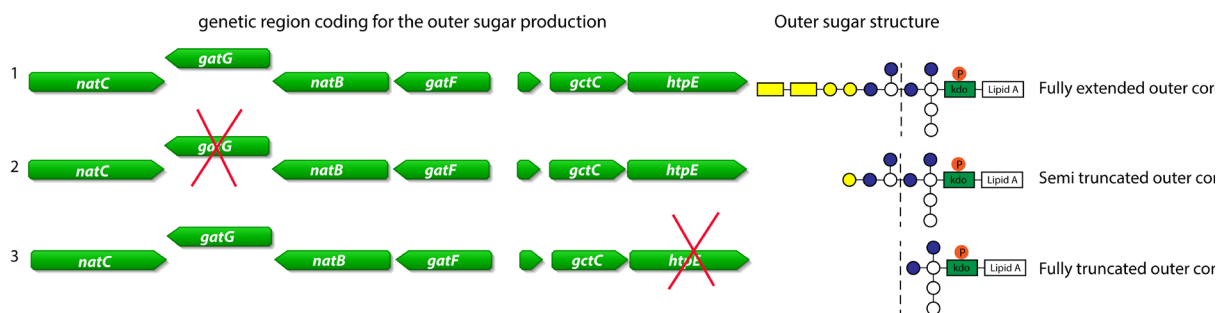


Figure 4-1 Schematic diagram demonstrating how mutations that prevent gene function affect the LPS type L3 outer core structure

4.3.2 Gel electrophoresis patterns

Typical carbohydrate gel patterns are shown in Figure 4-2. The markers are protein molecular weight markers and, as such, cannot predict the size of carbohydrate bands BUT do allow a standardisation of images across gels. After careful analysis, three major distinct patterns could be distinguished. In the first pattern, the dominant band was well below the 7 kDa marker band. Examples of this pattern occur in Lanes C, D and E of Figure 4-2. This pattern was associated with a fully truncated LPS.

Table 4-1 Summary of the results of the genotypic and phenotypic tools developed in this project in terms of their ability to predict the presence of full, semi-truncated or fully truncated LPS, as previously determined by chemical analysis by Harper et al. (2013)

Longest LPS outer core structure detected by Harper et al. (2013) (genetic explanation in brackets)	Isolate	Multiple Glycoforms	This study			
			LPS 3 Genotype*	Major mutations detected [#]	LPS form based on WGS [§]	LPS form based on SDS-PAGE ^{&}
No detectable outer core (Fully Truncated) (<i>htpE</i> disrupted)	PM1075	No	A	<i>Introduction of stop codon</i>	Fully Truncated	Fully Truncated
	PM1120	No	B	<i>Introduction of stop codon</i>	Fully Truncated	Fully Truncated
	PM1258	No	C	<i>1 nucleotide deletion</i>	Fully Truncated	Fully Truncated
Hep-[Glc]-Glc (Semi-Truncated) (PM3 and PM8 – no explanation) (PM18 and PM146 <i>gatF</i> disrupted)	PM3, PM8	Yes	D	<i>W251 to stop codon in gatG</i>	Semi-Truncated	Semi-Truncated
	PM18	No	E	NI	?	Semi-Truncated
	PM146	No	F	<i>SNPs in natC promoter</i>	?	Full LPS
Hep-[Glc]-Glc-Gal (Semi-Truncated) (<i>gatG</i> disrupted)	PM1098	No	G	NI	?	Semi-Truncated
	PM1205	No	H	Insertion of 42 nucleotide	Semi-Truncated	Semi-Truncated
	PM1268, PM1320	Yes	F'	<i>SNPs in natC promoter</i>	?	Semi-Truncated
	PM1439	No	I	14 nucleotide deletion	Semi-Truncated	Semi-Truncated
Hep-[Glc]-Glc-Gal-Gal (Semi-Truncated)	PM1369	Yes	E'	NI	?	Semi-Truncated
	PM1474	Yes	J	NI	?	Semi-Truncated
Hep-[Glc]-Glc-Gal-Gal-GalNAc (Semi-Truncated) (<i>natC</i> disrupted)	PM1	Yes	K	<i>SNPs in natC promoter</i>	?	Full LPS
	PM48, PM1434	Yes	L	<i>Introduction of stop codon</i>	Semi-Truncated	Semi-Truncated
	PM1099	Yes	M	unresolved	unresolved	Full LPS
	PM1103	Yes	N	183 bp deletion	Semi-Truncated	Semi-Truncated
	PM1441	Yes	O	<i>SNPs in natC promoter</i>	?	Full LPS
Hep-[Glc]-Glc-Gal-Gal-GalNAc-GalNAc (Full) (no mutations)	PM72	Yes	O	<i>SNPs in natC promoter</i>	?	Full LPS
	PM1153	Yes	P	<i>SNPs in natC promoter</i>	?	Full LPS
	PM1470	Yes	J'	NI	?	Full LPS

*Genotypes were defined as having an identical base sequence across the entire loci. Isolates with 100% similarity in the entire investigated region were assigned the same LPS L3 genotype. Isolates that showed no more than 2 base pair differences were differentiated from each other by the addition of a prime (e.g. genotypes E and E').

[#] Blue fonts indicate results obtained through WGS, which were not identified by Harper et al. (2013).

[§] Isolates were assigned to four categories of LPS structure – full length, full truncated, semi-truncated or ? (the latter is where the WGS data did not provide clarity).

[&] Red font indicates a major disagreement with the defined gold standard of the full chemical analysis of Harper et al. (2013).

[¶] NI - No major mutation resulting in a frame shift or introduction of a stop codon identified.



Figure 4-2 Nucleotide alignment of the LPS outer core biosynthesis loci and its flanking area of the *P. multocida* reference strain (Pm70), and those of the reference strains for Heddlestone serovars 3 (P1059) and 4 (P1662) and the 23 isolates of Harper et al. (2013), as well as PM1422 (red font) from the investigated chicken meat farm

All these isolates carry LPS outer core biosynthesis loci type L3. However, a lot of variation was found between the isolates carrying this LPS type.

