

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**Analysis of the *Helicobacter mustelae*
Surface Ring (*hsr*) Locus**

A thesis presented in partial fulfilment of
the requirement for the degree of
Master of Science in Biological Sciences
at Massey University, Palmerston North,
New Zealand

Natasha Talei Forester

2001

ABSTRACT

The DNA upstream of the gene encoding the *Helicobacter mustelae* surface ring (Hsr) protein of the ferret gastric pathogen *Helicobacter mustelae* was shown to contain several uncharacterised *hsr*-like repeat sequences in a total of 12 kb of *hsr*-related DNA, including the *hsr* gene (the *hsr* locus). The primary objective of this study was to analyse the *hsr* locus of *H. mustelae* strain 4298, in particular, to determine the extent and possible function of the *hsr*-related DNA repeat sequences.

H. mustelae was isolated from the stomachs of six New Zealand ferrets. This study represents the first successful isolation of *H. mustelae* from ferret stomachs from at least two geographically distinct locations in New Zealand. The Hsr proteins of the different *H. mustelae* strains exhibited variability in protein size and reactivity to anti-Hsr antisera. The DNA sequence of the strain 4298 15-kb *hsr* locus was completed and analysed for DNA repeats. There were 343 perfectly repeated sequences 12 – 741 bp in length, with up to 11 copies of each. Within the *hsr* gene of strain 4298, a 2.4 kb repeat region, the variable repeat region, was defined. The flanking *hsr*-related sequences were equally distributed and mostly (96%) inverted with respect to the *hsr* gene. DNA sequence alignments of nine different *H. mustelae* strains, showed a high level of sequence variation in the variable repeat region of the *hsr* gene, in contrast to the central and β domains. Alignment of sequenced DNA from the variable repeat region of different strains identified conserved-variable-conserved blocks (CVCs) of sequence, which may facilitate a recombination-based antigenic variation mechanism. Approximately 7 kb upstream and 3 kb downstream *hsr*-related flanking sequence may serve as a reservoir for sequence variation of the *hsr* gene. The searches for repeat elements have facilitated the identification of potential DNA regulatory elements involved in the abundant production of the Hsr protein.

The HSRL also contained an unrelated open reading frame, encoding Orf2, which had significant identity with LolA, a periplasmic lipoprotein carrier protein, but containing an N-terminal extension of 14 charged and polar amino acids. Insertional inactivation of *orf2* had no detectable effect on Hsr expression in the Hsr⁺ strain 4298.

ACKNOWLEDGEMENTS

To my supervisor, **Dr Paul O'Toole**. First, thank you for giving me the opportunity to work in your lab. I have enjoyed the experience. Thanks also for endeavouring to provide various resources for the lab to keep us all happy/quiet. Many thanks for the scientific guidance and taking time to read and provide helpful suggestions during the writing of this thesis. I apologise for the all the Tash-induced migraines received over the past few years.

Thanks also to **the Institute of Molecular BioSciences/ Department of Microbiology and Genetics** for allowing me to do this study part time while working. Thanks also to **Dr Kathryn Stowell** for cheerfully providing helpful advice on a number of occasions.

I would like to thank **Dr Kathy Parton** (IVABS, Massey University) for sourcing ferret samples, without which, I would not have been able to complete this work. Many thanks also to **Mr Terry Hynes** (District manager, Agriquality New Zealand) for the taking the time to collect and send ferret samples.

I am grateful to all Helipad members over the past four years (**Grover, Amanda, Kirsty, James, Mick, Basil, Michael (x2), Millis, Anja, Jasna, Stanmanda, Pania, Jakki, Todd, and Paul**) for your various contributions to my scientific (and social) development. Thanks also for the favours done here and there. A special thanks goes to **Dr Jasna Rakonjac**. Thank you so much Jasna, for your encouragement, proofing/editing, and helpful discussions/ tutorials over the past couple of years. It has been greatly appreciated.

To anyone else who contributed to my thesis in some small way, e.g., lending equipment, handy tips, administration, "mental health morning teas", and stuff like that - thanks heaps. Thanks to my family and friends for all of your help (in several forms).

