Molecular typing of *Salmonella enterica* serovars of significance in Australia using MLVA and bacteriophage genes

Chun Chun Young BHSc (Laboratory Medicine) (Hons.)

Discipline of Microbiology and Immunology

School of Molecular and Biomedical Science

University of Adelaide

A thesis submitted to the University of Adelaide for the degree of

Doctor of Philosophy

November 2012

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

aa	Amino acid
ATCC	American type culture collection
AFLP	Amplified fragment length polymorphism
AGRF	Australian Genomic Research Facility
AIDS	Acquired immune deficiency syndrome
ASRC	Australian *Salmonella* Reference Centre
bp; kb	Base pair(s); kilobase(s)
BHI	Brain heart infusion
BSA	Bovine serum albumin
CHEF	Clamped homogeneous electric field
DI	Simpson's index of diversity
DNA	Deoxyribonucleic acid
DOP-PCR	Degenerate-oligonucleotide primed – PCR
dNTP	Deoxyribobucleic acid
dsDNA	Double-stranded DNA
DT	Definitive type
EDTA	Ethylene-diaminetretra-acetic acid
g; mg; µg	gram; milligram; microgram
HBA	Columbia horse blood agar
IMVS	Institute of Medical and Veterinary Science
IS	Insertion sequence
l; ml; µl	Litre(s); millilitre(s); microlitre(s)
LB	Luria Bertani
LMP	Low melting point
M; mM; µM	Molar; millimolar; micromolar
MAPLT	Multiple amplification of prophage locus typing
MH	Mueller Hinton
MLEE	Multilocus enzyme electrophoresis
MLST	Multilocus sequencing typing
MLVA	Multiple-locus variable-number tandem repeat analysis
MOI	Multiplicity of infection
mPCR/RLB	Multiplex PCR-based reverse line blot hybridisation
MPU	Media Production Unit
n/a	Not applicable

NCTC	National collection of type cultures
NGS	Next generation sequencing
nt	Nucleotide
O/B	Outbreak
ORFs	Open reading frames
O/S	Overseas
PCR	Polymerase chain reaction
PFGE	Pulsed-field gel electrophoresis
pfu	Plaque forming units
PT	Phage type
RAPD	Randomly amplified polymorphic DNA
RDNC	Reacts does not conform
Rep-PCR	Repetitive element PCR
RFLP	Restriction fragment length polymorphism
SNP	Single-nucleotide polymorphism
SPI	*Salmonella* pathogenicity island
TAE	Tris- Acetate-EDTA
TE	Tris-EDTA
Tris	(hydroxymethyl) aminomethane
TRF	Tandem Repeat Finder
TSI	Triple sugar iron
TTSS	Type III secretion system
UN	Untypable
UPGMA	Unweighted-pair group method with arithmetic averages
VNTR	Variable-number tandem repeat
WGS	Whole genome sequencing
v/v	Volume per volume
w/v	Weight per volume
XLD	Xylose-lysine-desoxycholate

DECLARATION

I Chun Chun Young declare that the work described herein contains no material that has been previously submitted for the award of any degree or diploma in any university and to the best of my knowledge and belief contains no material previously published or written by another person, except where due reference is made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968. I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the library catalogue, and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Chun Chun Young
29th November 2012

ACKNOWLEDGMENTS

First of all I would like to acknowledge my supervisors Assoc. Professor Michael Heuzenroeder and Professor Mary Barton for taking me in and believing in me to undertake this PhD study. I would particularly like to extend my gratitude to the post-doc in our lab Dr Ian Ross for his help and time in having discussions on either the technical or theoretical part of my study. Thank you for also spending time and your patience in critically reading my thesis. Your red-pen writing all over my thesis draft has made it so much more colourful than just the boring black and white!

During my PhD study, I have met many scientists who have provided help that contribute to the completion of this thesis. Thanks to Ms Dianne Davos and Ms Helen Hocking and the staff from the Australian *Salmonella* Reference Centre, SA Pathology for providing phage typed strains for this study and performing serotyping on the doubtful isolates. Thanks to Dr Jane Arthur, Viral Epidemiology Lab, MID, SA Pathology who has helped with the DOP-PCR procedure, to Dr Mark Corbett, Neurogenetics Research Lab, WCH, and Ms Kate Hodgson, School of Pharmacy and Medical Sciences, UniSA for having discussions with me when I was clueless how to start analysing and annotating my sequenced phage genomes.

A sincere thank you also goes to everyone in Public Health Unit, MID, SA Pathology: Dr Robyn Doyle, Dr Wendy Hart, Mr Rolf Wise and Mr Allan Goodwin who have made my study more endurable for I am surrounded by a group of scientists who are not only the experts in molecular microbiology, but are also friendly, supportive and encouraging. In particular, I would like to thank Robyn for the helpful discussions to troubleshoot my failed cloning experiments and phage DNA extractions, to Rolf who helped when I have questions about BioNumerics. Thanks also to my fellow students in the lab, Dr Sophia Tan who helped me settling into the lab, and to Mr Geordie Morgan for his regular visits to my office in the WCH lab, taking me to coffee breaks when I was overwhelmed from thesis writing (or not). Special thank also goes to Ms Ena Ribic from the diagnostic microbiology lab, WCH for her patience and time listening to my blurb when I was writing in our shared office. Thanks to Ms Min Yan Teh from Dr Morona's lab, University of Adelaide for being a great fellow student and a friend who let me ask millions of questions when dealing with university documentations.