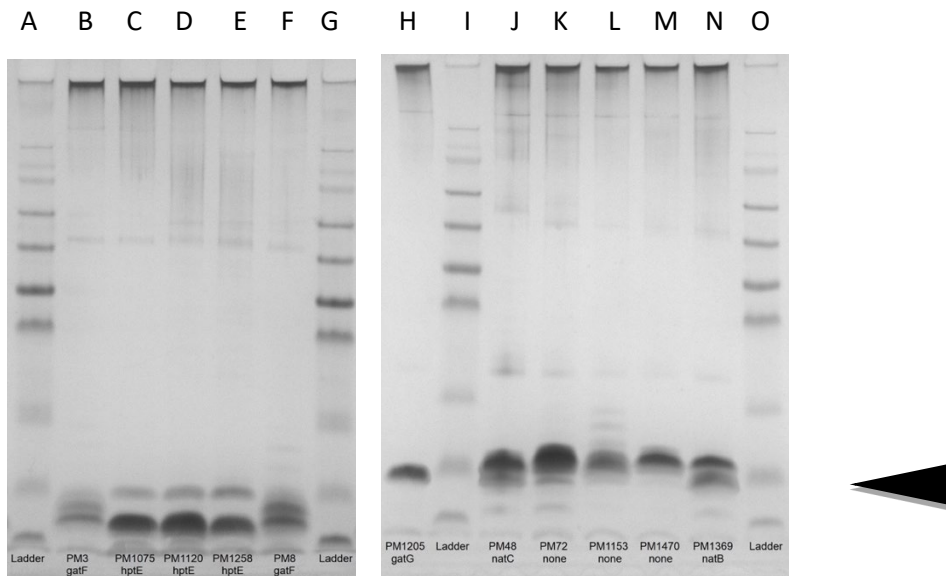


Figure 4-3 Representative SDS-PAGE gels to show the typical carbohydrate separation and silver staining of the LPS from the field isolates included in this study

Lanes A, G, I and O – Pre-stained protein marker (7 kDa marker indicated by arrowhead).

Lane B PM3 (*gatF* mutant).

Lane C PM1075 (*htpE* mutant).

Lane D PM1120 (*htpE* mutant).

Lane E PM1258 (*htpE* mutant).

Lane F PM8 (*gatF* mutant).

Lane H PM1205 (*gatG* mutant).

Lane J PM48 (*natC* mutant).

Lane J PM72 (no mutations).

Lane K PM1153 (no mutations).

Lane L PM1470 (no mutations).

Lane M PM1369 (*natB* mutant).

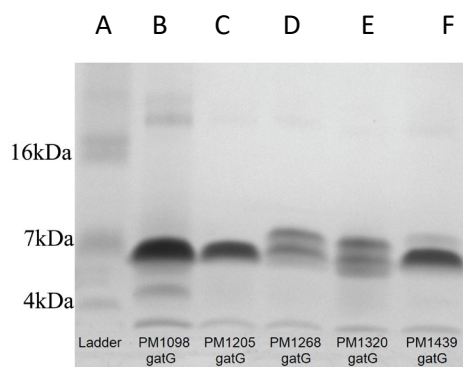


Figure 4-4 Comparisons of the carbohydrate silver stain pattern produced by 5 isolates that the mass spectrometry data suggested harboured a *gatG* mutation

Lane A – Pre-stained protein marker.

Lanes B, C and F showing the typical pattern our data indicates is associated with *gatG* mutation, while Lanes D and E have a different pattern.

The second characteristic pattern was one in which the dominant band occurred at a position slightly before the 7 kDa marker band. This pattern had no other dominant bands at the smaller molecular

weight sizes (see lanes J, K, L and M, Figure 4-2). This pattern represented an isolate with an apparently full LPS structure. Any other pattern was regarded as representing a semi-truncated LPS structure (lanes B, F and N, Figure 4-2).

The observed pattern in lane H (Figure 4-2) was associated with the *gatG* mutation. This pattern is shared by isolates in lanes B, C and F in Figure 4-3, all carrying a major mutation in their *gatG* gene resulting in frame shift. However, those isolates in lane D and E of Figure 4-3 had a different pattern to that expected from a *gatG* frame shift. This disagreement between our data and that obtained from the mass spectrometry led to grouping PM1439 together with the rest of the semi-truncated group in Table 4-1). Using the above criteria, each of the 23 isolates was then assigned to either fully truncated, semi-truncated or full LPS structure, with the results being shown in Table 4-1.

Comparisons of the carbohydrate separation and silver staining between the two isolates from Layer B with those of Harper et al. (2013) demonstrated that PM842 produced a pattern similar to PM1439, both having a 14 bp deletion in their *gatG* gene. It should be mentioned that the LPS of all the ST20/LPS type L3 carrying isolates in the three investigated farms (two layers and one meat) were 100% identical to that of PM1439 from Harper et al. (2013).

The second field isolate, PM1541, produced a pattern similar to PM1120 and PM1258 (Lanes C, D and E in Figure 4-4, respectively). These three isolates all carried a mutation in their *htpE* gene resulting in a frame shift and early introduction of stop codon, albeit at different locations.

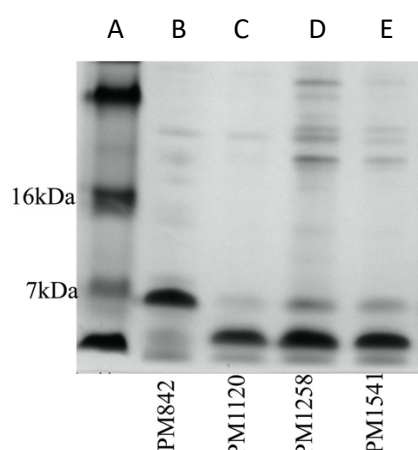


Figure 4-5 Comparisons of the carbohydrate silver stain pattern produced by 2 of the Layer B isolates (PM842 and PM1541) with 2 of those previously analysed by Harper et al. (2013) (PM1120 and PM1258) as described in Chapter 4

Lane A – Pre-stained protein marker.

Lane B shows a typical ST20 *gatG* mutation.

Lanes C, D and E show full truncation due to a *htpE* mutation with isolates in lanes C and D previously analysed by mass spectrometry.