Finally to my husband **Pete**. Thank you for your support, patience, taking care of Gracie in the evenings, and frequently locking me in the office. You can go back to the shed and play now. XXXX.

RELATED PUBLICATIONS

Some of the material presented in this thesis has been published.

Forester, N.T., Parton, K., Lumsden, J.S., and O'Toole, P.W. (2000). Isolation of *Helicobacter mustelae* from ferrets in New Zealand. *New Zealand Veterinary Journal* **48**:65-69.

Forester, N., Lumsden, J.S., O'Croinin, T., and O'Toole, P.W. (2001). Sequence and antigenic variability of the *Helicobacter mustelae* surface ring protein Hsr. *Infection and Immunity* **69**(5):3447 – 3450.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGEMENTS	iii
RELATED PUBLICATIONS	iv
TABLE OF CONTENTS	v
LIST OF FIGURES	ix
LIST OF TABLES	xi
1. INTRODUCTION	1
1.1 THE GENUS <i>HELICOBACTER</i>	1
1.1.1 General history	1
1.1.2 General characteristics of members of the genus <i>Helicobacter</i>	1
1.1.3 Medical significance of <i>Helicobacters</i> to Humans	2
1.2 <i>HELICOBACTER MUSTELAE</i>	3
1.2.1 General background	3
1.2.2 Ferrets and the ferret animal model	5
1.2.3 Characteristic features of <i>H. mustelae</i>	5
1.3 THE <i>HELICOBACTER MUSTELAE</i> SURFACE RING PROTEIN (HSR)	7
1.3.1 The Hsr protein	7
1.3.2 Hsr belongs to the family of Autotransporter proteins	9
1.3.3 The <i>hsr</i> gene and generation of the sequence of the 12 kb <i>hsr</i> locus – preliminary information	12
1.4 FUNCTIONAL ROLES OF SURFACE EXPOSED PROTEINS OF GRAM-NEGATIVE MUCOSAL PATHOGENS	14
1.5 GENERATION OF BACTERIAL PROTEIN VARIABILITY	15
1.5.1 Repeats sequences in prokaryotes	15
1.5.2 Pathoadaptive DNA rearrangements and mutations	16
1.5.3 <i>Helicobacter</i> natural competence	19
1.6 PRELIMINARY RESEARCH OBJECTIVES	20
2. MATERIALS AND METHODS	21
2.1 BACTERIAL STRAINS, CULTURE, AND STORAGE CONDITIONS	21
2.1.1 Bacterial strains	21
2.1.2 <i>Helicobacter</i> isolation from ferret stomachs	22
2.1.2.1 Processing of ferret stomachs	22
2.1.2.2 Biochemical analysis of putative <i>Helicobacter</i> isolates cultured from ferret stomachs	23

