https://theses.gla.ac.uk/

Theses digitisation:

https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses
https://theses.gla.ac.uk/
research-enlighten@glasgow.ac.uk

UNIVERSITY
of
GLASGOW

Genotypic diversity and epidemiological typing of *Bordetella pertussis*

Shona Elaine Neal B.Sc. (Hons)

Presented for the degree of Doctor of Philosophy to the Division of Infection and Immunity, Institute of Biomedical and Life Sciences, Faculty of Science, University of Glasgow

March 2004

ProQuest Number: 10390650

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 10390650

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

GLASGOW
UNIVERSITY
LIBRARY
13446

DECLARATION

Unless otherwise stated, this thesis is the original work of the author

Shona Elaine Neal

ACKNOWLEDGEMENTS

This PhD studentship was funded by the Public Health Laboratory Service (PHLS[1]) Central Research and Development Fund.

I would like to thank my supervisors for entrusting me with this captivating thesis and for all their support: Dr Norman Fry for his constant patience and guidance, Dr Tim Harrison for his considerate focus and criticism in discussions, and Dr Roger Parton, Dr John Coote and Prof Liz Miller for their invaluable advice and encouragement.

I would also like to thank the following people for providing me with a wide selection of *Bordetella pertussis* strains, which helped to establish the current *Bordetella* collection held at RSIL, HPA, and, is in itself, an achievement: Dr Barry Holmes (NCTC); Andy Robinson (CAMR); Ruth Matthews (Manchester University); Sandy Hanauer (Missouri State Public Health Laboratory, USA); Dorothy Xing (NIBSC), Elisabeth Njamkepo (Institut Pasteur, France); and Han van der Heide (RIVM), who also gave additional information on the IS*1002*-RFLP analysis procedure.

I am grateful to Nick Andrews, André Charlett, Pauline Rogers and Neville Verlander for their statistical advice and tutorials (Statistics Unit), to Steve Platt and Anthony Underwood for their bioinformatic support (Genomics, Proteomics and Bioinformatics Unit), to Jennie Lane for her utmost patience in amending my various diagrams (Department of Medical Illustration) and Natasha Crowcroft for information on recent UK vaccine use and on the PICU study (CDSC Immunisation). Part of this thesis would not have been possible without Julian Parkhill's kind permission to use, the then unpublished, annotated *B. pertussis* genome.

I could not have worked and played so hard without all my happy, friendly colleagues at RSIL and the HPA, but especially Androulla, Bee, Carmen, Chen, Craig, Michaela, Oceanis, Ru, Tony, and Teresa. They have always helped me de-stress when I needed it most. A thousand thank yous.

I would finally like to acknowledge my friends and family who have always been there through the good times and bad, especially Becky, Bhavna, Diane, Kath, Nicky, Phil, Shahla, Sheena and my sister Christine. To Ben, his patience, understanding and commitment have endured, and I'd like to thank him for this. My days off were few and far between. And to my Mum and Dad, I am eternally grateful for all their support in anything I choose to do. They have always trusted my judgement and ability, which has made me realise how anything can be accomplished. Thank you.

---

[1]The PHLS was integrated into the Health Protection Agency (HPA) on the 1st of April 2003. The HPA is a new independent organisation dedicated to protecting people's health. It brings together the expertise formerly in a number of official organisations, and it works in partnership with the National Radiological Protection Board.

CONTENTS

DECLARATION.....i
ACKNOWLEDGEMENTS.....ii
CONTENTS.....iii
LIST OF TABLES.....x
LIST OF FIGURES.....xiii
ABBREVIATIONS.....xvi
ABSTRACT.....xviii

1 INTRODUCTION.....1
1.1 *Bordetella* species.....2
1.1.1 Historical account of the *Bordetella* genus.....2
1.1.2 Amended definition of the *Bordetella* genus.....4
1.1.3 Differentiation of species within the *Bordetella* genus.....6
1.2 *Bordetella* virulence factors and pathogenesis.....6
1.2.1 Regulation of virulence factor expression.....6
1.2.2 Toxins.....8
1.2.2.1 Pertussis toxin.....8
1.2.2.2 Adenylate cyclase toxin.....10
1.2.2.3 Dermonecrotic toxin.....11
1.2.2.4 Tracheal cytotoxin.....11
1.2.2.5 Lipopolysaccharide.....12
1.2.3 Adhesins.....12
1.2.3.1 Filamentous haemagglutinin.....12
1.2.3.2 Fimbriae.....13
1.2.3.3 Pertactin.....15
1.2.3.4 Bordetella resistance to killing protein.....17
1.2.3.5 Tracheal colonisation factor.....17
1.2.3.6 Virulence-activated gene 8.....18
1.2.4 Other virulence factors.....18
1.3 Clinical aspects and diagnosis of pertussis.....19
1.3.1 Clinical symptoms, complications, and treatment of pertussis infection.....19
1.3.2 Laboratory diagnosis of *Bordetella pertussis* infection.....20
1.4 Epidemiology and prevention.....22
1.4.1 Pertussis vaccination in the United Kingdom.....22

1.4.2 Acellular vaccines 24
1.4.3 Pertussis in other countries....26
1.4.4 Recent changes in the dynamics of pertussis disease....28
1.5 Genotypic typing of bacteria....30
1.5.1 Caveats to typing bacteria....30
1.5.2 Genotypic methods....31
1.5.2.1 Plasmid typing....31
1.5.2.2 Restriction endonuclease analysis....33
1.5.2.3 Insertion sequence- restriction fragment length polymorphism analysis and ribotyping....33
1.5.2.4 Pulsed-field gel electrophoresis....34
1.5.2.5 PCR amplification methods....34
1.5.2.6 Amplified fragment length polymorphism analysis....35
1.5.2.7 DNA sequencing....35
1.5.3 Genetic events in bacteria visualised using genotypic methods....36
1.5.4 The "ideal" bacterial typing system....38
1.6 Genetic diversity and epidemiological typing of *Bordetella pertussis*....39
1.6.1 Genome plasticity of *Bordetella pertussis*....39
1.6.2 Genotypic diversity of *Bordetella pertussis*....39
1.6.3 Epidemiological typing of *Bordetella pertussis*....43
1.6.4 Gene sequence variation of *Bordetella pertussis*....46
1.7 Aims of thesis....50
1.7.1 To characterise the range and distribution of sequence variation in pertactin and pertussis toxin genes of *Bordetella pertussis* isolates from the United Kingdom....50
1.7.2 To investigate the genotypic diversity of a representative sample of historical, through to current, UK *Bordetella pertussis* isolates....51
1.7.3 To examine current *Bordetella pertussis* isolates for any association between clinical, epidemiological, and laboratory data....51
1.7.4 To characterise *Bordetella pertussis* strains collected by the Missouri State Public Health Laboratory from Missouri, USA....51
1.7.5 To evaluate molecular typing techniques for their indices of discrimination by using a panel of *Bordetella pertussis* isolates....51
1.7.6 To expand work on a DNA sequence-based approach to the genotypic characterisation of *Bordetella pertussis*....52
1.7.7 To identify novel molecular methods to type *Bordetella pertussis* isolates....52

1.7.8 To investigate a number of epidemiologically-significant markers by using clinical material obtained from *Bordetella pertussis*-positive patients 53

2 MATERIALS AND METHODS 54

2.1 *Bordetella* isolates 55

2.1.1 Historical *Bordetella pertussis* isolates 55

2.1.2 Contemporary UK *Bordetella pertussis* isolates 57

2.1.3 *Bordetella pertussis* vaccine and reference strains 57

2.1.4 *Bordetella pertussis* isolates from Missouri, USA 58

2.1.5 *Bordetella parapertussis* strains 58

2.2 Revival, growth and preservation of *Bordetella* species 58

2.2.1 Revival of *Bordetella* from freeze-dried ampoules 58

2.2.2 Culture conditions 59

2.2.3 Preservation of *Bordetella pertussis* isolates 59

2.3 Serotyping of *Bordetella pertussis* 60

2.4 DNA extraction 60

2.4.1 DNA purification from pure culture 60

2.4.2 Quantification of DNA 61

2.4.3 DNA purification from clinical samples 62

2.5 DNA sequencing 63

2.5.1 Design of oligonucleotide primers for PCR amplification and DNA sequencing 63

2.5.2 General methodology for PCR 63

2.5.3 Separation and visualisation of PCR products 64

2.5.4 PCR amplification of the pertactin and the pertussis toxin S1 subunit genes 68

2.5.4.1 Pertactin gene 68

2.5.4.2 Pertussis toxin S1 subunit gene 73

2.5.5 PCR product purification 76

2.5.5.1 Vacuum manifold method for PCR product purification 76

2.5.5.2 Syringe method for PCR product purification 77

2.5.6 Sequencing reaction and precipitation protocols 77

2.5.6.1 ABI Prism® BigDye™ Terminator Cycle Sequencing Ready Reaction Kit 78

2.5.6.2 CEQ™ 2000 Dye Terminator Cycle Sequencing with Quick Start Kit 79

2.5.7 Sequence analysis 80

2.6 IS*1002*-Restriction Fragment Length Polymorphism analysis 81

2.6.1 Digoxigenin-labelling of the IS*1002* probe via PCR amplification 81

2.6.2 Restriction endonuclease digestion with *Sma*I 82

2.6.3 Separation of DNA fragments by Field Inversion Gel Electrophoresis 82

2.6.4 Vacuum transfer of DNA to a nylon membrane 84

2.6.5 Pre-hybridisation and hybridisation of membranes 85

2.6.6 Washing and detection of digoxigenin-labelled bands 86

2.6.7 Analysis of IS*1002*-Restriction Fragment Length Polymorphism profiles 91

2.7 Pulsed-Field Gel Electrophoresis 94

2.7.1 Preparation of genomic DNA 94

2.7.2 Restriction endonuclease digestion with *Xba*I 94

2.7.3 Separation and detection of DNA fragments 95

2.7.4 Analysis of Pulsed-Field Gel Electrophoresis profiles 95

2.8 Data analysis 98

2.8.1 Calculation of genotypic diversity values 98

2.8.2 Identification of associations between clinical and laboratory data 98

2.8.3 Evaluation of typing methods 99

3 Population dynamics of *Bordetella pertussis* isolates in the UK 101

3.1 Strategy 102

3.1.1 DNA sequencing of the pertactin and pertussis toxin genes 102

3.1.2 Genotypic diversity of UK *Bordetella pertussis* strains between 1940 and 1999 102

3.1.3 Identification of associations between clinical and laboratory data 103

3.2 Methods 103

3.2.1 DNA sequencing of the pertactin and pertussis toxin genes 103

3.2.2 Random selection of UK *Bordetella pertussis* strains between 1940 and 1999 for IS*1002*-Restriction Fragment Length Polymorphism analysis 104

3.2.3 Identification of clinical and laboratory associations of pertussis cases from the UK between 1998 and 1999 104

3.3 Results 105

3.3.1 Polymorphism in the pertactin gene 105

3.3.2 Polymorphism of the pertussis toxin S1 subunit gene 106

3.3.3 Combined pertactin and pertussis toxin types 112

3.3.4 Genotypic diversity of UK *Bordetella pertussis* between 1940 and 1999 112

3.3.5 Epidemiological associations and trends of pertussis cases from the UK between 1998 and 1999 114

3.3.6 Laboratory typing results of the 90 UK cases from 1998 to 1999 126

3.3.6.1 Associations of *Bordetella pertussis* serotypes in the 90 pertussis cases . 132

3.3.6.2 Associations of *Bordetella pertussis* pertactin gene types in the 90 pertussis cases ... 134

3.3.6.3 Associations of *Bordetella pertussis* IS*1002*-Restriction Fragment Length Polymorphism types in the 90 pertussis cases ... 134

3.3.6.4 Associations of *Bordetella pertussis* combined types in the 90 pertussis cases ... 135

3.3.6.5 Age cohorts trend analysis ... 139

3.3.6.6 Temporal trend analysis ... 139

3.4 Discussion ... 142

3.4.1 Variation in the pertactin gene ... 142

3.4.2 Variation in the pertussis toxin S1 gene ... 145

3.4.3 Combined pertactin and pertussis toxin S1 types in the UK ... 146

3.4.4 Genotypic diversity of UK *Bordetella pertussis* between 1940 and 1999 ... 147

3.4.5 UK *Bordetella pertussis* sub-set from 1998 to 1999 ... 153

3.4.6 Associations within the pertussis epidemiological data of the 90 UK cases from 1998 to 1999 ... 155

3.4.7 Association between epidemiological data and typing results of the 90 UK *Bordetella pertussis* isolates from 1998 to 1999 ... 156

4 Characterisation of *Bordetella pertussis* isolates from Missouri, USA ... 162

4.1 Strategy ... 163

4.2 Methods ... 165

4.3 Results ... 165

4.3.1 Serotyping results ... 165

4.3.2 Polymorphism in the pertactin and pertussis toxin S1 genes ... 166

4.3.3 IS*1002*-Restriction Fragment Length Polymorphism analysis ... 174

4.3.4 Pulsed-Field Gel Electrophoresis analysis ... 177

4.3.5 Discriminatory indices and associations of the various typing methods ... 177

4.4 Discussion ... 181

4.4.1 Genotypic diversity of *Bordetella pertussis* isolates from Missouri, USA ... 182

4.4.2 Geographical distribution of *Bordetella pertussis* types ... 183

4.4.3 *Bordetella pertussis* isolates from Kansas City, Missouri, USA ... 185

5 Evaluation of typing methods for *Bordetella pertussis* ... 187

5.1 Strategy ... 188

5.1.1 Criteria used to select the representative *Bordetella pertussis* panel ... 188

5.2 Methods ... 189

5.2.1 The *in silico* digest of the Tohama-I genome ... 189

5.3 Results ... 192

5.3.1 Typability of typing methods 192

5.3.2 Reproducibility of typing methods 194

5.3.3 Discrimination indices of typing methods 194

5.3.4 Epidemiological concordance of typing methods 198

5.3.5 Linkage between typing methods 202

5.3.6 Comparison of IS*1002*-RFLP analysis and PFGE profiles by *in silico* and observed analysis of the Tohama-I strain 202