4.3.3 Comparison of WGS/bioinformatic analysis and gel electrophoresis patterns with the results from full chemical analysis

Table 4-1 provides a comparison of the results of the chemical analysis by mass spectrometry performed by Harper et al. (2013) and the two methods developed in this project –

WGS/bioinformatics analysis and SDS-PAGE gel patterns. In making this comparison, it is important to note that the full chemical analysis used by Harper et al. (2013) requires technology and expertise (in mass spectrometry) that is simply beyond the capacity of most laboratories around the world. Hence, while Harper et al. (2013) were able to actually identify the full chemical structure of the isolates (as shown in Table 4-1), both of the techniques we have developed have only been used to assign isolates to either three categories – full structure, semi-truncated structure and fully truncated structure (for the SDS-PAGE gel assay) or the same three plus an obvious *gatG* mutation. In addition to that, the WGS/bioinformatic analysis could also identify an uncertain group for the LPS structure. The additional uncertain group occurred in the WGS/bioinformatic analysis, as isolates in this group showed changes (as compared with the Pm70) that did not result in frame shifts or stop codons or which showed no insertions or deletions. While frame shifts, stop codons, insertions and deletions all clearly result in an inactive gene, the other base pair changes seen in these isolates cannot be clearly identified as causing gene disruption.

The WGS/bioinformatic approach aligned with the full chemical analysis for 12 isolates – three being identified as having fully truncated LPS because of disruptions to the *htpE* gene, and nine as having semi-truncated LPS because of clear disruptions to other genes later in the pathway (Table 4-1). A further 11 isolates could not be allocated to a clear category, with these problematic isolates falling into five of the six chemical structures detected by full chemical analysis (Table 4-1).

The SDS-PAGE gel analysis resulted in an agreement with the full chemical analysis for 19 isolates (three fully truncated, 13 semi-truncated and three with full LPS structure). There were four isolates that appeared to have a full LPS structure in the SDS-PAGE examination, which chemical analysis suggested had a semi-truncated LPS structure (Table 4-1).

We also examined the correlation of LPS genotypes with the results of full chemical analysis. There were six genotypes that had multiple isolates, with four of those genotypes (genotypes D, L and O, and F') each having two isolates with entirely identical LPS outer core loci. As expected, the two genotype D, the two genotype F' and the two genotype L isolates had the same LPS structure (within each genotype) as determined by chemical analysis. In contrast, the two isolates within genotype O with identical LPS loci produced two different longest LPS structures (isolates PM1441 and PM72). The three genotype pairs (E and E', F and F', J and J') only differed by two base pairs at most (highlighted by the addition of a prime to the genotype name). However, despite this very high similarity, all three genotype pairs (E and E', F and F' and J and J') had isolates that the chemical analysis suggested had differing longest LPS structures (Table 4-1).

Importantly, our bioinformatics analysis revealed some potential difficulties in the interpretations of the chemical analysis and the sequence variation data provided by Harper et al. (2013). These insights were made possible because our bioinformatics analysis was based on the entire LPS loci for all 23 isolates used by Harper et al. (2013). In contrast, Harper et al. (2013) only examined the sequence within the single gene that they predicted would explain the LPS structure determined by the chemical analysis. Examples of these additional insights were as follows:

- a) The truncated LPS structure of PM1 was associated by Harper et al. (2013) with deletions and/or SNPs in the *natC* promoter region. However, our study showed that seven more isolates showed 100% similarity to PM1 in their *natC* upstream region. Two of these isolates (PM72 and PM1153), however, had the capacity to produce full length LPS outer core. These results indicate that the shorter LPS produced by PM1 could have resulted from factors other than changes in the *natC* promoter region.

- b) Harper et al. (2013) identified PM72 and PM1441 as having different length LPS structures by chemical analysis with only PM72 producing a full length LPS structure. However, our WGS data revealed that these 2 isolates shared 100% sequence similarity in their LPS outer core biosynthesis loci and the flanking area. Interestingly, the SDS-PAGE gel result suggested that both isolates produced a full length LPS.
- c) Isolates PM3 and PM8 both obtained from turkeys were 100% identical in their LPS outer core biosynthesis loci and the flanking area. This is an expected result as a much earlier study from our group (Blackall et al., 1995) showed that these two isolates were clonal using both phenotypic and early genotypic methods. In the Harper et al. (2013) study, PM3 was shown to have the capacity to produce 2 different outer core glycoforms while PM8 produced 3 different glycoform sizes.
- d) Similarly, PM48 and PM1434 were 100% identical in their LPS outer core biosynthesis loci and the flanking area. In the Harper et al. (2013) study, while both isolates shared the same longest structure, PM48, but not PM1434, was found to produce multiple glycoforms of differing lengths.

4.4 Conclusions

In performing this study, our aim was to compare the knowledge gained from our two differing technologies (WGS/bioinformatic and SDS-PAGE) with that gained from a full chemical analysis when all these methods were applied to examining the LPS structure of *P. multocida* isolates. The importance of an understanding of the LPS structure is driven by the fact that a killed vaccine only provides protection against field isolates with an identical or nearly identical LPS structure (Harper et al., 2016).

The comparison of the information on LPS structure provided by our methods as compared with full chemical analysis was not easy for a number of reasons. The key difficulty is that many isolates of *P. multocida* produce different LPS structures at the same time. This characteristic (multiple glycoforms) was known to occur in 15 of the isolates examined in this study (Table 4-1) (Harper et al. 2013). Hence, in the original study, the chemical analysis data was presented only as the 'longest' structure detected (Harper et al. 2013). It is distinctly possible that the isolates with the potential to produce multiple glycoforms could have been producing different amounts of the various forms. While our study was performed on the same isolates as the Harper et al. (2013) study, our study occurred more than six years later and in a different laboratory setting. Hence, the differences in the results of our work to the Harper et al. (2013) study could be simply due to differing amounts of the multiple glycoforms being produced in the different time-separated studies.

On the basis of accepting the chemical analysis as the 'gold' standard, the SDS-PAGE methodology performed reasonably well – all three isolates that had a fully truncated LPS and all three isolates that had a full LPS were correctly identified. Of the 17 isolates that the chemical method identified as having a semi-truncated LPS structure, 13 were also identified as semi-truncated by SDS-PAGE, with four isolates apparently having a full LPS by SDS-PAGE but a semi-truncated LPS by the chemical analysis (PM1, PM146, PM1099, PM1441). A formal analysis of the performance of the SDS-PAGE method reveals a sensitivity of 80%, a specificity of 100%, a positive predictive value (PPV) of 100%, and a negative predictive value of 43% (see Table 5-1).