Last but definitely not the least, I would like to thank my caring and understanding parents for looking after me so well and for always supporting me to do things I like including this PhD study. Thanks also to my brother and sis-in-law for keeping me in their prayers, to my lovely niece and nephew for giving me laughter and making me do exercise (running after them). There are also my dear friends who I wish to say thank you to. Thanks to Kelly who has not only encouraged me as a friend, but also has helped by suggesting possible solutions to my experimental problems, to Jen and Lisa for always making time to have de-stress sessions with me, to Yvonne for reminding me to keep going and believing in myself. I wish to particularly say a big thank to Anna who has been a great friend supporting me and generously lending me ears listening to whatever I have to say even at the ridiculously late hours. Thank you for genuinely believing in my ability and giving me positivity whenever this PhD seemed impossible to complete.

PUBLICATIONS

Ross, I. R., C. C. Young and M. W. Heuzenroeder. 2012. New Options for Rapid Typing of *Salmonella enterica* Serovars for Outbreak Investigation, *Salmonella* - A Diversified Superbug, Yashwant Kumar (Ed.), InTech, Available from: http://www.intechopen.com/books/salmonella-a-diversified-superbug/new-options-for-rapid-typing-of-salmonella-enterica-serovars-for-outbreak-investigation

Young, C. C., I. R. Ross and M. W. Heuzenroeder. 2012. A new methodology for differentiation and typing of closely related *Salmonella enterica* serovar Heidelberg isolates, *Curr Microbiol,* DOI: 10.1007/s00284-012-0179-3

SUMMARY

Salmonella enterica subspecies *enterica* is the leading causes of food-borne infections and outbreaks worldwide and are therefore routinely monitored in many countries. Both epidemiological surveillance and tracking outbreaks require typing systems to characterise bacterial isolates. Pulsed-field gel electrophoresis (PFGE) is the most widely applied molecular method for *Salmonella* outbreak epidemiological typing despite that the method being relatively labour-intensive, generating restriction patterns that are subjective in nature and can be difficult to analyse. Furthermore, the recent emergence of homogeneous *Salmonella* phage groups (e.g. *S.* Typhimurium DT104) has hampered the discriminatory power of PFGE in outbreak settings.

The combination of these factors has led to the exploitation of alternative methods. Two of these emerging typing methods are multiple-locus variable-tandem repeat analysis (MLVA) and multiple amplification of prophage locus typing (MAPLT) that have been developed for the globally significant serovars: *S.* Typhimurium and *S.* Enteritidis. In this study, MLVA and MAPLT were being further investigated for use in other *Salmonella* serovars that are of significance to Australia. These serovars include *S.* Virchow, *S.* Bovismorbificans and *S.* Heidelberg. Overall, the developed MAPLT schemes demonstrated a comparable differentiating ability with PFGE for the serovars as a whole. However development of discriminative MLVA methods in the study was not always successful using the published MLVA loci, therefore prompting the need of examining the MLVA loci from the complete genomes of the serovars.

Differentiation within predominating phage types of the studied serovars was also examined. Both MAPLT and PFGE displayed similar levels of intra-phage type differentiation suggesting the usefulness for outbreak investigations. However strains were separated optimally when using MAPLT in association with discriminative MLVA primers. The composite assay of MAPLT/MLVA for *S.* Bovismorbificans and *S.* Virchow were further examined in the retrospective outbreak studies. It was observed that outbreak isolates generally showed identical or highly similar MAPLT/MLVA profiles (one locus difference). The results suggested that MAPLT / MLVA could confirm the close genetic relationship between outbreak isolates. However result interpretation guidelines might need to be established individually for each serovar due to the differences in the relative genetic similarity displayed within each serovar. It is expected that further enhancement of the composite assays could easily be carried out even when the complete genomic data of the *Salmonella* serovars are lacking. It is an advantage that MAPLT could be improved through using a DOP-PCR (degenerate oligonucleotide primed-PCR) procedure to detect additional prophage loci for further strain differentiation. Overall, the

study suggested that composite assays can also be developed for fine levels of differentiation of other *Salmonella* serovars to meet local epidemiological needs when required.

Development of MAPLT involved genetic characterisation of phages residing in the studied serovars. It was noted that DNA elements shared between phages Gifsy-1 and Gifsy-2 were frequently detected from these serovars. Complete genome analysis were subsequently performed on a *S.* Virchow phage (PV10) and a *S.* Heidelberg phage (PH03), which demonstrated their high DNA similarity to Gifsy-1 and Gifsy-2 respectively. Further analysis demonstrated that Gifsys-related phages are possibly widespread within *Salmonella*. Since the phages were observed to carry different sets of virulence genes between serovars, this study hypothesised that Gifsy-related phages may have a significant role in shaping the epidemiology of *Salmonella*.