5.4 Discussion 203

5.4.1 The suitability of the *Bordetella pertussis* typing panel 208

5.4.2 Evaluation of the typing methods 209

5.4.3 Recommendation of a typing system for *Bordetella pertussis* 214

6 Novel *Bordetella pertussis* gene targets for epidemiological typing 217

6.1 Strategy 218

6.2 Methods 218

6.2.1 Adenylate kinase gene 219

6.2.2 Filamentous haemagglutinin gene 219

6.3 Results 220

6.3.1 Adenylate kinase gene 220

6.3.2 Filamentous haemagglutinin gene 226

6.4 Discussion 232

7 Assessment of single-enzyme Amplified Fragment Length Polymorphism analysis of *Bordetella pertussis* 234

7.1 Strategy 235

7.2 Method 235

7.2.1 Single step restriction-ligation of genomic DNA 237

7.2.2 PCR amplification of tagged DNA fragments 237

7.2.3 Amplified Fragment Length Polymorphism analysis 239

7.3 Results 239

7.4 Discussion 247

8 Epidemiological typing direct from UK clinical samples 249

8.1 Strategy 250

8.2 Methods 250

8.2.1 Amplification of the pertactin gene 250

8.2.2 Amplification of the pertussis toxin S1 gene 251

8.3 Results 253

8.3.1 Detection of the pertactin gene, with primers AF and BR, in clinical samples from the UK 253

8.3.2 Detection of the pertussis toxin S1 gene, using primers SF and SR, in UK clinical samples 257
8.3.3 Detection of the pertactin and pertussis toxin S1 genes, with primers AF and AR, and S1FM and SR, respectively, in UK clinical samples 257
8.3.4 Pertactin and pertussis S1 gene types generated from PCR-positive clinical specimens 263
8.4 Discussion 265
9 CONCLUSIONS 267
10 FUTURE WORK 272
10.1 Pertactin and pertussis toxin gene sequencing of UK isolates 273
10.2 Pertactin and pertussis toxin gene sequencing in other countries 273
10.3 Pulsed-Field Gel Electrophoresis 273
10.4 Amplified Fragment Length Polymorphism typing 273
10.5 Epidemiological typing of *Bordetella pertussis* direct from clinical samples 274
11 REFERENCES 275
12 APPENDICES 292
12.1 Appendix 1. Amino acid abbreviations 293
12.2 Appendix 2. Address list of suppliers 294
12.3 Appendix 3. *Bordetella* isolates available for thesis 296
12.4 Appendix 4. Enhanced surveillance forms used by the Communicable Disease Surveillance Centre 318
12.5 Appendix 5. Publications arising from this thesis 321

LIST OF TABLES

Table 1.1. Differential properties of *Bordetella* species 5

Table 1.2. Virulence factors of *Bordetella* species 7

Table 1.3. Characteristics of genotypic bacterial typing systems 32

Table 1.4. Insertion elements present in *Bordetella* 41

Table 2.1. Summary of bacterial isolates available for this thesis 56

Table 2.2. *Bordetella pertussis* gene targets selected and sequences used to design primers.....65

Table 2.3. Oligonucleotide primers used for amplification and sequencing of the *Bordetella pertussis* genes 66

Table 2.4. Optimisation of IS*1002*-RFLP with membrane strips 88

Table 3.1. Epidemiological information available for 90 cases from whom *Bordetella pertussis* isolates were recovered between September 1998 and April 1999 107

Table 3.2. UK *Bordetella pertussis* isolates investigated for nucleotide polymorphisms outside regions 1 and 2[1] 109

Table 3.3. Diversity of *Bordetella pertussis* IS*1002*-RFLP types, according to the pertactin and pertussis toxin S1 gene types, and combined *prnA/ptxA* types 119

Table 3.4. Distribution of major epidemiological factors in the 90 UK pertussis cases analysed from September 1998 to April 1999 120

Table 3.5. Probability values of epidemiological associations of the 90 pertussis cases from 1998 to 1999, generated from Pearson's χ^2 analysis 122

Table 3.6. Sex distribution of pertussis cases in the nine geographical regions 123

Table 3.7. Distribution of typical pertussis symptoms through age cohorts 123

Table 3.8. Association of pneumonia with pertussis infection through age cohorts.......... 123

Table 3.9. Factors influencing likelihood of encountering a complication with a pertussis infection 127

Table 3.10. Factors influencing likelihood of encountering apnoea with a pertussis infection 127

Table 3.11. Association of hospital admissions with number of pertussis vaccine doses . 128

Table 3.12. Factors influencing likelihood of admission to hospital with a pertussis infection 128

Table 3.13. The distribution of typing information of the 90 UK pertussis cases analysed from September 1998 to April 1999 129

Table 3.14. Combined types of *Bordetella pertussis* from 1998 to 1999 in the UK, and their corresponding serotype, pertactin gene type, and IS*1002*-RFLP type 131

Table 3.15. Probability values between serotypes of *Bordetella pertussis* isolates and associated factors 133

Table 3.16. Probability values between pertactin gene types of *Bordetella pertussis* isolates and associated factors 136

Table 3.17. Probability values between IS*1002*-RFLP types of *Bordetella pertussis* isolates and associated factors 137

Table 3.18. Geographical distribution of IS*1002*-RFLP types 137

Table 3.19. Probability values between combined types of *Bordetella pertussis* isolates and associated factors 138

Table 3.20. Geographical distribution of combined types 138

Table 3.21 Distribution of *Bordetella pertussis* serotype through age cohorts 140

Table 3.22. Distribution of *Bordetella pertussis* pertactin gene type through age cohorts 140

Table 3.23 Distribution of serotypes in UK *Bordetella pertussis* isolates from September 1998 to April 1999 141

Table 3.24 Distribution of pertactin types in UK *Bordetella pertussis* isolates from September 1998 to April 1999 141

Table 3.25 Corresponding *Bordetella pertussis* IS*1002*-RFLP profiles between the UK and The Netherlands 149

Table 4.1. Distribution of serotype, pertactin and pertussis toxin S1 subunit variants from 44 *Bordetella pertussis* isolates from Missouri, USA, by isolation period 167

Table 4.2. Details of the six *Bordetella pertussis* isolates from Missouri, USA, for which over 90% of the coding region of the pertactin gene was determined 167

Table 4.3. Temporal distribution of IS*1002*-RFLP types in *Bordetella pertussis* isolates from Missouri, USA 176

Table 4.4. Temporal distribution of PFGE types, produced with *Xba*I, in *Bordetella pertussis* isolates from Missouri, USA 176

Table 4.5. Indices of discrimination of the typing methods performed on the 43 *Bordetella pertussis* isolates from Missouri, USA 180

Table 5.1. *Bordetella pertussis* panel for the evaluation of typing techniques 190

Table 5.2. Results of the typing methods derived from the *Bordetella pertussis* typing panel193

Table 5.3. Type designations of replicate isolates from the panel, with reproducibility values for each *Bordetella pertussis* typing method 195

Table 5.4. Type designations of 'epidemiologically-related' *Bordetella pertussis* isolates from the panel, with epidemiological concordance values 199

Table 5.5. A comparison of observed and predicted band sizes of fragments containing the IS*1002* elements of the *Bordetella pertussis* Tohama-I strain 204

Table 5.6. A comparison of observed and expected band sizes of the PFGE fragments of *Bordetella pertussis* Tohama-I with *Xba*I 206
Table 5.7. Summary of the evaluation of the typing methods used on *Bordetella pertussis* .. 211
Table 7.1. Restriction endonucleases, adapters and primers used for AFLP on *Bordetella pertussis* 238
Table 7.2. Type designations of the *Bordetella pertussis* isolates from Missouri, USA and loading order of the AFLP gels, shown in Figures 7.2 to 7.6 240
Table 8.1. *Bordetella pertussis* clinical specimens investigated for PCR amplification of the pertactin gene, as visualised in Figure 8.3 255
Table 8.2. *Bordetella pertussis* clinical specimens investigated for PCR amplification of the pertactin and pertussis toxin S1 genes, as visualised in Figure 8.4 and 8.5 258
Table 8.3. *Bordetella pertussis* clinical DNA extracts investigated for detection of the pertactin gene, as visualised in Figure 8.6 261
Table 8.4. Pertactin and pertussis toxin S1 gene types from clinical specimens 264

LIST OF FIGURES

Figure 1.1. Virulence-activated factors of *Bordetella pertussis* 9

Figure 1.2. Alignment of the *Bordetella pertussis* pertactin gene (*prnA*) showing the tri-peptide motif RGD which mediates adhesion 16

Figure 1.3. Pertussis incidence and vaccine coverage in the United Kingdom, from 1940 to 2003 25

Figure 1.4. Alignment of the pertussis toxin subunit 1 (*ptxA*) gene of *Bordetella pertussis*, showing the polymorphic regions 48

Figure 2.1. Schematic of the *Bordetella pertussis* pertactin gene (*prnA*) showing regions of polymorphism 69

Figure 2.2. PCR amplification of the *Bordetella pertussis* pertactin gene with primers AF and BR, at various DNA concentrations and annealing temperatures 70

Figure 2.3. PCR amplification of the 5'-end of the *Bordetella pertussis* pertactin gene with primers PR-8F and PRN707R 71

Figure 2.4. PCR amplification of the 3'-end of the *Bordetella pertussis* pertactin gene with primers PRN1976F and PR5R, at various annealing temperatures 72

Figure 2.5. Schematic of the *Bordetella pertussis* pertussis toxin subunit 1 gene 74

Figure 2.6. PCR amplification of the *Bordetella pertussis* pertussis toxin S1 gene with primers SF and SR, at various magnesium chloride concentrations and in the presence or absence of dimethyl sulphoxide 75

Figure 2.7. PCR amplification of the *Bordetella pertussis* insertion element IS*1002*, at various annealing temperatures 83

Figure 2.8. IS*1002*-RFLP of *Bordetella pertussis*, with various hybridisation temperatures and concentrations of the IS*1002* probe 89

Figure 2.9. IS*1002*-RFLP of *Bordetella pertussis* and *Bordetella parapertussis*, with various washing conditions 90

Figure 2.10. Normalised IS*1002*-RFLP *Bordetella pertussis* patterns determined by Dice, UPGMA, and a 91.8% similarity cut-off 93

Figure 2.11. Normalised PFGE *Bordetella pertussis* patterns determined by Dice, UPGMA, and a 94% similarity cut-off 97

Figure 3.1. National Health Service regional boundaries of England and Wales before 1st April 1999 108

Figure 3.2. Pertactin gene variants between 1920 and 2002, in *Bordetella pertussis* isolates from the UK 110

Figure 3.3. Pertussis toxin S1 gene variants between 1920 and 2002 in *Bordetella pertussis* isolates from the UK 111
Figure 3.4. Temporal trends of combined *prnA/ptxA* types of UK *Bordetella pertussis* isolates 115
Figure 3.5. Normalised IS*1002*-RFLP types of UK *Bordetella pertussis* isolates from 1940 to 1999 116
Figure 3.6. Temporal trends of *Bordetella pertussis* according to the IS*1002*-RFLP band number in the UK 117
Figure 3.7. Changes in the population structure of *Bordetella pertussis* in the UK by IS*1002*-RFLP type 118
Figure 3.8. IS*1002*-RFLP patterns of *Bordetella pertussis* isolates from 1998 to 1999 in the UK 130
Figure 3.9 *Bordetella pertussis* IS*1002*-RFLP profiles from the UK and The Netherlands... 149
Figure 3.10. *Bordetella pertussis* serotype distribution in England and Wales, from 1995 to 2002 161
Figure 4.1. The state of Missouri, USA, with main cities located, including Kansas City . 164
Figure 4.2. Nucleic acid alignment of the *Bordetella pertussis* pertactin gene types, *prnA*(2) and *prnA*(9) 168
Figure 4.3. IS*1002*-RFLP types of *Bordetella pertussis* isolates from Missouri, USA 175
Figure 4.4. PFGE types, with *Xba*I, of *Bordetella pertussis* isolates from Missouri, USA 178
Figure 4.5. A dendrogram of *Bordetella pertussis* isolates from Missouri, USA, analysed by PFGE with *Xba*I, focusing on the clade containing PFGE-2 and PFGE-3 isolates 179
Figure 5.1. Reproducibility of *Bordetella pertussis* genotyping by IS*1002*-RFLP analysis...... 196
Figure 5.2. Reproducibility of *Bordetella pertussis* genotyping by PFGE using *Xba*I........ 197
Figure 5.3. Cluster analysis of 'epidemiologically-related' *Bordetella pertussis* isolates by IS*1002*-RFLP analysis 200
Figure 5.4. Cluster analysis of 'epidemiologically-related' *Bordetella pertussis* isolates by PFGE using *Xba*I 201
Figure 5.5. The observed IS*1002*-RFLP profile of the *Bordetella pertussis* Tohama-I strain 205
Figure 5.6. A comparison of observed and expected band sizes of the PFGE fragments of *Bordetella pertussis* Tohama-I with *Xba*I 207
Figure 6.1. Schematic of the *Bordetella pertussis* filamentous haemagglutinin gene (partial) 221
Figure 6.2. PCR amplification of the *Bordetella pertussis* adenylate kinase gene with primers adkF and adkR, at various annealing temperatures 222
Figure 6.3. Nucleic acid alignment of the *Bordetella pertussis* adenylate kinase gene........ 223