Table 4-2 Summarised results of the performance of the SDS-PAGE analysis for LPS truncation as compared with the gold standard chemical analysis of Harper et al. (2013)

SDS-PAGE Result	Chemical Analysis (Harper et al. 2013)
-----------------	--

	Truncated	Full
Truncated	15	0
Full	5	3

Characteristics of the SDS-PAGE analysis			
Sensitivity	Specificity	PPV	NPV
80%	100%	100%	43%

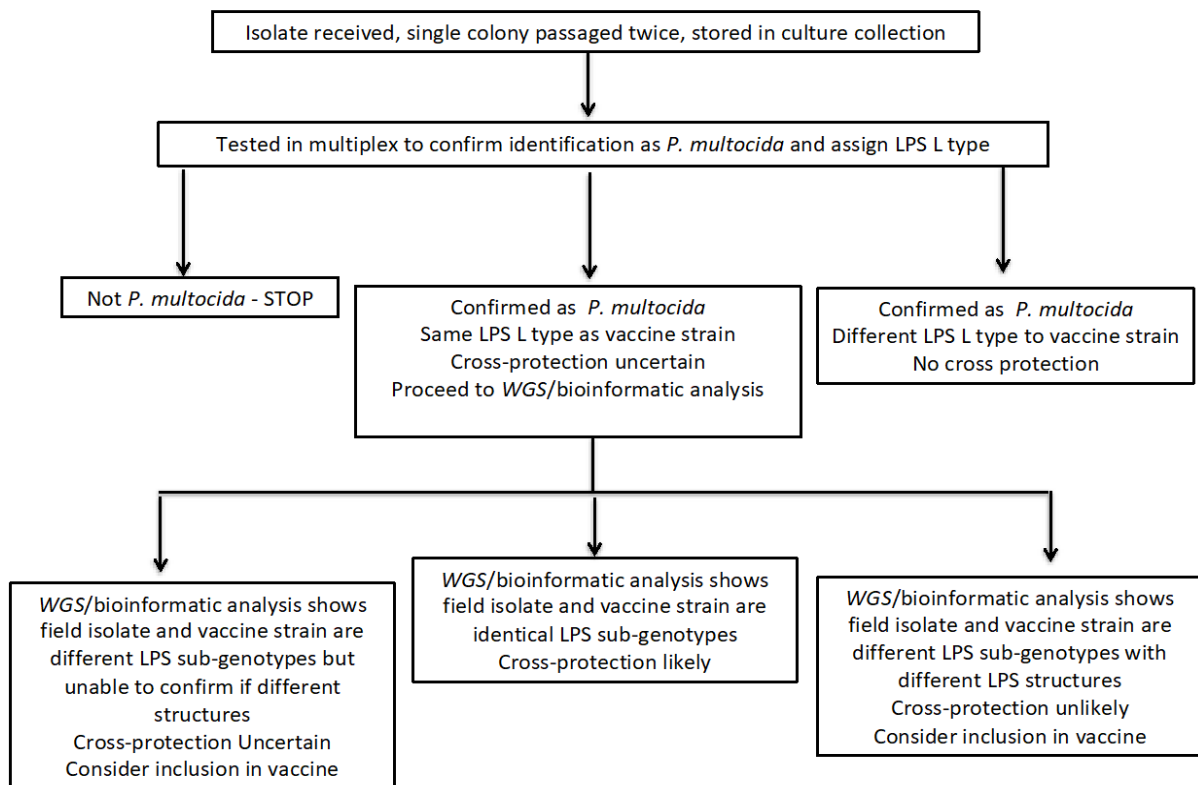
Our WGS/bioinformatics analysis approach was able to recognise the three isolates with a fully truncated LPS and also correctly recognised nine of the 17 isolates that had semi-truncated LPS structures. However, the WGS/bioinformatics analysis could not provide a definitive LPS structure for the remaining 11 isolates (8 of which had a semi-truncated LPS structure, and 3 with a full LPS structure). These 11 problematic isolates all had sequence differences in the LPS loci as compared with the Pm70 reference strain, but it was not possible to determine if these changes would result in gene inactivation. More advanced analysis techniques would be required to determine if the changes detected in these 11 problematic isolates were of significance.

Overall, the combined use of SDS-PAGE and WGS/bioinformatics provides insight into the LPS structure that has not previously been available to the Australian poultry industry. Neither of the techniques can provide the detailed LPS structure of all isolates. Indeed, at this point in time, no laboratory can offer a routine service that allows the detailed LPS structure to be determined, as the technology and expertise required for the mass-spectrometry approach used is simply not available except in a few highly specialised laboratories located outside Australia.

The SDS-PAGE analysis was able to quickly identify (with high specificity and reasonable sensitivity) those isolates that had a full LPS structure and those that had a totally truncated LPS structure. At the moment, the SDS-PAGE technique cannot be used to recognise the degree of LPS truncation. This is a critical issue as isolates that differ in the degree of LPS truncation are not cross-protective in terms of the immunity provided by a killed vaccine (Harper et al. 2016). As well, the interpretation of SDS-PAGE patterns remains a subjective process that requires an experienced technical operator for confident interpretation. Given this need for skilled human interpretation and the inability of the SDS-PAGE technology (in the current form used in our work) to separate the various degrees of LPS structure truncation, we cannot recommend the routine use of this technology to guide vaccine strain selection. It is possible that, with further work, a more informative SDS-PAGE gel technology could be established. This is an attractive future option as SDS-PAGE analysis has both a lower cost and a shorter turnaround time than the other option of WGS/bioinformatics analysis.

At the moment, we would recommend that the optimal current approach to guiding strain selection for fowl cholera vaccines is the WGS/bioinformatics approach. Field isolates should first be confirmed as *P. multocida* and allocated to an LPS type by PCR. If the field isolate and the vaccine strain are different LPS types, there will be no cross-protection from a killed vaccine. For those field isolates that share the same LPS type as the vaccine strain, a further stage is required. At this next stage, the isolate should be subjected to WGS/bioinformatics analysis. If the field isolate and the current vaccine strain are the same LPS type and the same LPS locus subtype, then the isolate and the vaccine strain are cross-protective. If the field isolate and the vaccine strain are different LPS genotypes, there is a possibility that the isolates are not cross-protective. In this case, our recommendation would be to ensure that the LPS genotype of the field challenge strain and the vaccine strain were the same. If the vaccine and the field isolate are the same LPS genotype, there will be cross-protection (assuming an effective vaccination program is in place).