2.2	MEDIA AND ADDITIVES.....	23
2.3	OLIGONUCLEOTIDE PRIMERS.....	24
2.4	VECTORS AND RECOMBINANT PLASMIDS.....	26
2.5	ANTISERA.....	26
2.6	DNA PREPARATION.....	27
2.6.1	Plasmid preparation.....	27
2.6.1.1	Easy plasmid miniprep (Easyprep).....	27
2.6.1.2	WIZARD™ plasmid miniprep.....	27
2.6.1.3	CONCERT™ Rapid Plasmid Purification Miniprep System.....	27
2.6.2	Preparation of genomic DNA from <i>Helicobacter</i> cells.....	28
2.6.3	Extraction of bacterial DNA from stomach tissue for PCR analysis.....	28
2.7	DNA ANALYSIS METHODS.....	29
2.7.1	DNA agarose gel electrophoresis.....	29
2.7.2	DNA restriction endonuclease treatment.....	30
2.7.3	DNA quantification.....	30
2.7.4	Southern blotting and hybridisation.....	31
2.7.5	DNA sequencing.....	32
2.8	DNA AMPLIFICATION BY POLYMERASE CHAIN REACTION (PCR).....	33
2.9	CLONING STRATEGIES.....	34
2.9.1	DNA preparation.....	34
2.9.1.1	Vector DNA.....	34
2.9.1.2	Insert DNA.....	35
2.9.2	Ligation.....	35
2.9.3	Transformation.....	36
2.9.3.1	Preparation of competent bacterial cells.....	36
2.9.3.2	Transformation of <i>E. coli</i>	37
2.9.3.3	Transformation of <i>H. mustelae</i>	37
2.10	PROTEIN SAMPLE PREPARATION.....	38
2.11	PROTEIN ANALYSES.....	39
2.11.1	Protein electrophoresis.....	39
2.11.2	Western blotting and immunodetection.....	39
2.12	MICROSCOPY.....	40
3.	RESULTS.....	41
3.1	BASIC ANALYSIS OF THE 12 KB <i>HSR</i> LOCUS (HSRL).....	41
3.2	SEQUENCE VARIATION OF THE <i>HSR</i> GENE IN <i>H. MUSTELAE</i>	43
3.2.1	<i>Helicobacter</i> isolation from stomachs of New Zealand ferrets.....	43
3.2.2	Molecular analysis of the New Zealand (N. Z.) <i>H. mustelae</i> isolates.....	44
3.2.2.1	16S ribosomal gene analysis of <i>Helicobacter</i> strains isolated from N.Z. ferrets.....	44

3.2.2.2	Total cellular protein profiles and Western blot analysis of <i>H. mustelae</i> isolates obtained from N.Z. ferrets.....	48
3.2.3	Southern blot analysis to investigate organisation and conservation of the HSRL of seven <i>H. mustelae</i> isolates	50
3.2.4	DNA sequence analysis of the variability in the repeat region of the HSRL of different <i>H. mustelae</i> strains.....	53
3.2.4.1	PCR amplification and DNA sequencing of sections of the <i>hsr</i> gene from nine <i>H. mustelae</i> strains.....	53
3.2.4.2	Sequence alignments of regions of the <i>hsr</i> gene from nine <i>H. mustelae</i> strains.....	55
3.2.4.3	Search for conserved-variable-conserved motifs (CVCs) in the <i>hsr</i> -like sequence flanking the <i>hsr</i> gene of strain 4298.....	57
3.2.4.4	Properties of the peptide encoded by part of the variable repeat region of the <i>hsr</i> gene of nine <i>H. mustelae</i> strains.....	59
3.3	COMPLETION OF THE DNA SEQUENCE OF THE HSRL.....	61
3.3.1	The lambda clone λ E2	61
3.3.2	Restriction endonuclease mapping and subcloning of λ E2 for DNA sequence determination 3' of the <i>hsr</i> gene of <i>H. mustelae</i> strain 4298	61
3.3.3	DNA sequence analysis of pUC19- λ E2/E1	64
3.3.4	DNA Sequence analysis of pUC19- λ E2/E2	64
3.3.4.1	Sequencing of the <i>hsr</i> -like sequence in the λ E2/E2 DNA fragment.....	64
3.3.4.2	Other non- <i>hsr</i> related features of the λ E2/E2 DNA fragment	65
3.3.5	Sequence arrangement with respect to the <i>hsr</i> locus of strain 4298 and completion of the 14919 bp <i>hsr</i> locus DNA sequence.....	66
3.4	DNA SEQUENCE ANALYSIS OF THE 14919 BP <i>HSR</i> LOCUS (HSRL).....	68
3.4.1	Open reading frame analysis.....	68
3.4.2	Repeat sequences of the HSRL.....	69
3.4.2.1	Analysis of dispersed DNA sequence repeats in the <i>hsr</i> locus.....	69
3.4.2.2	Tandem repeats in the <i>hsr</i> locus.....	73
3.4.2.3	Imperfect inverted repeats and potential stem-loop structures in the <i>hsr</i> locus	74
3.4.2.4	Analysis of imperfect repeats in the <i>hsr</i> locus.....	75
3.5	KNOCKOUT MUTAGENESIS OF THE <i>ORF2</i> GENE IN <i>H. MUSTELAE</i> STRAIN 4298.....	77
3.5.1	Analysis of the <i>orf2</i> gene and gene products	77
3.5.1.1	The <i>orf2</i> gene	77
3.5.1.2	The Orf2 protein.....	78
3.5.2	Generation of the <i>orf2</i> knockout plasmid pHM205 Δ ORF2.....	79
3.5.3	Transformation of <i>H. mustelae</i> strain 4298 with pHM205 Δ ORF2.....	79