Figure 6.4. PCR amplification of the *Bordetella pertussis* filamentous haemagglutinin gene with primers fha-1705F and fha-3482R, at various annealing temperatures.............. 227
Figure 6.5. PCR amplification of the *Bordetella pertussis* filamentous haemagglutinin gene with primers fha-224F and fha-1366R, at various magnesium chloride concentrations and annealing temperatures 228
Figure 6.6. PCR amplification of the *Bordetella pertussis* filamentous haemagglutinin gene with primers fha-1219F and fha-2361R, at various magnesium chloride concentrations and annealing temperatures 229
Figure 6.7. PCR amplification of the *Bordetella pertussis* filamentous haemagglutinin gene with primers fha-2223F and fha-3482R, at various annealing temperatures.............. 230
Figure 6.8. Contiguous sequences of the *Bordetella pertussis* filamentous haemagglutinin gene.. 231
Figure 7.1. Pictorial representation of the single-enzyme AFLP method 236
Figure 7.2. AFLP profiles of *Bordetella pertussis*, using *Apa*I and selective primers Apa-C (a) and Apa-GC (b).......... 241
Figure 7.3. AFLP profiles of *Bordetella pertussis*, using *Eco*RI and selective primers Eco-C (a) and Eco-G (b) 243
Figure 7.4. AFLP profiles of *Bordetella pertussis*, using *Mse*I and selective primers Mse-A (a) and Mse-AT (b).......... 244
Figure 7.5. AFLP profiles of *Bordetella pertussis*, using *Pst*I and selective primers Pst-A (a) and Pst-C (b).......... 245
Figure 7.6. AFLP profiles of *Bordetella pertussis*, using *Taq*I and selective primer Taq-A.... 246
Figure 8.1. PCR amplification of the *Bordetella pertussis* pertactin gene with primers AF and AR, at various annealing temperatures 252
Figure 8.2. PCR amplification of the *Bordetella pertussis* pertussis toxin S1 gene with primers S1FM and SR, at various annealing temperatures.......... 254
Figure 8.3. Detection of the *Bordetella pertussis* pertactin gene with primers AF and BR, in UK clinical specimens 256
Figure 8.4. Detection of the *Bordetella pertussis* pertactin gene with primers AF and BR, in UK clinical specimens 259
Figure 8.5. Detection of the *Bordetella pertussis* pertussis toxin S1 gene with primers SF and SR, in UK clinical specimens 260
Figure 8.6. Detection of the *Bordetella pertussis* pertactin gene with primers AF and AR, in UK clinical specimens 262

ABBREVIATIONS

16S rRNA	16S ribosomal ribonucleic acid
ACV	Acellular vaccine
adk, ADK	Adenylate kinase, gene and protein
AFLP	Amplified fragment length polymorphism
bp	Base pairs
DNA	Deoxyribonucleic acid
CAMR	Centre for Applied Microbiology and Research
CDC	Centers for Disease Control and Prevention
CDSC	Communicable Diseases Surveillance Centre
CHO	Chinese hamster ovary cells
CI	Confidence interval
CPHL	Central Public Health Laboratory
DF	Degrees of freedom
fhaB, FHA	Filamentous haemagglutinin, gene and protein
fim, FIM	Fimbriae, gene and protein
GD	Genetic diversity
HPA	Health Protection Agency
IS	Insertion sequence element
IS*1002*-RFLP	IS*1002*-based restriction fragment length polymorphism
Kb	Kilobase
KDa	Kilo Dalton
l	Litre
M	Molar
mol% G+C	Guanine and cytosine percentage of DNA, as a relative molar amount
PCR	Polymerase chain reaction
PFGE	Pulsed-field gel electrophoresis
PHLS	Public Health Laboratory Service
PICU	Paediatric Intensive Care Unit
PRL	Pertussis Reference Laboratory
prnA, PRN	Pertactin, gene and protein
ptxA, PTX	Pertussis toxin S1 subunit, gene and protein
NCTC	National Collection of Type Cultures
NIBSC	National Institute for Biological Standards and Control
RGD	Arginine-glycine-aspartic acid motif

RSIL	Respiratory and Systemic Infection Laboratory
SNP	Single nucleotide polymorphism
T	Type strain
T_d	Temperature of disassociation
T_m	Melting temperature
U	Units
WCV	Whole-cell vaccine

ABSTRACT

The resurgence of pertussis in several highly vaccinated countries has stimulated this study of the genotypic diversity and epidemiological typing of *Bordetella pertussis*. The genotypic variation of *B. pertussis* was investigated in 318 UK clinical isolates from 1920-2002, using pertactin (*prnA*) and pertussis toxin S1 subunit (*ptxA*) gene typing. Isolates before 1979 were all (97/97) *prnA*(1), and identical to the three UK Wellcome whole-cell vaccine (WCV) strains. Since 1982, the novel types *prnA*(2) and *prnA*(3) were recovered, so that from 1998-2002, 81/133 (61%) isolates were non-*prnA*(1) (P=0.0006). Two *ptxA* variants, *ptxA*(1) and *ptxA*(2), were found and these both occur in the UK Wellcome WCV. The temporal expansion of *ptxA*(1) in UK isolates is also significant (P<0.0001), from 40/73 (55%) isolates between 1920-1956, to 133/133 (100%) *ptxA*(1) isolates between 1998-2002. These *ptxA* UK data argue against the theory of vaccine-driven evolution, proposed after a study in The Netherlands reported a shift in circulating *B. pertussis* strains towards non-vaccine types. To challenge this theory further, 149 isolates were randomly-picked from 1940-1999 to calculate the genotypic diversity (GD) in each decade using IS*1002*-based restriction fragment length polymorphism (RFLP) analysis. Of the 24 different RFLP types identified, RFLP-1 was present in all decades in 64/149 (43%) isolates. The three types RFLP-1, RFLP-2, and RFLP-3, which are present in the Wellcome WCV, constituted a major proportion of isolates in the 1990s at 32/36 (89%). The calculated GD values revealed a drop after the introduction of the pertussis WCV in 1957 from the 1940s (0.81) to the 1950s (0.64), (P=0.092), and a significant decrease from the 1980s to the 1990s (0.84 compared to 0.57, P<0.01). An association was seen between the UK pertussis notification figures and the GD values of each decade, hence, when pertussis notifications were high, the GD increased, which could be due to the clonal expansion of *B. pertussis*. Comparisons with other countries concerning *prnA* and *ptxA* gene types, RFLP types and genotypic diversity are discussed.

Ninety UK isolates, from September 1998 to April 1999, were investigated, along with corresponding clinical and epidemiological information, for any associations with the laboratory typing data. The most significant findings were that the under two month age group is mostly at risk of pertussis infection, and that cases admitted to hospital are more likely to be infected with a serotype 1,2 (P=0.017) and/or a *prnA*(1) (P=0.025) strain. Incidentally, there was a very significant association between serotype 1,2 with *prnA*(1), and serotype 1,3 with *prnA*(2) (P<0.001). These associations may be viewed as indicators of severe pertussis disease.

Isolates (n=44) from Missouri, USA were also characterised, after an increase in pertussis in the summer months of 1999. Epidemiological typing confirmed that RFLP-1 and RFLP-2 are successful global lineages of *B. pertussis*, and that the number of RFLP types dropped from 10 types in the 1980s to 4 types in the 1990s. A novel pertactin gene type, *prnA*(9), was found in two of the isolates and differs from *prnA*(2) by an extra GGFGP repeat. Ten isolates from Kansas City were all serotype 1,2, *prnA*(2), *ptxA*(1), and RFLP-1, but were divided into two PFGE types, PFGE-2 and PFGE-3, which emphasises the difficulty of choosing an optimum typing method for epidemiological investigations.

To date, the evaluation of typing methods used for *B. pertussis* has not been performed extensively. Therefore, in this thesis, the recognised methods, namely serotyping, pulsed-field gel electrophoresis (PFGE) using *Xba*I, and IS*1002*-RFLP analysis, were evaluated according to guidelines recommended by Struelens, M. J. and the Members of the European Study Group on Epidemiological Markers 1996, *Clin Microbiol Infect*, vol. 2, no. 1, pp. 2-11. It was found that, PFGE typing gave a good discrimination index value, but gave a low score for reproducibility, and further work is required to optimise this method. Furthermore, if *prnA* and *ptxA* gene typing were included in the evaluation, combined with serotyping, this combination would equal the discrimination of PFGE.

Other typing methods attempted for *B. pertussis* included direct sequencing of adenylate kinase (*adk*) and filamentous haemagglutinin genes (*fhaB*), and single-enzyme amplified fragment length polymorphism (AFLP) analysis with a selection of enzymes and selective primers. The type strain and a clinical strain, generated one and six single nucleotide polymorphisms (SNPs) in *adk* and *fhaB*, respectively. The discriminatory ability of the single-enzyme AFLP analysis was not satisfactory, as only a few different profiles were seen in the nine isolates tested. However, at least four AFLP profiles were generated using *Pst*I enzyme, and the selective primer Pst-C.

The direct detection and epidemiological typing of *B. pertussis* in 20 previously untypable clinical samples was attempted using *prnA* and *ptxA* as targets. Six clinical extracts generated *prnA* and *ptxA* (5/6) sequence data, therefore confirming that these typing procedures on *B. pertussis* PCR-positive clinical specimens is worthwhile in order to generate the *prnA* and *ptxA* distributions from babies or adults presenting atypical symptoms. This strategy should also be encouraged in other countries that have studied *prnA* and *ptxA* allele distributions, in order to update the representation of the genetic diversity of *B. pertussis*.

1 INTRODUCTION

1.1 *Bordetella* species

At the time of writing this thesis, eight *Bordetella* species had been described (Table 1.1, page 5). A chronological account of how these species became members of the *Bordetella* genus is given below.

1.1.1 Historical account of the *Bordetella* genus

An epidemic of whooping cough was first described in 1578, by the French physician de Baillou (Holmes, 1940.). In 1906, Joules Bordet and Octave Gengou isolated the causative agent from a patient in the acute stage of the disease (Table 1.1, page 5) (Bordet and Gengou, 1906). This organism, originally known as *Haemophilus pertussis* (Leslie and Gardner, 1931), is now the type species of the *Bordetella* genus (Pittman, 1984). The host range of *B. pertussis* is limited to humans and it is the only species in the genus to produce pertussis toxin, which may contribute to the prolonged paroxysmal coughing that is most severe with *B. pertussis* infection. *Bordetella bronchiseptica,* described in 1912, causes respiratory disease in a broad range of hosts. These diseases include kennel cough in dogs, atrophic rhinitis in pigs, and other illnesses in many other domestic and wild animals, including horses and rabbits (Ferry, 1912). However, it is an infrequent opportunist in humans (Pittman, 1984). *Bordetella parapertussis,* first described in 1938, generally causes a milder form of whooping cough in humans, and was originally differentiated from *B. pertussis* by its more rapid growth, and its ability to produce a brown pigment on tyrosine agar (Eldering and Kendrick, 1938). *Bordetella parapertussis* has also been isolated from sheep (Cullinane *et al.*, 1987). However, human and sheep isolates of *B. parapertussis* have distinct characteristics (van der Zee *et al.*, 1996a).

Until recently, classification of the genus *Bordetella* has been based primarily on the three 'classical' species described above, *B. pertussis, B. parapertussis* and *B. bronchiseptica* (Gerlach *et al.*, 2001). However, in 1984 Kersters and colleagues proposed the species *Bordetella avium* (originally identified as *Alcaligenes faecalis* (Simmons *et al.*, 1980)) to include organisms isolated from the respiratory tract of turkeys and other birds presenting turkey coryza or rhinotracheitis (see Table 1.1, page 5) (Kersters *et al.*, 1984). At this time, Volume I of Bergey's Manual of Systematic Bacteriology, described *B. avium* as a "*Bordetella bronchiseptica*-like bacterium", and defined the *Bordetella* genus as:

> "Minute coccobacillus, 0.2-0.5 μm in diameter and 0.5-2.0 μm in length, often bipolar stained, and arranged singly or in pairs, more rarely in chains. Gram-

negative. Nonmotile, two species; motile, one species, by peritrichous flagella. Strictly aerobic. Optimum temperature, 35-37°C. Colonies on Bordet-Gengou medium are smooth, convex, pearly, glistening, nearly transparent and surrounded by a zone of hemolysis without definite periphery. Metabolism respiratory, never fermentative. Chemoorganotrophic, require nicotinamide, organic sulfur (e.g., cysteine) and organic nitrogen (amino acids). Utilize oxidatively glutamic acid, proline, alanine, aspartic acid and serine with production of ammonia and CO_2. Litmus milk is made alkaline. Mammalian parasite and pathogen. Localize and multiply among the epithelial cilia of the respiratory tract. The mol% G+C of the DNA is 66-70 (T_m)."

(Pittman, 1984)

After prior assignments to *Alcaligenes*, *Brucella* and *Haemophilus* (Pittman, 1984), de Ley and colleagues (1986), proposed that the four *Bordetella* species should be included in a new family, *Alcaligenaceae*, together with *Achromobacter* and *Alcaligenes* species, based on DNA-rRNA hybridisation and phenotypic analyses.