A flow chart showing the recommended typing for the determination of potential cross-protection between a vaccine strain and a recent field isolate is provided below. For this comparison to be effective, the vaccine strain in use must have been fully characterised already OR be characterised along with the field isolate. In this flow chart, the SDS-PAGE assay has not been considered, as we believe that further development of this assay is necessary before the assay can be used with confidence.



5 Comparison of costs for different *P. multocida* typing methods and possible models for provision of typing services to industry

5.1 Introduction

In this Chapter, we provide a current estimated cost for the various typing methods currently available for poultry *P. multocida* isolates. As well, we then propose several models for how these assays could be provided, and conclude with our recommended approach for the provision of these assays.

5.2 Costs

The costs of the various assays available prior to this project, and developed in this project, have been calculated using consumable costs and the relevant salaries for a research technical officer, a bioinformatics specialist and a senior researcher. The exercise has been performed with the intention of showing comparative costs and is not an indication of the costs that any particular institute or organisation would impose. The sections below show the costs for individual steps in the various individual components. A summary table that shows the total costs of the various options for typing is then provided. The operating costs of WGS analysis are based on a batch approach, which requires consolidation of isolates and not a single isolate run.

Initial processing

In this component, an incoming culture is subjected to two serial single colony passages (to ensure purity) and then storage at -80°C for future work. This aspect of the work is standard for all typing methods.

Salary – \$120; Consumables – \$6; Total Cost – \$126

PCR identification and allocation to an LPS type

This component consists of a single multiplex PCR that confirms the identity at the species level and then assigns the isolate to one of the eight LPS types.

Salary – \$180; Consumables – \$2; Total Cost – \$182

rep-PCR identification for genotyping

This component consists of a single rep-PCR that confirms the genotype of a field isolate with a vaccine strain.

Salary – \$240; Consumables – \$2; Total Cost – \$242

MLST typing by conventional PCR and sequencing of amplicons

This component consists of seven individual PCR reactions, a clean-up of each amplicon, transfer of the amplicons to the sequencing facility, sequencing of the amplicons, analysis of the sequencing results, and assignment to an MLST type.

Salary – \$721; Consumables – \$226; Total Cost – \$947

WGS analysis to generate MLST result and LPS sub-genotype

This component consists of a DNA extraction, clean-up of the DNA, costs of a whole genome sequence, and the analysis time to determine the MLST type and the LPS type and sub-type.

Salary – \$440; Consumables – \$205; Total Cost – \$645

SDS-PAGE analysis of LPS patterns

This component consists of an LPS extraction, the running of an SDS-PAGE gel, and a comparison of the resultant pattern with known standard patterns or with a vaccine strain.

Salary – \$240; Consumables – \$25; Total Cost – \$265

Direct comparison of the various typing options

Table 5-1 provides a cost comparison of the various approaches to the typing of field isolates by a range of options. An assessment of the ability of the options to provide guidance for killed vaccination programs is also included.

Table 5-1 Costs and suitability of various typing options to guide fowl cholera vaccination programs*

Typing Option	Cost	Ability to guide killed vaccination programs
Identification and LPS typing	\$308	Limited ability as LPS types are known not to be cross-protective
Identification, LPS typing and rep-PCR	\$550	Limited ability as rep-PCR has been shown to fail to recognise different LPS types and this cannot predict cross-protection (see Chapter 3)
Identification, LPS typing and conventional MLST	\$1,255	Limited ability as MLST has been shown to fail to recognise different LPS genotypes and thus cannot predict cross-protection (see Chapter 3)
Identification, LPS typing and WGS to generate MLST and LPS sub-genotype	\$953	Optimal current technology – if vaccines are matched to the LPS sub-genotypes, then cross-protection will occur. Weakness is that some isolates that have different LPS sub-genotypes will be cross-protective but current WGS technologies cannot predict this situation. Hence, this approach will result in more vaccine changes than may be absolutely necessary.
Identification, LPS typing and SDS-PAGE	\$573	Unable to recognise the LPS structures in detail – cannot currently provide definitive knowledge on cross-protection

*The costs provided here are intended to allow comparison of the various options. The costs are not a formal quote, and are not predictive of the charges that any organisation may impose for providing these services.

Recommended typing option

As outlined in the work described in Chapters 2, 3 and 4, none of the suite of assays developed in this project (and previous industry funded projects) is fully able to accurately identify if a particular killed vaccine and a particular field isolate are cross-protective.

Currently, as shown in Chapters 2, 3 and 4, the best available option is a progression of two sets of assays – first confirmation of identity as *P. multocida* by PCR and then allocation to an LPS type, again by PCR. This flow of work has been provided as a flow chart in Chapter 4. In the following text, we set that work in the context of the relative costs.

If the isolate and the vaccine strain belong to different LPS types, then there is no cross-protection and the total estimated cost of the work is \$308.

If the isolate and the vaccine strain are in the same LPS type, then WGS and bioinformatics analysis is necessary to determine the LPS sub-genotype. If the isolate and vaccine are the same sub-genotype, then cross-protection is expected. If the isolate and the vaccine strain are of different sub-genotypes, and the bioinformatics analysis predicts that the sub-genotypes are likely to be result in different LPS structures (e.g. a stop codon in *hptE* in the field isolate but no such stop codon in the vaccine strain), then cross-protection will probably not occur and a change in vaccine composition should be considered. If the isolate and the vaccine strain are of different sub-genotypes, and the bioinformatics analysis cannot predict that the sub-genotypes are likely to be result in different LPS structures, then a prediction about cross-protection is not possible. In this situation, a conservative approach would be to change the vaccine to ensure a match with the LPS sub-genotype of the field isolate. This conservative approach is supported by the work described in Chapters 2, 3 and 4. Overall, the full flow of work to this LPS sub-genotyping by WGS/bioinformatic has an estimated cost of \$953.

5.3 Ongoing provision of the typing tools developed in this and previous projects

In the initial project documentation, the research team flagged that the assays developed in this project would be moved across to the ‘user pays’ service provided by our research group. This is the model we have used since the initial poultry industry research project commenced in our group in 1979. Given our ability to maintain this approach for over 40 years, we believed that this was a realistic option.