4. DISCUSSION.....	81
4.1 <i>HELICOBACTER MUSTELAE</i> IS PRESENT IN FERRETS FROM AT LEAST TWO GEOGRAPHICALLY DISTINCT LOCATIONS IN NEW ZEALAND.	81
4.2 DISTRIBUTION OF THE <i>HSR</i> -RELATED REPEATS.....	81
4.2.1 The <i>hsr</i> gene is flanked by multiple repeats of <i>hsr</i> sequence, in the 15 kb <i>hsr</i> locus of <i>H. mustelae</i> strain 4298	81
4.2.1.1 The distribution and complexity of repeat sequences in the 15 kb <i>hsr</i> locus of <i>H. mustelae</i> strain 4298.....	81
4.2.1.2 Repeat features with possible roles in <i>hsr</i> expression.....	86
4.2.2 Distribution of <i>hsr</i> repeat sequences within the <i>hsr</i> gene	89
4.3 HSR VARIABILITY	90
4.3.1 Variability of the <i>hsr</i> gene of different <i>H. mustelae</i> strains	90
4.3.2 Serological analysis of Hsr protein variability of <i>H. mustelae</i> strains	94
4.3.3 Supporting evidence for antigenic variation.....	95
4.4 GENOMIC ORGANISATION FLANKING THE HSRL	95
4.5 ORF2 IS RELATED TO THE LOLA FAMILY OF LIPOPROTEIN CARRIER PROTEINS.	97
4.6 FUNCTION OF THE HSR PROTEIN.....	99
5. SUMMARY OF THE MAIN OUTCOMES OF THIS STUDY AND SUGGESTED FUTURE STUDY DIRECTIONS	104
5.1 <i>H. MUSTELAE</i> AND NEW ZEALAND MUSTELIDS.....	104
5.2 HSR RING STRUCTURE AND SUGGESTED FUNCTIONS.....	104
5.3 <i>HSR</i> -LIKE REPEATS AND ANTIGENIC VARIATION	105
5.4 <i>HSR</i> EXPRESSION	105
5.5 ORF2 – THE LOLA HOMOLOGUE	106
APPENDIX 1 Physical maps of vectors used in this study	107
APPENDIX 2 The complete sequence of the <i>hsr</i> locus.....	111
APPENDIX 3 Sequences repeated in the <i>hsr</i> locus of <i>H. mustelae</i> strain 4298...	127
APPENDIX 4 DNA and protein sequence alignments	144
REFERENCES.....	150