A group of isolates which were Gram-negative, non-oxidising, soluble, brown pigment-producing rods were initially classified in the Centers for Disease Control and Prevention (CDC), Atlanta, United States of America (USA) as CDC non-oxidiser group 2 (NO-2) (Weyant *et al.*, 1995). These isolates were found mostly in the blood of young adults, and therefore were not initially considered as members of the *Bordetella* genus, possibly due to the lack of association with respiratory disease (Weyant *et al.*, 1995). However, after genetic and chemotaxonomic analyses, these isolates were described as *Bordetella holmesii* (Table 1.1, page 5) (Weyant *et al.*, 1995). Subsequently, they have been isolated from nasopharyngeal specimens (Mazengia *et al.*, 2000; Yih *et al.*, 1999). Isolation of *B. holmesii* has been described in patients with certain underlying conditions such as sickle-cell anaemia (Njamkepo *et al.*, 2000) and those undergoing haemodialysis (Greig *et al.*, 2001). Another *Bordetella* species, *Bordetella hinzii* was also described in 1995, and is similar to *Bordetella avium* in that they have both been isolated from birds (Vandamme *et al.*, 1995). *Bordetella hinzii* has also been isolated from immuno-compromised patients, such as those with acquired immune deficiency syndrome (AIDS) (Cookson *et al.*, 1994) and cystic fibrosis (Funke *et al.*, 1996). *Bordetella trematum* was described in 1996 (Vandamme *et al.*, 1996) and has to date, not been isolated from the respiratory tract, but from wounds and ear infections of humans. Genotypic and phenotypic comparisons of these three *Bordetella* species have placed them in the *Bordetella* genus.

Prior to 2001, all members of the *Bordetella* genus were obligatory parasites of either humans, other mammals or birds, and mostly isolated from the epithelial cilia of the respiratory tract (Table 1.1, page 5). *Bordetella pertussis, B. parapertussis* and *B. bronchiseptica* are capable of intracellular survival in host cells (Forde *et al.*, 1998; Friedman *et al.*, 1992; Lee *et al.*, 1990). Furthermore, *B. bronchiseptica* and *B. parapertussis* have been shown to grow and survive in natural waters and other low-nutrient conditions, suggesting that they might have an environmental reservoir (Porter *et al.*, 1991; Porter and Wardlaw, 1993). This suggestion was further supported by the description in 2001 of *Bordetella petrii*, which was isolated from river sediment and enriched in an anaerobic, dechlorinating bioreactor (von Wintzingerode *et al.*, 2001). This new species, through phylogenetic and chemotaxonomic comparisons, was found to be most closely related to the *Bordetella* species (von Wintzingerode *et al.*, 2001). Although *B. petrii* is an environmental isolate and a facultative anaerobe, this designation has been accepted, but its pathogenicity and animal hosts, if any, are unknown.

1.1.2 Amended definition of the *Bordetella* genus

The description of *B. petrii* has caused the definition of the *Bordetella* genus to be modified to:

> "Gram-negative, catalase-positive, asaccharolytic coccobacilli with an DNA [mole percent of guanine and cytosine content] G+C content of 60-69 mol%. Growth occurs strictly aerobically, facultatively anaerobically and non-fermentatively. One species [*B. petrii*] is able to grow anaerobically under conditions favouring respiratory nitrate and selenate reduction. Species assimilate citrate under aerobic conditions, one species [*B. petrii*] assimilates D-gluconate. … … … Strains were isolated from humans and warm-blooded animals, one strain [*B. petrii*] was isolated from the environment. Based on comparative 16S rRNA gene sequence analysis, the species forms a distinct cluster separated from members of the genera *Achromobacter* and *Alcaligenes* within the *β-Proteobacteria*. The type species of the genus is *B. pertussis*."

(von Wintzingerode *et al.*, 2001)

Thus, the *Bordetella* genus has expanded to include recently described species.

Table 1.1. Differential properties of *Bordetella* species

Characteristic	*B. pertussis*	*B. bronchiseptica*	*B. parapertussis*	*B. avium*	*B. hinzii*	*B. holmesii*	*B. trematum*	*B. petrii*
Year of description	1906[1]	1912	1937	1984	1995	1995	1996	2001
Host range	Humans	Broad range (e.g., dogs, pigs, horses, rabbits, humans)	Humans, sheep	Birds	Birds, humans	Humans	Humans	n/a[2]
Site(s) of isolation	RT[3]	RT	RT	RT	RT	RT, blood	Wounds, ear	n/a[2]
Motility	–	+	–	+	+	–	+	-
Growth (days)[4]	3-4	1	1-2	1	2	2-3	1	2
Growth on MacConkey's agar	–	+	+	+	+	+/-	+	+
Brown pigment[5]	–	–	+	–	–	+	–	–
Urease	–	+	+	–	–	–	–	–
Oxidase	+	+	–	+	+	–	–	+
Nitrate reduction	–	+	–	–	–	–	+/-	–
G+C content (mol %)	67.7-68.9	68.2-69.5	68.1-69.0	61.6-62.6	65.0-67.0	61.5-62.3	64.0-65.0	63.3-64.3

[1]Year *B. pertussis* was first isolated as *H. pertussis* and later re-classified as the type species of the *Bordetella* genus in 1952 (Pittman, 1984). [2]n/a=not applicable, as *B. petrii* was isolated from the environment. [3]RT=respiratory tract. [4]To form visible colonies, on appropriate media. [5]When grown on media containing tyrosine.

Adapted from Gerlach *et al.* (2001), Parton (1998), von Wintzingerode *et al.* (2001) and (2002).

1.1.3 Differentiation of species within the *Bordetella* genus

Bordetella species can be differentiated by a number of phenotypic tests including motility, presence of flagella, and the time taken for appearance of visible colonies on agar plates (Table 1.1, page 5). Other useful tests include the lack of growth on MacConkey agar for *B. pertussis* and some *B. holmesii* strains, and that only *B. parapertussis* and *B. holmesii* are able to produce a brown pigment when grown on agar containing tyrosine. Biochemical tests such as urease and oxidase production are also useful for differentiating the *Bordetella* species (Roberts and Parton, 2001; von Wintzingerode *et al.*, 2001 and 2002).

1.2 *Bordetella* virulence factors and pathogenesis

Consistent with their close genetic relationship, the three "classical" *Bordetella* species, *B. pertussis*, *B. parapertussis* and *B. bronchiseptica*, have many virulence factors in common (Table 1.2, page 7). These can be divided into two classes dependent on their function: toxins including pertussis toxin (PTX) and adenylate cyclase toxin (ACT), and adhesins such as filamentous haemagglutinin (FHA) and fimbriae (FIM). Most of the important virulence factors have been characterised for *B. pertussis*, and those central to this thesis are described further.

1.2.1 Regulation of virulence factor expression

The regulation of the expression of *Bordetella* virulence factors has been studied for many years. Infrequent changes in virulence properties of fresh isolates were first described by Leslie and Gardner (1931) as 'phase variation'. Lacey, in 1960, introduced the term 'antigenic modulation', where *B. pertussis* isolates, when grown under submodulating conditions, can reversibly convert from an X mode (virulent), to the C mode (avirulent) via an intermediate mode (I) (Lacey, 1960). In 1984, Weiss and Falkow (1984) brought together the infrequent action of phase variation and the freely reversible antigenic modulation by suggesting a common regulatory mechanism was involved, identified as the virulence regulatory locus, *vir*, and later re-named the bordetella virulence gene (*bvg*).

Table 1.2. Virulence factors of *Bordetella* species

Virulence factor		*B. pertussis*	*B. parapertussis*	*B. bronchiseptica*	*B. avium*	*B. hinzii*	*B. holmesii*	*B. trematum*	*B. petrii*
Pertussis toxin	expression[1]	+	–	–	–	–	–	–	NT
	genes[2]	+	+	+	–	–	–	–	NT
Adenylate cyclase toxin	expression[1]	+	+	+	–	–	–	–	NT
	genes[2]	+	+	+	–	–	–	–	NT
Dermonecrotic toxin	expression[1]	+	+	+	+	NT	NT	NT	NT
	genes[2]	+	+	+		–	–	–	NT
Tracheal cytotoxin	expression[3]	+	+	+	+	NT	NT	NT	NT
Endotoxin[4]	expression	+LOS	+LPS	+LPS	+	NT	NT	NT	NT
Filamentous haemagglutinin	expression[1]	+	+	+	–	–	–	NT	NT
	genes[2]	+	+	+	–	–	–	NT	NT
Fimbriae	expression[1]	+	+	+	+	NT	NT	NT	NT
	genes[2]	+	+	+		NT	NT	NT	NT
Pertactin	expression[1]	+	+	+	–	NT	–	NT	NT
	genes[2]	+	+	+	–	–	NT	NT	NT
Bordetella resistance to killing protein	expression[1]	+	NT	some strains	NT	NT	NT	NT	NT
	genes[2]	+	+	+	–	NT	NT	NT	NT
Tracheal colonisation factor	expression[1]	+	–	–	NT	NT	NT	NT	NT
	genes[2]	+	+	+	NT	NT	NT	NT	NT
Virulence-activated gene 8	expression[1]	+	–	+	NT	NT	NT	NT	NT
	genes[2]	+	+	+	NT	NT	NT	NT	NT

[1]Detected by Western blotting; [2]detected by DNA sequence analysis or Southern blotting; [3]method of detection not stated; [4]LOS=lipooligosaccharide, LPS=lipopolysaccharide; NT=not tested.

Adapted from Parton (1998) and Gerlach *et al.* (2001).

The *bvg* locus is present in *B. pertussis*, *B. parapertussis*, *B. bronchiseptica* and *B. avium* (Arico *et al.*, 1991; Gentry-Weeks *et al.*, 1991), and codes for two proteins, BvgA and BvgS, as part of a two-component system (Figure 1.1, page 9). The *bvg* locus is activated (Bvg$^+$ phase) when the temperature is above 30°C and no nicotinic acid or magnesium sulphate ($MgSO_4$) is present. Transcription of many virulence-activated genes (*vags*) seems to be activated at different times after transfer to the non-modulating conditions (Uhl and Miller, 1994; Uhl and Miller, 1995). 'Early activated' promoters include those genes encoding adhesins FIM and FHA, whilst the 'late activated' *vag* genes encoding PTX and ACT, are expressed two hours later (Scarlato *et al.*, 1991). The gene encoding pertactin (PRN), an adhesin, is intermediately transcribed (Kinnear *et al.*, 1999). At low temperatures, or in the presence of nicotinic acid or $MgSO_4$, the *vag* genes are silent, and another set of genes, the virulence-repressed genes (*vrg*) are expressed (Bvg$^-$ phase). These genes, studied more extensively in *B. bronchiseptica*, are involved in motility, iron-scavenging, urease and phosphatase activity (Akerley *et al.*, 1992; Giardina *et al.*, 1995; McMillan *et al.*, 1996; Chhatwal *et al.*, 1997). An intermediate *bvg* phase (Bvg^i phase) has been reported recently, and is characterised by the expression of BipA, a bordetella Bvg-intermediate phase protein (Stockbauer *et al.*, 2001).

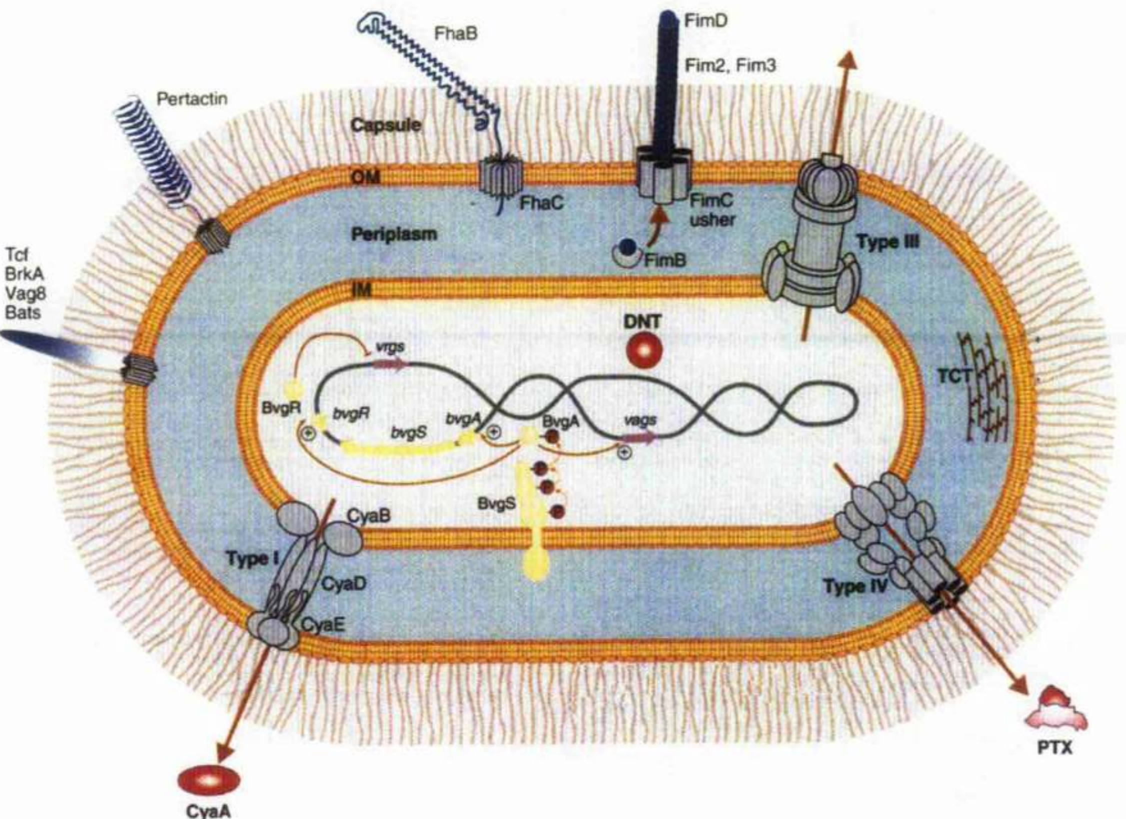
A second two-component regulatory system (regulator of intracellular response, the *ris* locus) has been described recently in *B. bronchiseptica* and *B.pertussis*, although its proposed role in bacterial resistance to oxidative stress, the production of acid phosphatase and *in vivo* persistence has been observed only in *B. bronchiseptica* (Jungnitz *et al.*, 1998).