However, there has been a recent significant change in our ability to attract industry research funds. By November 2020 (or slightly latter due to COVID-19 imposed delays), our research group will have no industry research funds from the Australian chicken meat, the egg or the pig industries (for the first time in over 40 years). This lack of funding means that the group has no capacity to support technical and professional staff who are essential for the performance of the user pays service.

Using reserves, we will provide an ongoing access to these services (after the COVID-19 restrictions are lifted) till January 2021. By that time, our reserves will have been depleted and we will have to close access to the ‘user pays’ service.

As all our of methods and tools have been fully published, it is possible that other Australian laboratories may choose to adopt our technologies and provide these services.

6 General conclusion and recommendations

In this section, highlights of the results each of the Chapters, which align with the project objectives, are provided.

In Chapter 1, we were able to demonstrate our capacity to perform WGS and bioinformatic analysis on a small subset of *P. multocida* isolates. We established a successful collaboration with the laboratory of Associate Professor Scott Beatson (The University of Queensland), which enabled us to have full access to their servers and computational support. We established and then used WGS and bioinformatics to show a set of six isolates were ST20 and LPS type L3. These results aligned with the MLST and LPS type performed via conventional PCR and amplicon sequencing.

In Chapter 2, we examined the genetic diversity occurring within a fowl cholera outbreak on free range layer Farm A. We examined 37 isolates from outbreaks of fowl cholera from 1994 to 2016. The isolates included multiple isolates from the same chicken, and also isolates from different chickens within an outbreak as well as the vaccine strains. The *in silico* LPS typing clearly showed that all isolates across the 22 years were carrying LPS type L3, although they belonged to two different STs, 20 and 9. While the ST20 isolates (earlier outbreaks) were more closely related to each other, the ST9 isolates appeared to be more diverse. This was also reflected in the fact that ST9 isolates in this farm had the capacity to carry three different LPS type L3 subtypes, with each subtype having the capacity to produce a different LPS outer structure based on the genomic data. Our results also showed that isolates with different mutations in their LPS type L3 loci, could produce the same rep-PCR pattern and belong to the same ST. Hence even a combination of rep-PCR fingerprinting, MLST and LPS multiplex can still be misleading in guiding vaccine selection. More importantly, the results of this report demonstrated for the first time that *P. multocida* isolates from the same outbreak can produce different LPS structure. This indicates the importance of analysis of more than one isolate per outbreak, especially if the outbreak is happening in more than one shed of the farm. This lack of discrimination capacity by the typing schemes in-use – rep-PCR, MLST, and LPS multiplex PCR – meant that some of the outbreaks of fowl cholera on farms might not be understood due to an antigenically very different challenge strain.

In Chapter 3, we looked at a layer Farm B, which was a large complex with multiple farms some distance from each other. The complex has been using the same autogenous vaccine for all four farms at rearing. However, our results showed that although all the investigated *P. multocida* isolates were carrying LPS type L3, each farm had its own strains with different LPS type L3 subtypes. Some of these isolates from different farms were carrying missense mutation in the same gene and had the capacity to produce the same outer structure (same phenotype), although the mutation was in different positions of the LPS region (different genotype).

In Chapter 4, we compared full chemical analysis by mass spectrometry, WGS/bioinformatic analysis and SDS-PAGE examination in terms of an ability to describe the LPS structure of isolates of *P. multocida*. We found that neither of our methods alone could be used to confidently identify the LPS structure of a *P. multocida* isolate. This is not an unexpected finding, as detailed structural analysis is not possible without sophisticated mass spectrometry methods; methods currently beyond the capacity of diagnostic laboratories. Promising results were provided by the SDS-PAGE technology. However, as the SDS-PAGE method had an ability to definitely recognise only full truncated and full structures and not the partial structures, the approach cannot be recommended currently. However, a combination of WGS/bioinformatic and SDS-PAGE methods would be a far better approach for providing information on the LPS structure within a given LPS type. Hence we recommend that further research be undertaken

towards improving the SDS-PAGE method for carbohydrate analysis.

In Chapter 5, we examined the comparative cost structures of the various typing options for *P. multocida*. This analysis demonstrated that the WGS/bioinformatic approach was a cost effective method that generated the current best possible information to guide vaccination programs. As well, if supported by an improved SDS-PAGE methodology, a powerful set of tools would be available to the industry. A step-by-step process for optimal use of the typing tools developed in this and previous industry funded projects was provided. While the planned outcome of this project was to move the typing assays across to the user pays service operated by the research group, the failure of the research group to attract any industry funds for research beyond November 2020 means that this planned outcome is no longer possible. A temporary user pays service will be maintained till the end of December 2020. It is possible that other laboratories will be able to provide these typing services, as all of the methods are fully described in our formal peer review publications.

Our concept of using WGS to explore the LPS genetic and the resulting phenotypic diversity is highly likely to be broadly applicable across different LPS genotypes and different host species. As examples we have already used this same approach to investigate pasteurellosis outbreaks in different hosts including LPS type L1 carrying isolates in meat chickens (Omaleki et al. 2020b), LPS types L3, L4 and L6 carrying isolates in captive marsupials (Omaleki et al. 2002a) and LPS type L3 carrying isolates in pinnipeds (Crawford et al. 2019). In addition to this work on field isolates, we have briefly looked at the LPS type L1 outer core biosynthesis loci, as well as the potential presence of the *petG* gene, reported to be responsible for addition PEtn to the outer core of the *P. multocida* reference strain X73 (Omaleki et al. 2002b).