LIST OF FIGURES

Figure 1.1	<i>H. mustelae</i> cells morphology and Hsr protein rings	4
Figure 1.2	Autotransporter domains and membrane topology.....	10
Figure 1.3	Genomic organisation of the <i>hsr</i> locus of <i>H. mustelae</i> strain 4298.....	13
Figure 3.1	Occurrence of repetitive DNA sequences in the 12 kb <i>hsr</i> locus (HSRL) ..	42
Figure 3.2	Multiple alignment of partial 16S DNA sequences from <i>Helicobacter mustelae</i> strains isolated from New Zealand ferret stomachs	46-47
Figure 3.3	Hsr is present in New Zealand ferret <i>H. mustelae</i> isolates	49
Figure 3.4	HSRL restriction patterns of <i>H. mustelae</i> isolates are different	51
Figure 3.5	Comparative DNA sequence analysis of three regions of the <i>hsr</i> gene of nine <i>H. mustelae</i> isolates.....	54
Figure 3.6	Conserved-variable-conserved motifs in the <i>hsr</i> locus.....	58
Figure 3.7	Kyte & Doolittle scale mean hydrophobicity profiles of part of the variable protein sequences of nine <i>H. mustelae</i> isolates.....	60
Figure 3.8	Restriction endonuclease and DNA sequence mapping of λ E2	63
Figure 3.9	Multiple alignment of the λ E2/E2 ORF against <i>H. pylori</i> Glr proteins	65
Figure 3.10	PCR confirmation of the assembly of the DNA sequences of the HSRL of <i>H. mustelae</i> strain 4298.....	67
Figure 3.11	Repeat distribution in the <i>hsr</i> locus of <i>H. mustelae</i> strain 4298.....	71
Figure 3.12	Physical organisation and repeat sequences in the <i>hsr</i> locus (HSRL) of <i>H. mustelae</i> strain 4298.....	72
Figure 3.13	Distribution of the HSR region direct repeats in the HSRL	73
Figure 3.14	Distribution of potential stem-loop structures in the <i>hsr</i> locus.....	75
Figure 3.15	Local DNA features of the <i>orf2</i> gene	77
Figure 3.16	Alignment of Orf2 translated product against <i>H. pylori</i> LolA proteins. ...	78
Figure 3.17	PCR confirmation of kanamycin resistant 4298 Δ ORF2 transformants.....	80
Figure 3.18	SDS-PAGE analysis of whole cell lysates of six kanamycin resistant 4298 Δ ORF2 transformants	80

Figure 4.1	Overview of the <i>hsr</i> locus and genetic elements.	83-84
Figure 4.2	Potential regulatory elements of the <i>hsr</i> gene.....	87
Figure 4.3	Potential mechanisms for generation of diversity in the <i>hsr</i> locus.....	92-93
Figure 4.4	Comparison of the genetic organisation around the <i>H. mustelae</i> strain 4298 HSRL with respect to the corresponding <i>H. pylori</i> strain 26695 DNA and protein homologues.	96
Figure 4.5	Model of potential Hsr structure and function.....	101-102

LIST OF TABLES

Table 1.1	List of <i>Helicobacter</i> species and their hosts.....	2
Table 1.2	Characteristics of <i>H. mustelae</i> compared with <i>H. pylori</i>	6
Table 1.3	Summary of the amino acid characteristics of the Hsr protein.....	8
Table 1.4	Autotransporter proteins occur in many Gram-negative bacteria and have functions promoting disease in host organisms	10
Table 2.1	Bacterial strains used in this study.....	21
Table 2.2	Recipes for media used in this study.	24
Table 2.3	Oligonucleotide primers used in this study.	25
Table 2.4	Plasmids and other vectors used in this study.	26
Table 2.5	Antisera used in this study.....	26
Table 3.1	Summary of the repeat sequence frequency.	42
Table 3.2	Ferret stomach processing details.....	44
Table 3.3	Summary of 16S DNA alignment results for the <i>Helicobacter</i> isolated from ferrets from two separate regions of New Zealand.....	45
Table 3.4	Restriction analysis and Southern blotting results for strain 4298	52
Table 3.5	Summary of PCR products amplified from three different regions of the HSRL of nine strains of <i>Helicobacter mustelae</i>	53
Table 3.6	Summary of DNA sequence alignment of three different regions of the <i>hsr</i> gene of nine <i>Helicobacter mustelae</i> isolates	56
Table 3.7	Repetitive nature of the conserved and variable sequence blocks from the repeat region of the <i>hsr</i> gene of strain 4298.	59
Table 3.8	PCR confirmation of the arrangement of λ E2 with respect to the <i>hsr</i> locus (HSRL)....	67
Table 3.9	Open reading frames of the <i>hsr</i> locus.....	68
Table 3.10	Summary of the repetitive DNA sequences in the 14919 bp <i>hsr</i> locus (HSRL) .	70
Table 3.11	Tandem repeats in the <i>hsr</i> locus	74
Table 3.12	Potential stem loop (PSL) structures in the <i>hsr</i> locus.....	75