1.2.2 Toxins

1.2.2.1 Pertussis toxin

Pertussis toxin is produced only by *B. pertussis*, and although *B. bronchiseptica* and *B. parapertussis* both possess the genes for this complex protein, they are silent and therefore the protein is not transcribed. The protein PTX is a 117 KDa A-B exotoxin, similar to those produced by *Corynebacterium diphtheriae* and *Vibrio cholerae*. It is made up of five different polypeptides (S1 to S5) in a 1:1:1:2:1 ratio, with the S1 subunit comprising the A protomer, and the B cell-binding subunit involving the other polypeptides S2 to S5 (Roberts and Parton, 2001).

Figure 1.1. Virulence-activated factors of *Bordetella pertussis*



Adapted from Locht *et al.* (2001) with kind permission. *Bordetella pertussis* is depicted with inner and outer membranes (IM and OM), a periplasm and a capsule. The adhesins fimbriae (Fim2, Fim3 and FimD), filamentous haemagglutinin (FhaB), pertactin, tracheal colonisation factor (Tcf), bordetella resistance to killing (BrkA), virulence-activated gene 8 (Vag8) and bordetella autotransporters (Bats) are shown in blue; the toxins pertussis toxin (PTX), adenylate cyclase toxin (CyaA, ACT in text), dermonecrotic toxin (DNT) and tracheal cytotoxin (TCT) are in red; the accessory proteins FhaC (translocation of FhaB), FimB and FimC (fimbrial biogenesis), Type I (translocation of CyaA), Type III and Type IV (translocation of PTX), secretion apparatus are in grey; and the bordetella virulence gene regulatory system (BvgA, BvgS and BvgR) are in beige.

The B oligomer forms a triangular structure for the S1 subunit for correct insertion into the cytoplasm of mammalian cells (Stein *et al.*, 1994). The S2 and S3 subunits constitute the receptor-binding domain, preferentially binding to glycoproteins. After receptor binding, the S1 subunit dissociates from the B oligomer and is translocated through the endoplasmic reticulum membrane, probably via the retrograde transport system (el Baya *et al.*, 1997). A set of nine pertussis toxin liberation genes (*ptl*) are involved in the transportation of the pertussis toxin through the inner and outer membranes of *B. pertussis* (Farizo *et al.*, 2002). Transportation of PTX through the outer membrane could be due to the formation of a pilus-like structure of the *ptl* products, where extra PTL proteins are added on to the structure to push out the pertussis toxin (see Figure 1.1, page 9). The *ptxA* and *ptl* loci are present in *B. parapertussis* and *B. bronchiseptica*, but due to mutations in the *ptxA* promoter gene, they are not transcribed and therefore, there is no expression (Locht and Antoine, 1999).

The S1 subunit (the A protomer) possesses an enzymatic activity, specifically causing the adenosine diphosphate (ADP)-ribosylation of a host-cell G protein, G_i, that normally inhibits mammalian adenylate cyclase (AC) activity. This results in continuous activation of AC by another stimulatory G protein, G_S. The S1 subunit has various pathophysiological and immunomodulating activities, such as histamine and serotonin sensitisation, activation of insulin secretion, stimulation of haemagglutination and inhibition of neutrophil oxidative burst (Mills, 2001). The immunomodulating activities make PTX a good candidate for inclusion in acellular vaccines (ACVs) (Munoz, 1985; Mills, 2001). Formaldehyde-toxoided PTX (PTXd) gives protection in mice against both intra-cerebral (i.c.) and aerosol challenges with *B. pertussis* (Sato and Sato, 1984). Genetically-inactivated pertussis toxin developed by Peppoloni and colleagues (1995) for inclusion in ACVs, contains three T cell antigenic sites in the S1 subunit: 1 to 42, 181 to 211, and 212 to 235 amino acids. These sites in the S1 subunit have been shown by Mooi *et al.* (1998) to exhibit polymorphisms at the genetic level, and this will be described later (1.6.4, page 46).

1.2.2.2 Adenylate cyclase toxin

Adenylate cyclase toxin (ACT) is a bifunctional 177 KDa protein belonging to the RTX (repeat in toxin) family, and is produced by *B. pertussis*, *B. parapertussis* and *B. bronchiseptica* (Roberts and Parton, 2001). It possesses weak haemolytic activity and, because of its calmodulin-dependent adenylate cyclase activitiy, is capable, after gaining entry, of intoxicating a wide range of eucaryotic cells, including neutrophils, monocytes, macrophages and natural killer cells (NKCs) (Hewlett and Gordon, 1988). Antibodies

against adenylate cyclase are produced during pertussis infection and in humans vaccinated with pertussis whole-cell vaccines (WCVs) (Farfel *et al.*, 1990). Adenylate cyclase toxin is protective in mice against respiratory tract challenge with *B. pertussis* (Guiso *et al.*, 1991; Hormozi *et al.*, 1999), and there have been suggestions to include a detoxified form of ACT in ACVs.

1.2.2.3 Dermonecrotic toxin

Dermonecrotic toxin (DNT), becomes completely inactive when heated at 56°C for 10 minutes, hence the alternative name, heat-labile toxin (Nakase and Endoh, 1988). The toxin was later renamed dermonecrotic toxin, because it can cause necrotic lesions when injected subcutaneously into rabbits, mice or guinea pigs (Locht and Antoine, 1999). Dermonecrotic toxin is highly lethal when injected intravenously in mice (LD_{50}=6.0 ng), and the mouse toxicity test for WCV safety was introduced by Pittman and Cox (1965), chiefly to ensure DNT was destroyed during vaccine production. This 160 KDa toxin is produced in *B. pertussis*, *B. parapertussis* and *B. bronchiseptica*, and was one of the first bordetella toxins to be described. *Bordetella avium* also produces DNT, but at a slightly lower mass of 140 KDa (Gentry-Weeks *et al.*, 1988). Dermonecrotic toxin is not secreted extracellularly and is localised in the cytoplasm. This, and also the very low expression level of DNT, has made it problematic to purify the protein (Smith *et al.*, 2001).

1.2.2.4 Tracheal cytotoxin

Tracheal cytotoxin (TCT) was discovered by Goldman and colleagues (1982), due to its ability to cause ciliostasis and ciliated cell extrusion in hamster tracheal organ cultures, and inhibition of DNA synthesis in hamster tracheal epithelial cells. A similar action is seen in the primary cytopathology of pertussis in humans and in turkey coryza (Goldman, 1988). A non-protein toxin, TCT is a low mass (921 Da) disaccharide-tetrapeptide, derived from the peptidoglycan of the bacterial cell envelope and is released extracellularly (Goldman, 1988). Interleukin-1 production is induced by TCT, which then activates host cell nitric oxide (NO) production, which is the actual cytotoxic factor (Heiss *et al.*, 1994). Cytokine-inducible NO synthase is, however, detected only when TCT and *B. pertussis* lipopolysaccharide (described below) are both present (Flak and Goldman, 1999). Tracheal cytotoxin also inhibits neutrophil chemotaxis and oxidative metabolism, which may contribute to *B. pertussis* survival (Mills, 2001). Tracheal cytotoxin production, unlike other virulence factors of *Bordetella*, is not controlled by the BvgA/S regulatory system (1.2.1, page 6) (Locht, 1999).

1.2.2.5 Lipopolysaccharide

In common with other Gram-negative bacteria, *Bordetella* species produce lipopolysaccharide (LPS) or endotoxin (Roberts and Parton, 2001). Known as lipooligosaccharide in *B. pertussis* due to the absence of the polysaccharide O antigen moiety, it is a major component of the outer membrane and comprises of a lipid and a core oligosaccharide. The lipid domain possesses the potent endotoxin activity, has adjuvant properties and can induce resistance to virus infection (Kerr and Matthews, 2000). The core oligosaccharide induces pyrogenicity, and mitogenicity is seen in spleen cell cultures (Locht and Antoine, 1999). *Bordetella pertussis* has two different forms of LPS, (LPS-A and LPS-B) and these differ in the structure of the core oligosaccharide moiety (Chaby and Caroff, 1988). The ratio of LPS-A to LPS-B can also differ between strains of *B. pertussis* (Locht and Antoine, 1999). Lipooligosaccharide from *B. pertussis* together with TCT have been shown to induce inflammation in the respiratory mucosa and cause damage to ciliated epithelial cells (Flak and Goldman, 1999). The stimulation of secretion of both pro-inflammatory and anti-inflammatory cytokines, makes lipopolysaccharide a major factor in the systemic and neurological responses to infection (Mills, 2001).

1.2.3 Adhesins

1.2.3.1 Filamentous haemagglutinin

The most studied and major adhesin of *B. pertussis*, filamentous haemagglutinin (FHA), is also produced by *B. parapertussis* and *B. bronchiseptica*, and an unrelated haemagglutinin is present in *B. avium* (Parton, 1998). Filamentous haemagglutinin is produced as a 370 KDa precursor in *B. pertussis* and is proteolytically processed at both the amino (N-) and carboxyl (C-) terminals, resulting in the mature 220 KDa form. The predicted structure of FHA based on electron microscopy, circular dichroism spectroscopy and computational analysis, is a filamentous monomeric molecule of approximately 2-4 nm in diameter and 40-50 nm long (Figure 1.1, page 9) (Makhov *et al.*, 1994). The protein contains two large repeat regions: one from 344-1065 amino acids consisting of 38 copies of a 19-residue sequence, and the other between 1440-1688 amino acids of 13 copies of a different 19-residue sequence. These two regions are thought to form short β-strands connected by two β-turns, which may form a hairpin-like structure, with the head composed of the two terminal regions, the shaft consisting of the β-sheets, and the tail containing the important arginine-glycine-aspartic acid (RGD) motif.

Filamentous haemagglutinin is encoded by the 10.1 Kb gene *fhaB*, and consists of an N-proximal haemolysin-homologue region, the two large repeats, an immunodominant domain and the C-terminal region (Delisse-Gathoye *et al.*, 1990). The secretion and post-translational maturation requires another gene, *fhaC*. These two genes are separated by genes encoding proteins involved in fimbrial synthesis, and are known as the *fim/fha* gene cluster. This indicates a close relationship between the expression of these adhesins. The protein FhaC interacts with the 115-residue haemolysin-homologue region and forms a β-barrel, outer-membrane pore for FHA to translocate through the outer membrane (Smith *et al.*, 2001). The 150 KDa C-terminal domain of FHA is then cleaved and the mature FHA protein can form the hairpin structure.

The N-terminus of FHA exhibits heparin-binding to epithelial cells such as WiDr (a human colon carcinoma cell line) and Chinese hamster ovary (CHO) cells (Brennan *et al.*, 1991; Menozzi *et al.*, 1994). The RGD binding site of FHA promotes adhesion to the macrophage integrin CR3, which may lead to phagocytosis without triggering an oxidative burst. The immuno-dominant domain of FHA via carbohydrate-binding mediates the attachment of FHA to cilia, and has been demonstrated in human ciliated cells *in vitro* and ciliated cells in animals (Weiss, 1997). Filamentous haemagglutinin inhibits the classical pathway of complement activation by binding to C4b-binding protein (Berggard *et al.*, 1997). Secretion of IL-10 by macrophages, which is mediated by FHA, inhibits IL-12 production, and may explain the suppressed T-cell response in the lungs of mice during *B. pertussis* infection (Mills, 2001). A strong immunogen, purified FHA protects mice against aerosol challenge with *B. pertussis*. Filamentous haemagglutinin has therefore been included in many second-generation ACVs due to its adhesin and protective properties.

1.2.3.2 Fimbriae

Bordetella pertussis, B. parapertussis, B. bronchiseptica, and *B. avium* all produce fimbriae, also known as pili (Parton, 1998). In *B. pertussis*, they are composed of major and minor subunits, the major subunits forming a pentameric repeat unit, each 13 nm long and about 5 nm in diameter (Steven *et al.*, 1986). The major subunits, Fim2 and Fim3 are 22.5 KDa and 22.0 KDa, and are identical to serotype 2 and serotype 3 agglutinogens respectively. The minor subunit FimD, is 40.0 KDa, and is the adhesin located at the tip of the Fim2 and/or Fim3 stacked structures.

The *fim2* and *fim3* genes lie independently in the genome, as does the silent pseudogene *fimX*. Accessory proteins essential for fimbrial biogenesis include a periplasmic chaperone (FimB), an outer membrane transport/anchor protein (FimC), and the minor fimbrial subunit FimD, described above (Willems *et al.*, 1992).

Agglutination properties of *B. pertussis* are now known to be attributed, at least in part, to the fimbriae, after evidence that anti-fimbrial antibodies in the sera of infected or vaccinated children agglutinated the bacteria (Locht, 1999). This property has been developed as a serotyping scheme to follow epidemiological trends in *B. pertussis* populations (Eldering *et al.*, 1957). These schemes initially used more than three agglutinogens, but presently, France, The Netherlands and the UK detect three agglutinogens by slide agglutination using either antisera from rabbits or monoclonal antibodies (Preston, 1963; Mooi, 1994; Guiso *et al.*, 2001b). The serotypes of *Bordetella pertussis* are controlled by the expression of Fim2 and Fim3; either Fim2 or Fim3 determine serotypes 1,2 or 1,3, respectively; expression of both Fim2 and Fim3 simultaneously, results in a serotype 1,2,3 strain; and when expression of both Fim2 and Fim3 is absent, serotype 1 strains are recorded (Patton, 1998). This is dependent on the level of transcription, and the switching between the activation of the *fim2* and *fim3* promoters, in a process called serotype variation. The fimbrial promoters each possess a long run of cytosine residues, which can vary in number due to insertions or deletions of extra cytosine residues, caused by slipped-strand mispairing during chromosome replication (Mooi, 1994). The crucial aspect is thought to be an optimum distance between the putative –10 box and the activator (BvgA) binding site for transcription to occur, which results in the expression of fimbriae (Smith *et al.*, 2001).