7 References

- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19, 455-477.
- Blackall, P.J., Hofacre, C.L. 2020. Fowl Cholera In: Swayne, D.E., Boulianne, M., Logue, C.M., McDougald, L.R., Nair, V.L., Suarez, D.L. (Eds.) *Diseases of Poultry*. John Wiley and Sons Inc, Hoboken, NJ, USA, 831-845.
- Blackall, P.J., Mifflin, J.K., 2000. Identification and typing of *Pasteurella multocida* : a review *Avian Pathol.* 29, 271-287.
- Blackall, P.J., Pahoff, J.L., Marks, D., Fegan, N., Morrow, C.J., 1995. Characterisation of *Pasteurella multocida* isolated from fowl cholera outbreaks on turkey farms. *Aust Vet J* 72, 135-138.
- Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114-2120.
- Brogden, K.A., Rhoades, K.R., Heddleston, K.L., 1978. A new serotype of *Pasteurella multocida* associated with fowl cholera. *Avian Dis.* 22, 185-190.
- Carver, T., Harris, S.R., Berriman, M., Parkhill, J., McQuillan, J.A., 2011. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* 28, 464-469.
- Crawford, R.L., Blyde, D., Blackall, P.J., Forde, B.M., Beatson, S.A., Harris, L.M., Turni, C., Omaleki, L., 2019. Novel insights into pasteurellosis in captive pinnipeds. *Vet. Microbiol.* 231, 232-237.
- Gunawardana, G.A., Townsend, K.M., Frost, A.J., 2000. Molecular characterisation of avian *Pasteurella multocida* isolates from Australia and Vietnam by REP-PCR and PFGE. *Vet. Microbiol.* 72, 97-109.
- Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072-1075.
- Hadfield, J., Croucher, N.J., Goater, R.J., Abudahab, K., Aanensen, D.M., Harris, S.R., 2017. Phandango: an interactive viewer for bacterial population genomics. *Bioinformatics* 34, 292-293.
- Harper, M., John, M., Edmunds, M., Wright, A., Ford, M., Turni, C., Blackall, P.J., Cox, A., Adler, B., Boyce, J.D., 2016. Protective efficacy afforded by live *Pasteurella multocida* vaccines in chickens is independent of lipopolysaccharide outer core structure. *Vaccine* 34, 1696-1703.
- Harper, M., John, M., Turni, C., Edmunds, M., St Michael, F., Adler, B., Blackall, P.J., Cox, A.D., Boyce, J.D., 2015. Development of a rapid multiplex PCR assay to genotype *Pasteurella multocida* strains by use of the lipopolysaccharide outer core biosynthesis locus. *J. Clin. Microbiol.* 53, 477-485.
- Harper, M., St Michael, F., John, M., Vinogradov, E., Steen, J.A., van Dorsten, L., Steen, J.A., Turni, C., Blackall, P.J., Adler, B., Cox, A.D., Boyce, J.D., 2013. *Pasteurella multocida* Heddleston serovar 3 and 4 strains share a common lipopolysaccharide biosynthesis locus but display both inter- and intrastrain lipopolysaccharide heterogeneity. *J Bacteriol* 195, 4854-4864.
- Heddleston, K.L., Gallagher, J.E., Rebers, P.A., 1972. Fowl cholera: gel diffusion precipitin test for serotyping *Pasteruella multocida* from avian species. *Avian Dis.* 16, 925-936.
- Hitchcock, P.J., Brown, T.M., 1983. Morphological heterogeneity among *Salmonella lipopolysaccharide* chemotypes in silver-stained polyacrylamide gels. *Journal of bacteriology* 154, 269-277.
- Imperiali, B., 2019. Bacterial carbohydrate diversity — a Brave New World. *Current Opinion in Chemical Biology* 53, 1-8.
- Jolley, K.A., Chan, M.S., Maiden, M.C., 2004. mlstdbNet - distributed multi-locus sequence typing (MLST) databases. *BMC Bioinformatics* 5, 86.

- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947-2948.
- Nascimento, M., Sousa, A., Ramirez, M., Francisco, A.P., Carriço, J.A., Vaz, C., 2017. PHYLOViZ 2.0: providing scalable data integration and visualization for multiple phylogenetic inference methods. *Bioinformatics* 33, 128-129.
- Omaleki, L., Beatson, S.A., Thomrongsuwanakij, T., Blackall, P.J., Buller, N., Hair, S., Vitali, S., Wallace, A., Turni, C., Forde, B.M., 2020a. Phase variation in *latB* gene associated with a fatal *Pasteurella multocida* outbreak in captive squirrel gliders. *Vet. Microbiol.* 243.
- Omaleki, L., Blackall, P.J., Cuddihy, T., Beatson, S.A., Forde, B.M., Turni, C., 2020b. Using genomics to understand inter- and intra- outbreak diversity of *Pasteurella multocida* isolates associated with fowl cholera in meat chickens. *Microb Genom* 6.
- Salipante, S.J., Hall, B.G., 2011. Inadequacies of Minimum Spanning Trees in Molecular Epidemiology. *J. Clin. Microbiol.* 49, 3568-3575.
- Singh, R., Blackall, P.J., Remington, B., Turni, C., 2013. Studies on the presence and persistence of *Pasteurella multocida* serovars and genotypes in fowl cholera outbreaks. *Avian Pathol.* 42, 581-585.
- Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312-1313.
- Subaaharan, S., Blackall, L.L., Blackall, P.J., 2010. Development of a multi-locus sequence typing scheme for avian isolates of *Pasteurella multocida*. *Vet. Microbiol.* 141, 354-361.
- Treangen, T.J., Ondov, B.D., Koren, S., Phillippy, A.M., 2014. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol.* 15, 524.
- Turni, C., Singh, R., Blackall, P.J., 2018. Genotypic diversity of *Pasteurella multocida* isolates from pigs and poultry in Australia. *Aust Vet J* 96, 390-394.
- Wick, R.R., Schultz, M.B., Zobel, J., Holt, K.E., 2015. Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics* 31, 3350-3352.
- Wood, D.E., Salzberg, S.L., 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* 15, R46.
- Zhang, P., Fegan, N., Fraser, I., Duffy, P., Bowles, R.E., Gordon, A., Ketterer, P.J., Shinwari, W., Blackall, P.J., 2004. Molecular epidemiology of two fowl cholera outbreaks on a free-range chicken layer farm. *J. Vet. Diagn. Invest.* 16, 458-460.