The minor fimbrial subunit FimD, binds to VLA-5 receptors and promotes attachment of fimbriae to monocytes (Hazenbos *et al.*, 1993), and the authors have proposed that this interaction could be critical for *B. pertussis* intracellular survival. Two regions of *fim2*, H1 and H2, can also bind heparin, a sulphated sugar which is ubiquitous in the respiratory tract (Geuijen *et al.*, 1998). Recent studies using epithelial cell lines derived from the human tract indicate that fimbriae from *B. pertussis* play a role in the infection of the laryngeal mucosa (Locht *et al.*, 2001).

A correlation was also observed between the presence of anti-fimbrial antibodies and subsequent protection in infected or vaccinated children (Locht, 1999), and purified fimbriae protect mice against respiratory infection with *B. pertussis* (Ashworth *et al.*, 1988).

Due to their protective and antigenic properties, both Fim2 and Fim3 have been included in some ACVs (Parton, 1998).

1.2.3.3 Pertactin

Only produced by *B. pertussis, B. parapertussis,* and *B. bronchiseptica,* the surface-associated adhesin pertactin (PRN) was initially designated P69 or OMP69 in *B. pertussis* due to its apparent relative molecular mass of 69 KDa, following sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) (Charles *et al.*, 1989). The PRN proteins of *B. parapertussis* and *B. bronchiseptica* are 70 KDa and 68 KDa respectively and all are transcribed as larger precursor proteins. The PRN precursor in *B. pertussis* is 93.5 KDa, and is processed at both the N-terminus, to cleave the 3 KDa signal peptide responsible for translocation of the protein across the inner membrane, and at the C-terminus, which forms a channel through the outer membrane for the mature pertactin protein to reach the cell surface (Charles *et al.*, 1994). Homologues of this C-terminal domain are also seen in other Gram-negative proteins that are collectively known as autotransporters (Henderson *et al.*, 1998), since they do not require accessory proteins for biogenesis, unlike PTX and FHA (Weiss *et al.*, 1993; Smith *et al.*, 2001).

The crystal structure of the mature pertactin protein has been elucidated as a monomer composed of 16 parallel strands that form a β-helix, which is V-shaped in cross-section (Figure 1.1, page 9). Loops protrude from the β-helix core and they have been associated with the biological activities of pertactin (Emsley *et al.*, 1996). These include an RGD motif, also found in FHA (1.2.3.1, page 12), and mediates binding of pertactin to CHO cells and possibly integrins on mammalian cell surfaces (Kerr and Matthews, 2000). Two other loops contain proline-rich repeat regions: the first repeat, region 1, is located close to the RGD motif and has the sequence Gly-Gly-x-x-Pro (GGxxP), and the second repeat, region 2, is located close to the C-terminus and bears the sequence Pro-Gln-Pro (PQP) (Figure 1.2, page 16) (Emsley *et al.*, 1996). Both regions are immunodominant B-cell epitopes, as revealed by mapping these regions with pertactin-specific monoclonal antibodies (Charles *et al.*, 1991). These two proline-rich, immunodominant regions have recently been found to be polymorphic within strains of *Bordetella*, differing in either the number of repeats or the amino acids present in the repeat. Most polymorphisms are located at region 1 in *B. pertussis* (Mooi *et al.*, 1998), and at region 2 in *B. bronchiseptica* (Boursaux-Eude and Guiso, 2000), and their significance will be discussed later (1.6.4, page 46).

Figure 1.2. Alignment of the *Bordetella pertussis* pertactin gene (*prnA*) showing the tri-peptide motif RGD which mediates adhesion. Part A, illustrates region 1 coding for $(GGxxP)_{4-7}$ and part B shows region 2 coding for $(PQP)_{4-5}$

GenBank Accession Number	prnA type	305	390	778	791	region 1	994	1010	1212	1595	region 2	2558
AJ011091	prnA(1)	TCC S	GTG V	CGC GGG GAC R G D	CCT P		TCG S	TCT S	TCC S	CTA L		CAC H
AJ011092	prnA(2)	...	...	....	...		...	...	...	...		...
AJ011093	prnA(3)	...	...	....	...		...	...	...	...		...
AJ011015	prnA(4)	...	...	....	...		...	...	...	...		...
AJ011016	prnA(5)	...	...	....	...		...	...	...	...		...
AJ132095	prnA(6)	.T. P	..G	....	...		...	.T. F	..T	.G. R		.G. R
AJ133784	prnA(7)	...	...	....	...		...	...	...	.G. R		...
AJ133245	prnA(8)	...	...	-----------	...		...	...	...	.G. R		...
AF218785	prnA(9)	...	...	....	...		...	...	...	...		...
AJ507542	prnA(11)	...	...	....	...		...	...	...	...		...
J04560	prnA(1)	...	...	....	.T. L		C.. P	...	...	...		...

A. Region 1 of the pertactin gene of *B. pertussis*

```
                 795
                 |
AJ011091 prnA(1) GGCGGTGCGGTTCCC CGCGGTGCGGGTCCC GGTGGTGCGGTTCCC --------------- --------------- GGCGGCTTCGGTCCC GGCGGCTTCGGTCCC
                  G  G  A  V  P   G  G  A  V  P   G  G  A  V  P                                    G  G  F  G  P   G  G  F  G  P
AJ011092 prnA(2) ............... ............... --------------- GGCGGCTTCGGTCCC GGCGGCTTCGGTCCC ............... ...............
                                                                  G  G  F  G  P   G  G  F  G  P
AJ011093 prnA(3) ............... ............... --------------- --------------- ............... ............... ...............
AJ011015 prnA(4) ............... ............... --------------- --------------- --------------- ............... ...............
AJ011016 prnA(5) ............... --------------- --------------- --------------- ............... ............... ...............
AJ132095 prnA(6) ......G........ ............... ..G............ --------------- --------------- ............... ...............
                       G
AJ133784 prnA(7) ............... ............... ............... --------------- --------------- ............... ...............
AJ133245 prnA(8) ............... ..T............ --------------- --------------- --------------- ............... ...............
AF218785 prnA(9) ............... ............... GGCGGCTTCGGTCCC ............... ............... ............... ...............
                                                  G  G  F  G  P
AJ507642 prnA(11) ............... ............... ..G............ GGTGGTGCGGTTCCC --------------- ............... ...............
                                                                  G  G  A  V  P
```

B. Region 2 of the pertactin gene of *B. pertussis*

```
                  1735
                  |
AJ011091 prnA(1)  CCGCAGCCG GGT CCCCAGCCC CCCCAGCCG CCGCAGCCG CAGCCGGAAGCGCCGGCG CCGCAACCG
                   P  Q  P   G   P  Q  P   P  Q  P   P  Q  P   Q  P  E  A  P  A   P  Q  P
AJ011092 prnA(2)  ......... ... ......... ......... ......... .................. .........
AJ011093 prnA(3)  ......... ... ......... ......... ......... .................. .........
AJ011015 prnA(4)  ......... ... ......... ......... ......... .................. .........
AJ011016 prnA(5)  ......... ... ......... ......... ......... .................. .........
AJ132095 prnA(6)  ......... ... ......... --------- ......... .................. .........
                   P  Q  P   G   P  Q  P   -  -  -   P  Q  P   Q  P  E  A  P  A   P  Q  P
AJ133784 prnA(7)  ......... ... ......... ......... ......... .................. .........
AJ133245 prnA(8)  ......... ... ......... ......... ......... .................. .........
AF218785 prnA(9)  ......... ... ......... ......... ......... .................. .........
AJ507642 prnA(11) ......... ... ......... ......... ......... .................. .........
```

This figure shows ten pertactin types. Only *prnA*(1) through to *prnA*(5) were described before work on this thesis began. Non-silent mutations are shown in bold and associated amino acid changes are indicated beneath the relevant codons. Shaded nucleotides indicate silent mutations. Dots indicate sequence identity and dashes indicate that the sequence is not found. Numbers indicate the position of nucleotides relative to the start codon of AJ011091. Please refer to Appendix 1 for amino acid designations.

Pertactin-deleted mutants of *B. pertussis* were found to be capable of colonising and multiplying in the respiratory tract of mice, implying that PRN is not essential for bacterial persistence (Roberts *et al.*, 1991). Other research using site-directed mutagenesis to change a PRN-derived RGD motif (Arg-Gly-Asp) to a RGE motif (Arg-Gly-Glu), also showed no difference in their ability to promote adhesion to Hep-2 or CHO cells (Everest *et al.*, 1996). There has been some evidence that PRN and FHA functionally interact for efficient cell binding (Arico *et al.*, 1993), but the exact role and mechanism of pertactin in adherence has yet to be elucidated.

Pertactin was initially discovered during the development of a veterinary vaccine against *B. bronchiseptica* and found to be a good protective immunogen (Locht, 1999). Pertactin from *B. pertussis* has also been demonstrated to protect mice against aerosol challenge with *B. pertussis* and, together with FHA, also protects mice against i.c. challenge (Roberts *et al.*, 1991). This immunogenicity has made pertactin become one of the main constituents in the preparation of ACVs.

1.2.3.4 Bordetella resistance to killing protein

The second autotransporter to be described, bordetella resistance to killing protein (BrkA), has 29% overall similarity to the nucleotide sequence of pertactin and shares 55% amino acid similarity in the C-terminal autotransporter domain (Fernandez and Weiss, 1994). The *brkA* gene, also translated by *B. parapertussis* and some strains of *B. bronchiseptica*, contains two RGD motifs and two potential binding sites for sulphated glycoconjugates, suggesting its role in adherence to host cells. *Bordetella pertussis* BrkA mutants are approximately ten-fold more susceptible to classical-pathway complement-dependent killing by human sera. It was also shown that *B. pertussis* strains carrying an extra copy of the *brk* locus secreted increased levels of BrkA (Fernandez and Weiss, 1998). The *brk* locus contains two genes that are transcribed in opposite directions, *brkA* and *brkB* (Fernandez and Weiss, 1994), and are essential for serum resistance and the secretion of the BrkA protein. The immunogenic properties of BrkA have so far not been studied, and this protein has therefore not been considered as an ACV component.

1.2.3.5 Tracheal colonisation factor

The third autotransporter to be identified was the tracheal colonisation factor (TCF) (Finn and Stevens, 1995). While it is only expressed in *B. pertussis*, the gene coding for TCF is also present in *B. parapertussis* and *B. bronchiseptica*. A *B. pertussis* mutant lacking TCF was found

to have a reduced ability to colonise the trachea, but not the lungs of mice (Finn and Stevens, 1995). The C-terminal domain shows 50% amino acid similarity to the pertactin equivalent (1.2.3.3, page 15) and thus is predicted to translocate the mature 64 KDa TCF protein through the outer membrane. The protein also contains an RGD motif and is proline-rich, similar to PRN and BrkA (Finn and Stevens, 1995). Pertactin, BrkA, and TCF are all surface-associated proteins. However, TCF has also been detected in culture supernates as a slightly smaller 60 KDa secreted form (Finn and Stevens, 1995). There is a significant reduction of *B. pertussis* tracheal colonisation in mice when immunised with attenuated *V. cholerae* expressing TCF (Chen *et al.*, 1998), although the protective potential of TCF has yet to be fully determined.

1.2.3.6 Virulence-activated gene 8

Through screening for *bvg*-controlled genes that were positively regulated, the virulence-activated gene 8 (*vag8*) was identified (Knapp and Mekalanos, 1988). The gene is expressed in *B. pertussis* and *B. bronchiseptica* as a 95 KDa protein. A *vag8* gene homologue is also present in *B. parapertussis*, but is not transcribed. Classed as an autotransporter, Vag8 possesses similar characteristics to PRN, BrkA, and TCF, although it is not known whether the C-terminal domain in Vag8 is cleaved, as the predicted size corresponds to the size determined by SDS-PAGE (Finn and Amsbaugh, 1998). Conflicting results have been reported concerning the pathogenesis of Vag8 using a mouse aerosol model (Finn and Amsbaugh, 1998). A mutant Vag8 *B. pertussis* strain was defective in colonisation in both the lungs and the trachea, but strains with an internal deletion in the *vag8* structural gene could colonise as well as the wild-type. The authors suggest that this loss in function of *vag8* may be compensated for by other virulence factors (Finn and Amsbaugh, 1998). In *B. bronchiseptica*, it has recently been reported that Vag8 is required for type III secretion (Foreman-Wykert *et al.*, 2002), but type III secretion is not expressed in *B. pertussis* (Yuk *et al.*, 1998), and hence the role of Vag8 in *B. pertussis*, if any, is unknown.

1.2.4 Other virulence factors

With the aid of a new suicide vector, pFUS2 (Antoine *et al.*, 2000), microarray analysis (Hot *et al.*, 2003), and the recent availability of the three *Bordetella* genomes, *B. pertussis, B. parapertussis,* and *B. bronchiseptica* (Parkhill *et al.*, 2003), novel genes have been discovered recently with similarities to well-characterised virulence factors. These include the bordetella autotransporter protein C (*bapC*, originally described as *bap5*), a 79.5 KDa predicted protein possessing an RGD motif (Blackburn, 2000). Recent work has also

demonstrated that a *bapC* deficient strain of *B. pertussis* was less able to colonise the respiratory tract in mice than the wild-type and that the mutant also had reduced complement-mediated killing (Bokhari, 2002). A positively regulated BvgAS, iron-mediated outer membrane protein, Vir90, has also been described, and Passerini de Rossi *et al* (2003) predict one function of this newly characterised protein as a receptor for ferrichrome or a siderophore. Fifteen other virulence factor genes have been submitted to GenBank, including one coding for a cold-shock protein (Phg) (Antoine *et al.*, 2000).

1.3 Clinical aspects and diagnosis of pertussis

1.3.1 Clinical symptoms, complications, and treatment of pertussis infection

Transmission of *B. pertussis* from an infected to a susceptible person is via droplet spread from the respiratory tract, followed by an incubation period of seven to 14 days. Classical pertussis can be split into three stages: catarrhal, paroxysmal, and convalescent (Roberts and Parton, 2001). The catarrhal stage may last from three days to two weeks, and resembles a "common" viral cold, with the infected individual showing symptoms such as a mild cough, profuse mucoid rhinorrhoea, and sometimes a mild fever. These symptoms can be easily confused with those caused by other respiratory pathogens but, with prior knowledge of contact history of the patient, treatment can be given to reduce or stop the organism's spread from person to person. Prophylaxis (e.g., erythromycin) is best given at this time to help shorten or even abort the pertussis illness (Linnemann *et al.*, 1975; Dodhia *et al.*, 2002).

Without medical treatment, the mild cough develops into a paroxysmal cough, which can last from one to six weeks. A paroxysmal episode consists of six to ten forceful expiratory coughs, followed by the characteristic "whoop" sound of the inhalation, caused by the air rushing in the partially closed glottis (Walker, 1988). Up to 20 or more of these paroxysmal spasms can occur in 24 hours, where they are more common at night, and they are frequently interrupted by vomiting. The World Health Organisation (WHO) case definition of pertussis is a minimum of paroxysmal coughing for 21 days with laboratory confirmation or epidemiological linkage (World Health Organisation, 1991). This severe stage is both detrimental and tiring for the patient, and most complications occur during this stage. Apnoea, subconjunctival haemorrhage, hernia, and rectal prolapse are some of the complications which develop due to the pressure effects of the paroxysms. Other complications can result in pneumonia and otitis media, and very rarely, central nervous

system complications such as seizures and encephalopathy, can also occur (Roberts and Parton, 2001).

Many cases of pertussis infection require hospitalisation, especially in infants under six months old, and up to 50% can be hospitalised (Health Protection Agency, 2003c). Convalescence follows the paroxysmal stage and can last between a few weeks to many months, depending on the severity of disease and the complications encountered. Death due to complications can occur, with the majority of cases being under six months old (Edwards *et al.*, 1999). A recent study using data from the UK arrived at an estimate of 28 deaths due to pertussis between 1995 and 1997 (van Buynder *et al.*, 1999).

Clinical symptoms of pertussis are lessened in vaccinated children, adolescents and adults. In adults, a persistent cough with a mean duration of 36 to 48 days is typical, and sleep can be affected. Whooping, vomiting and sweating attacks can also occur, with complications such as rib fracture, incontinence, back pain and secondary infection being reported (von Konig *et al.*, 2002).

1.3.2 Laboratory diagnosis of *Bordetella pertussis* infection

Culture and isolation of *B. pertussis* from suitable respiratory samples is considered to be the gold standard for the laboratory diagnosis of *B. pertussis* infection (Roberts and Parton, 2001). The medium first used for isolation contained blood, potato extract, and glycerol and was named after the developers, Bordet and Gengou (BG) (Bordet and Gengou, 1906). Bordet-Gengou medium is still used today, as well as Regan-Lowe medium, a charcoal agar supplemented with 10% horse blood. Liquid media include Stainer-Scholte broth containing (2,6-O-dimethyl)-β-cyclodextrin (Roberts and Parton, 2001). The antibiotic cephalexin can also be added to the media to inhibit normal nasopharyngeal flora growth. The sensitivity of culture can differ widely, and isolation rates are higher; (i) before the paroxysmal stage; (ii) from an aspirate rather than a swab specimen; (iii) after a shorter transport time; and (iv) from younger, unvaccinated patients than from older, vaccinated ones (Roberts and Parton, 2001). Antibiotic therapy, if given, will also reduce the rate of *B. pertussis* isolation (Birkebaek, 2001). Consequently, sensitivity of culture can be as low as 20-40% (Heininger, 2001) and can take three to four days for colonies to appear (Parton, 1998). Thus, faster, more sensitive methods of diagnosis would help in the administration of correct treatment, and hence reduce transmission.

Serological assays have been evaluated for laboratory diagnosis of *B. pertussis* for over 30 years, originally using complement fixation and agglutination assays (Anon, 1970). These have been replaced roughly 20 years ago with antigen-specific assays, namely, enzyme-linked immunosorbent assays (ELISAs) (Viljanen *et al.*, 1982). These assays have been valuable in vaccine efficacy and epidemiology studies due to their sensitivity (57-100%) (von Konig *et al.*, 2002). For diagnosis, ELISAs have been used to detect IgG antibodies specific to PT, FHA, PRN, or FIM. Only PT is specific for *B. pertussis*, as other bordetellae can produce the other named factors that can induce an immune response, such as antibodies to FHA, which can be detected after an infection with *B. parapertussis* (Michaels, 1998). Paired sera are desirable to show a recent infection, the first serum sample ideally taken two weeks after the onset of cough, and the second about two to three weeks later (Kerr and Matthews, 2000). For diagnosis however, a single serum sample can be used, but this requires prior knowledge of background values from healthy control individuals and defined cut-off values (de Melker *et al.*, 2000b; Nardone *et al.*, 2004). Due to these difficulties, and that ELISAs require highly skilled personnel to generate and analyse the data, inter-laboratory comparisons can be problematic. Serology is useful in showing recent infections of *B. pertussis*, but does not affect the treatment of the individual, and therefore a more rapid method is required to affect the transmission of *B. pertussis*.

The development of diagnostic polymerase chain reaction (PCR) assays over recent years has allowed health care professionals to rapidly detect a wide range of bacteria and viruses causing disease in man (Louie *et al.*, 2000; Elnifro *et al.*, 2000). This rapid and sensitive method has been applied to the fastidious, slow-growing *B. pertussis*, targeting genes such as the pertussis toxin promoter (Houard *et al.*, 1989), and the repeated insertion sequence, IS*481* (Glare *et al.*, 1990). Using the same respiratory sample as for culture isolation, *B. pertussis* DNA can be extracted, amplified and detected in less than a day. Consensus recommendations have been published to minimise that no false positives are detected (Meade and Bollen, 1994), partly by including internal and external controls. Other advantages of PCR-based assays include the detection of *B. pertussis* in clinically atypical, vaccinated or antibiotic-treated patients. Sensitivity rates can reach three-fold higher for PCR detection compared to culture isolation (Fry *et al.*, 2004; Heininger, 2001). Major technological advances in recent years have yielded machinery capable of simultaneously amplifying and detecting *B. pertussis* DNA within four hours (Reischl *et al.*, 2001; Kosters *et al.*, 2001).

Problems with PCR techniques include DNA extraction efficiency, where there are relatively low amounts of *B. pertussis* DNA compared to host (i.e., human) DNA, technical difficulties with the assay, PCR inhibitors and false positives. This latter problem has been found recently in a real-time PCR targeting for IS*481* (Reischl *et al.*, 2001). The *B. pertussis* Tohama-I genome contains 238 copies of the IS*481* element (Parkhill *et al.*, 2003), and this element has also been detected in the newly discovered *Bordetella* species, *B. holmesii*. Only eight to ten copies of IS*481* were found in *B. holmesii* strains, but this is sufficient to produce a positive result in the sensitive real-time PCR assay targeting the IS*481* element. The authors therefore recommend cautious use of this IS*481* assay, and suggest the concurrent use of another target (Reischl *et al.*, 2001). The IS*481* element was also found in two *B. bronchiseptica* strains isolated from humans, although the insertion element has not been found in other *B. bronchiseptica* strains (Gladbach *et al.*, 2002). Despite this, PCR is a powerful tool in detecting *B. pertussis* within a sufficient time frame to affect transmission and treatment (Birkebaek, 2001). However, it has been recommended that culture isolation should continue, as strain variation, antibiotic resistance and other characteristics may not be studied by PCR-based systems alone (Hewlett, 2000).

1.4 Epidemiology and prevention

Before mass vaccination against pertussis was introduced, *B. pertussis* was responsible for childhood morbidity and mortality worldwide (Roberts and Parton, 2001). An endemic disease, pertussis epidemics were seen every three to four years, occurring in approximately 20% of infants younger than one year, and up to 60% in children between one to four years, peaking at three years old (Edwards *et al.*, 1999). Pertussis was an early childhood infectious disease and mortality rates were highest in neonates. Due to this severe and highly transmissible disease, vaccines against pertussis were formulated soon after the isolation of the causative organism in 1906, and the first whole-cell pertussis vaccine (WCV) was licensed in the USA in 1914 as a suspension of killed *B. pertussis* organisms (Edwards *et al.*, 1999).

1.4.1 Pertussis vaccination in the United Kingdom

The WCV was gradually introduced in the UK from 1937 and immunisation against pertussis became national policy in 1957 (Grant and Cherry, 2002). Initially the Burroughs Wellcome WCV was comprised of two *B. pertussis* strains, CN4132 and CN2992, exhibiting serotypes 1,2 or 1,2,3, respectively. One study showed that circulating strains before 1958 were serotypes 1,2 and 1,2,3, and, by 1963 to 1964, strains had shifted to mostly serotype

1,3, and it was suggested that a vaccine containing serotype 1,2 could not protect against a pertussis infection by a serotype 1,3 strain (Preston, 1965). Furthermore, a study by the Public Health Laboratory Service (PHLS) Whooping-cough Committee and Working Party reported that serotype 1,2,3 was isolated more frequently from unvaccinated than from vaccinated children (Abbott *et al.*, 1973). Therefore, a third strain, CN5476 (serotype 1,3), was added in 1966 to include all three serotypes and the potency of pertussis vaccine was increased according to the *British Pharmocopoeia* requirements. In 1979, the WHO recommended that all pertussis vaccine preparations should contain serotype antigens 1,2, and 3 (World Health Organisation, 1979). The pertussis WCV is administered as part of a combined vaccine with diphtheria and tetanus toxoids (DTP). In the UK, since 1990, the previous schedule at three, five, and ten months has been replaced by the accelerated immunisation schedule at two, three, and four months of age.

Following the introduction of WCV in the UK, cases of pertussis dropped significantly (Figure 1.3, page 25), although the three-yearly cyclic peaks are still present. This indicates that, whilst the number of pertussis cases have been reduced, transmission of the organism is still occurring, albeit at a reduced rate (Rohani *et al.*, 2000). The vaccine uptake was over 80% between 1969 to 1973, and pertussis notifications fell from 85,000 in 1957, to 2400 in 1974, thus lowering the incidence of pertussis in all age groups (Immunisation Division, CDSC, HPA, 2004a).

Despite the success of the WCV, adverse effects, such as redness, swelling and pain at the site of injection in half of the children in receipt of the vaccine, have been reported (Edwards *et al.*, 1999). Other common reactions are fever, irritability, drowsiness and excessive crying. Convulsions and seizures have also been associated with the WCV, but are rare (Edwards *et al.*, 1999). In 1974 and 1977, two reports linked the pertussis component of the DTP vaccine with acute neurological disorders in children in the UK (Department of Health and Social Security, 1977). This was widely publicised in the media, and public confidence towards pertussis vaccination fell. Vaccine uptake then dropped dramatically to 30% in 1978. This low vaccine coverage led to two pertussis epidemics, peaking in 1978 and 1982 (Figure 1.3, page 25) (Immunisation Division, CDSC, HPA, 2004a).

Further studies by a number of groups, including the American Academy of Neurology, the British Paediatric Association, the National Advisory Committee on Immunization (Canada) and the Advisory Committee on Immunization Practices (USA), concluded that

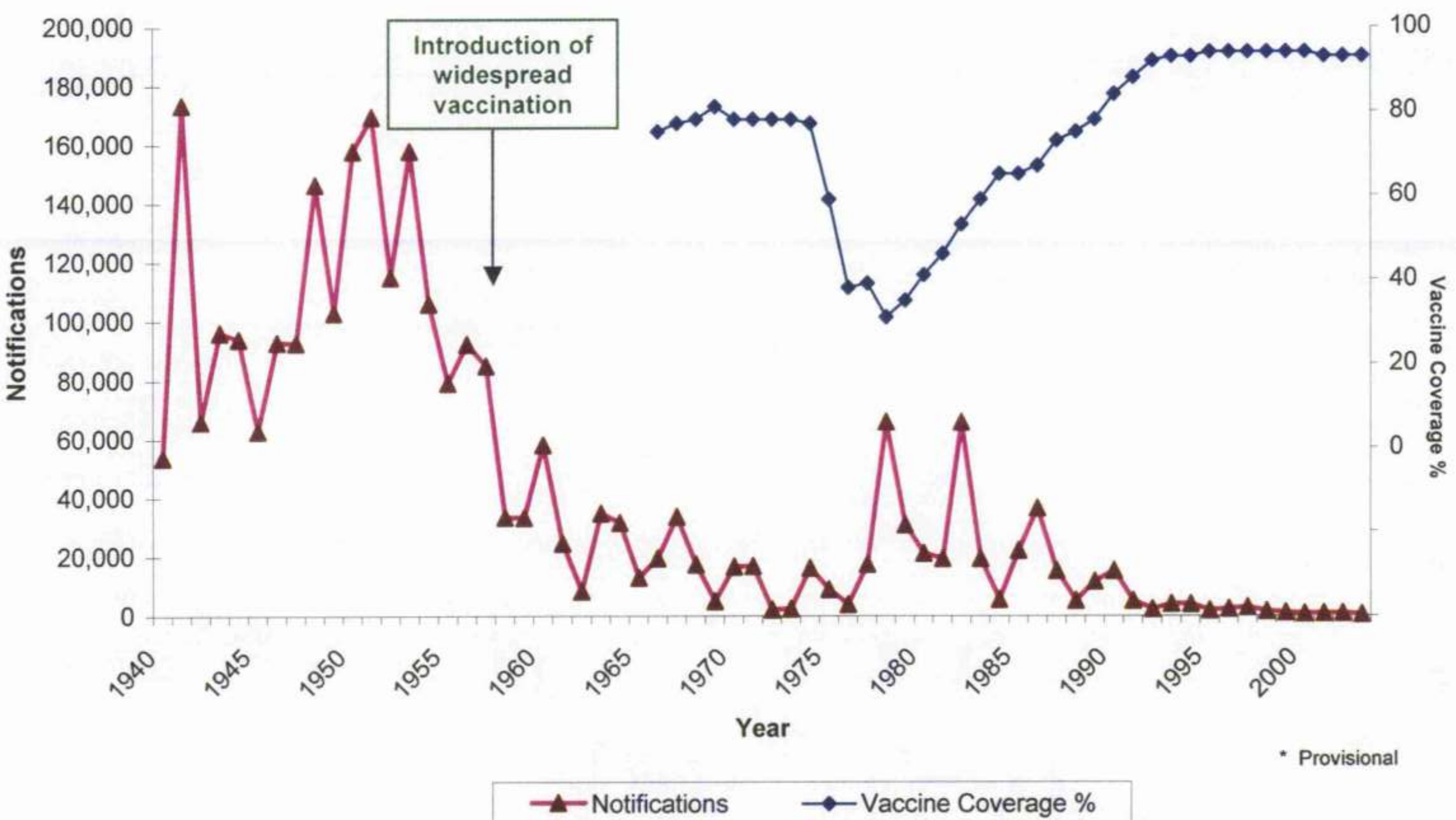
these data did not link encephalopathy with the pertussis vaccine (Edwards *et al.*, 1999). Gradually, public confidence was regained and vaccine uptake slowly recovered to current levels of 93% (Immunisation Division, CDSC, HPA, 2004a). The reported incidence of pertussis in the UK is now historically low, and only 409 cases in England and Wales were reported in 2003 (Figure 1.3, page 25). This drop in incidence may however be due to under-reporting of pertussis cases, which is described later (1.4.4, page 28) (Crowcroft and Britto, 2002).