8 Plain English Summary

Project Title:	Clouds and <i>Pasteurella multocida</i>
Australian Eggs Limited Project No.	1BS702UQ
Researchers Involved	L. Omaleki, P. Blackall, C. Turni
Organisations Involved	QAAFI, The University of Queensland, GPO Box 267, Brisbane Qld 4001
Phone	0406 411 712 (L. Omaleki) 07 3255 4296 (C. Turni)
Fax	NA
Email	l.omal eki@uq.edu.au c.turni1@uq.edu.au
Objectives	1. Establish a pipeline to allow effective rapid whole genome sequencing and then full bio-informatic analysis to be able to understand and measure genetic diversity within a disease outbreak. 2. Use the above pipeline and analysis approach on a well characterised and extensive collection of <i>Pasteurella multocida</i> from two prolonged fowl cholera outbreaks – one involving an organic broiler operation and the other involving a free range layer operation (including the various killed vaccine strains used over the period of the two outbreaks). 3. Use the sequencing and bioinformatics analysis to understand the contribution of clouds (genetic diversity) to evasion of vaccine protection and to thus assist in ensuring a match between vaccine protection and field challenge. 4. Transfer the established technologies to the user pays service provided by the Microbiology Research Group for ongoing use by the poultry industry beyond the life of the project.
Background	Fowl cholera is one of the major threats to the fast-growing chicken meat industry in Australia. Currently, killed <i>P. multocida</i> vaccines are specifically formulated and produced for individual farms using the bacteria that has been previously found on the farm in prior disease outbreaks. However, as the industry shifted to free range/organic production systems, the chickens suffered increasing exposure to wildlife and predators, both carrying different types of <i>P. multocida</i>. This has resulted in a significant increase of the disease, with more diversity observed in <i>P. multocida</i> causing fowl cholera outbreaks. This variation in the bacteria has caused major problems in terms of vaccine efficacy, with problems linked to vaccination failure impacting this fast growing industry.
Research	This project established and validated a flow of work that involved whole genome sequencing and subsequent analysis. This validated workflow was then used on 72 isolates of <i>P. multocida</i> that had been previously obtained

from cases of fowl cholera on two free range layer farms. As well, the vaccine strains used on the farms were examined.

In our genomic work, we focused on the total genome as well as that part of the genome that codes for the carbohydrates that comprise the outer surface of *P. multocida*. This outer surface is called the lipopolysaccharide (LPS) and is a critical issue in the protection provided by killed fowl cholera vaccines. A killed vaccine strain provides protection only against field isolates that have the same carbohydrate structures – with even a single sugar change resulting in a loss of protection.

As well as the genomic work, we evaluated the use of a method to visualise the carbohydrates in the LPS by a gel electrophoresis technology termed SDS-PAGE.

Outcomes

This project firstly established and validated a flow of work that involved whole genome sequencing and bioinformatics analysis.

Our genomic analysis of the 75 field isolates demonstrated a high degree of diversity in the LPS type L3. We were able to clearly identify when an outbreak was as a result of introduction of a new strain or the persistence of the previous strains on the farm.

Importantly, we identified that the currently in-use DNA fingerprinting techniques, rep-PCR and MLST, can fail to identify the variations in the carbohydrates in the LPS type L3 isolates. As the protection afforded by killed vaccines is limited to isolates with identical LPS structures to the vaccine strains, this inability to detect LPS structure differences is a major failing of both rep-PCR and MLST.

Our results clearly demonstrated that each individual farm within a large complex may have a different *P. multocida* population with completely different LPS outer structure. This enhances the need to treat each farm individually or have a multi-valent vaccine containing strains expressing all the different LPS structures found across the farms of the complex.

We found that our genomic analysis as well as a supporting phenotypic method based on gel electrophoresis are capable of differentiating between some forms of truncations in the LPS type L3 outer structure. However, neither of our methods could be used to confidently identify all of the various truncations of LPS structure of *P. multocida* isolates at this stage. We have provided a recommended workflow for the use of the typing technologies developed in this project (and in previous industry funded projects).

In the final stage of our project, we compared the relative cost of the various typing assays developed in this project (as well as those we have developed in previous projects). Both this costing exercise, and the prior sections on the depth and quality of the information provided by each method, will help guide industry in deciding which tools provide the relevant information, and the relative cost of that tool.

Implications

Our work has shown for the first time that there is high inter- and intra-chicken genetic diversity in *P. multocida* isolates during an outbreak of fowl

cholera. Importantly, it demonstrates the benefit of a genomics/bioinformatics analysis approach as a means of providing detailed information on the possible structure of the outer LPS. As killed fowl cholera vaccines only provide protection against strains with an outer core LPS that is identical or nearly identical, our approach is the optimal technology that can guide strain selection for killed autogenous fowl cholera vaccine. Importantly, our genomics/bioinformatics approach was able to explain why the outbreaks of fowl cholera on the study farms were not controlled by vaccination.

Our study clearly demonstrates the advantages of genomic analysis over the conventional PCR based approaches in providing clear insights in terms of the linkage of isolates within and between outbreaks.

Recent failures to attract ongoing research funding have prevented us from providing a firm plan to support the ongoing provision of these *P. multocida* typing tools to industry via a 'user pays' service provided by our research group. In the absence of industry funding, we have constructed a temporary arrangement that will allow the services to be provided for the rest of 2020. Access to the technologies from 2021 onwards is doubtful. It is possible that other Australian laboratories will use the formal publications describing our work to provide access to the assays.

Key Words

fowl cholera, *Pasteurella multocida*, whole genome sequencing, bioinformatics, vaccines

Publications

Crawford, R.L., Blyde, D., Blackall, P.J., Forde, B.M., Beatson, S.A., Harris, L.M., Turni, C. and Omaleki, L. (2019) Novel insights into pasteurellosis in captive pinnipeds. *Vet Microbiol* 231, 232-237.

Omaleki, L., Beatson, S.A., Thomrongsuwannakij, T., Blackall, P.J., Buller, N.B., Hair, S.D., Vitali, S.D., Wallace, A.M., Turni, C. and Forde, B. (2020) Phase variation in *latB* associated with a fatal *Pasteurella multocida* outbreak in captive squirrel gliders. *Vet Microbiol* 243, 108612.

Omaleki, L., Blackall, P.J., Cuddihy, T., Beatson, S.A., Forde, B.M. and Turni, C. (2020) Using genomics to understand inter- and intra-outbreak diversity of *Pasteurella multocida* isolates associated with fowl cholera in meat chickens. *Microbial Genomics* DOI 10.1099/mgen.0.000346.

Sinclair, H.A., Chapman, P., Omaleki, L., Bergh, H., Turni, C., Blackall, P.J., Papacostas, L., Braslins, P., Sowden, D. and Nimmo, G.R. (2019) Identification of *Lonepinella* sp. in koala bite wound infections, Queensland, Australia. *Emerg Inf Dis* 25, 153-156.