1.4.2 Acellular vaccines

Concerns about the safety and adverse reactogenicity of the WCV (1.4.1, page 22) led to the development of acellular vaccines (ACVs). First introduced in 1981 in Japan, six licensed ACVs initially consisted of formaldehyde-toxoided PT (PTXd) and FHA, with some also containing PRN and FIM as contaminants. Due to the exclusion of LPS, these vaccines were less reactogenic (Edwards *et al.*, 1999). Results from these first ACVs convinced other countries and companies to develop second-generation vaccines using other detoxification methods, including gluteraldehyde, hydrogen peroxide, and genetic engineering (Edwards *et al.*, 1999). Large-scale trials in Sweden, Germany, Italy and Senegal, tested the efficacy and reactogenicity of various ACVs. These ACVs differed in amount and number of components, the purification method and adjuvant used. The trials were also different in design and the vaccine regimen tested, so comparing results directly is difficult. In general, the efficacy of ACVs ranged from 59% to 93%, vaccines containing three to five antigens were better than those comprising only one or two, and ACVs were less reactogenic than WCVs (Edwards *et al.*, 1999).

With the success of these trials, pertussis ACVs combined with vaccines against diphtheria and tetanus toxoids (DTaP) have been licensed and used in many countries, including the USA, Japan and most European countries (Edwards *et al.*, 1999). These have been successful in reducing pertussis notifications and, due to their low reactogenicity, ACVs can be used as boosters (Edwards *et al.*, 1999). Pertussis boosters have been introduced in other European countries such as France, Germany, and The Netherlands (Therre and Baron, 2000). In addition to DTaP, multi-combination vaccines have been developed. Several countries, including Denmark, Austria and Italy are using tetravalent or pentavalent vaccines with *Haemophilus influenzae* b and/or inactivated poliovirus (DTaPHib, DTaPIPV or DTaPHibIPV) (Therre and Baron, 2000).

Figure 1.3. Pertussis incidence and vaccine coverage in the United Kingdom, from 1940 to 2003*



* Provisional

Graph obtained from the Immunisation Division, Communicable Diseases Surveillance Centre, Health Protection Agency (2004a).

The enhanced surveillance system was introduced in 1994 to monitor whooping cough in England and Wales, the vaccine efficacy, and to assess the need for infant pertussis booster vaccination (Miller *et al.*, 1994). This surveillance entails the collaboration between the Communicable Diseases Surveillance Centre (CDSC) and Manchester University to follow up all culture-confirmed cases of pertussis for information on vaccination history, severity, antibiotic prophylaxis/treatment and history of exposure. Early analysis showed that mortality and morbidity was still prevalent in the very young and suggested booster vaccination could reduce this (van Buynder *et al.*, 1999). Further studies supported this proposal through unimmunised infant case reports (Ranganathan *et al.*, 1999; Crowcroft *et al.*, 2003), surveillance data (Beard and Finn, 2000), and potential health gain and cost consequence estimates (Edmunds *et al.*, 2002; Stevenson *et al.*, 2002). In November 2001, the Department of Health introduced a pre-school acellular vaccine (ACV) booster (around four to five years old) in England and Wales, with a recommendation to introduce the booster to Scotland and Northern Ireland (Health Protection Agency, 2003c). This intervention has yet to be evaluated for its impact on pertussis notifications and improvement in vaccine efficacy.

1.4.3 Pertussis in other countries

Currently, 80% of the world's children are vaccinated against pertussis, mostly with WCV. Despite this, there are still 20 to 40 million cases and 200,000 to 400,000 deaths due to pertussis each year (Anon, 1999a). Over 90% of these cases are from developing countries and the WHO Expanded Programme of Immunisation (EPI) is slowly reducing this, but their goal of lowering the incidence of pertussis to below one in 100,000 by 2010 may not be achieved.

In contrast, many developed countries with good immunisation programmes have managed to reduce pertussis incidence to low levels (Anon, 1999a). However, pertussis vaccination programmes have been dramatically affected by the media and anti-vaccine movements, such as in the UK (1.4.1, page 22), Japan and Sweden. In Japan, nearly 80% vaccine coverage was achieved in 1974, but because of the resulting low incidence, and reports of neurological reactions resulting after DTP vaccination, one province switched to a vaccine containing just diphtheria and tetanus toxoids (DT) (Gangarosa *et al.*, 1998). The diphtheria and tetanus vaccine was then quickly introduced throughout the whole country after two infants died within 24 hours of receiving DTP. With no pertussis vaccine, an epidemic occurred in 1979 culminating in 13,000 cases and 41 deaths. Japan later

introduced ACVs, and pertussis incidence fell sharply to levels seen before the interruption. Sweden's vaccine was also suspended in 1979 after suggestions that the WCV had poor efficacy and links to neurological events. Pertussis incidence increased four-fold, major outbreaks occurred and incidents of serious complications approached global rates. After several major clinical studies of various ACVs in Sweden, the country introduced widespread immunisation with an ACV in 1996, and notification figures are now falling (Anon, 1999b). Ireland and Australia were also indirectly affected after the adverse media coverage in the UK. Expressed as "passive anti-vaccine movements", this resulted in low vaccine coverage, a raised incidence and major outbreaks in Ireland and Australia (Gangarosa *et al.*, 1998). Events in these countries illustrate how important it is to sustain high vaccination rates to help reduce the possibility of pertussis epidemics from occurring.

In some countries with high, sustained vaccination coverage, there have been reports of outbreaks and rising pertussis incidence (Cherry, 1996), including Canada (De Serres *et al.*, 1995) and the USA (Guris *et al.*, 1999b). In Quebec, and other parts of Canada, outbreaks occurred yearly from 1988 to 1994, but there was no explanation for their cause (De Serres *et al.*, 1995). With the replacement of ACV in 1997, the resurgence of pertussis in Canada subsided, but a large outbreak in British Columbia in 2000 has re-established concern. Recent data indicate that there is an increase of pertussis in ten to 19 year olds, and a decrease amongst infants and pre-school children (Skowronski *et al.*, 2002). These authors now believe that the introduction of the ACV, a waning immunity and changes in laboratory methods may account for this apparent resurgence of pertussis (Skowronski *et al.*, 2002). Like Canada, the USA has also seen an increase of pertussis in older children (ten to 14 years) from 1980 to 2000, and consequently, pertussis is regarded as a re-emerging infection (Roberts and Parton, 2001). An increase of pertussis has also been seen in infants under four months of age since 1991 (Guris *et al.*, 1999b). A change from WCV to ACV has not affected this increase, so improved diagnosis and reporting in these two age groups are thought to be contributing to this apparent resurgence (Guris *et al.*, 1999b).

Australia has reported pertussis epidemics since 1993, despite vaccination coverage approaching 90%. Due to an observed increased rate in children aged five to 14 years, a fifth WCV dose was introduced in 1995 for children aged four to five years, and this appears to have lessened the burden in younger children (Andrews *et al.*, 1997). The introduction of improved diagnostic methods such as serology in the 1980s and PCR later, may have also contributed to the apparent increased incidence in pertussis (World Health Organisation, 2001). The Netherlands experienced a pertussis epidemic in 1996, with 2,771

cases compared to just 319 in 1995. Vaccine coverage after three doses of WCV was at 96% when the epidemic happened, but the cause was unknown (de Melker *et al.*, 1997).

All of these countries with apparently high pertussis vaccine coverage have reported an increased incidence in pertussis. Many factors have been suggested for this resurgence, including; (i) waning vaccine-induced immunity; (ii) a decrease in vaccine quality; (iii) improved surveillance and laboratory diagnosis; (iv) a recent change in case definition of pertussis, amended in 1991 by WHO; and (v) the expansion of strains which are antigenically distinct from the vaccine strains (Mooi *et al.*, 1998). The last theory, proposed by Mooi and his colleagues in The Netherlands (1.6.4, page 46) (1998), was tested during the course of this thesis and the results and conclusions are described later.

1.4.4 Recent changes in the dynamics of pertussis disease

Mortality and morbidity due to pertussis have decreased due to immunisation, with a shift in the age distribution of those affected, from mostly infants and young children, to younger unvaccinated neonates, adolescents and adults (Crowcroft and Britto, 2002). In the USA, 28% of reported cases occurred in subjects ten years or older in 1992 to 1994, compared to just 3% in over 15 year olds before vaccination was introduced (Edwards *et al.*, 1999). This increase in adolescents and adults has also been seen in Australia, Canada, the UK and other countries in Europe, to 10 to 12% of all cases reported in over fifteen year olds (Anon, 1999a). Better laboratory diagnosis using serology and PCR-based assays may account for this reported increase in adolescents and adults, where culture isolation alone may have caused under-reporting. However, the adult population, initially assumed to be immunised, is now thought to represent those with waning immunity and can be susceptible to subsequent pertussis infections (von Konig *et al.*, 2002). Furthermore, when infected, these patients' symptoms may be atypical and the disease can go undetected. Through studies of households, hospitals, and outbreaks, it is now thought that these adolescents and adults transmit pertussis to vulnerable unimmunised young infants or those that are not yet fully immunised (Rosenthal *et al.*, 1995; von Konig *et al.*, 2002; Christie *et al.*, 2001; Crowcroft *et al.*, 2003). Together with the high mortality rate in these young infants, the possible prolonged illness, and loss of work hours, various proposals to introduce boosters for adolescents have been put forward (Edwards, 2001; Heininger, 2001; von Konig *et al.*, 2002). In 1998, France introduced an ACV booster for 11 to 13 year olds (Fletcher *et al.*, 2001) and, very recently, Germany and Canada have also recommended adolescent boosters (Anon, 2003a and 2003b). Administering pertussis vaccine to health-

care workers has also helped in containing outbreak situations. In Germany, paediatric workers are required by law to receive boosters (von Konig *et al.*, 2002).

Although the reported incidence of pertussis in England and Wales is low, a *B. pertussis* surveillance study, compiling data from various sources in England and Wales between 1995 to 1997, analysed the composite data-set for any significant associations between clinical, epidemiological and laboratory data (van Buynder *et al.*, 1999). Death due to pertussis occurred only in infants less than six months old and was thought to be under-estimated. This has been exemplified with a recent report concluding that national mortality statistics under-estimate deaths from pertussis in England, and that the enhanced laboratory surveillance identifies most of these deaths (Crowcroft *et al.*, 2002). Unvaccinated subjects were more likely to be hospitalised than vaccinated subjects, and severity of pertussis disease was reduced when appropriate antibiotics were given to case contacts before onset (van Buynder *et al.*, 1999). An association between serotype 1,2 and severe disease was also found, together with an increase of serotype 1,2 in strains isolated. This prevalence of serotype 1,2 has also been found in The Netherlands with an increase of disease (de Melker *et al.*, 2000a; Mooi, 1994).

Pertussis epidemiology appears to be dynamic, dependent on vaccination status, diagnostic methods and surveillance programmes. In Europe, new projects such as the European seroepidemiology network (ESEN) and the European network for vaccine preventable diseases (EUVAC-NET) will aid standardisation of diagnosis, vaccine coverage and surveillance across several countries (World Health Organisation, 2001; Schmidt *et al.*, 2001). In developing countries, however, pertussis causes high morbidity and mortality, although reports in South-East Asian and African regions of vaccine coverage are sparse (World Health Organisation, 2001). The introduction of ACVs has increased public confidence towards pertussis vaccination due to its efficacy and low reactogenicity. However, the high cost of ACVs compared to WCVs may be unaffordable to developing countries. Whilst there is dependence on private pharmaceutical companies to produce both, an assurance that WCVs will be available to poorer countries is needed. When vaccination was introduced, the eradication of *B. pertussis* was envisaged, but because of resurgence in developed countries and poor vaccine distribution in developing countries, this seems to be unlikely. Basic laboratory surveillance and control measures need to be focused in countries with high mortality to reduce pertussis globally (Crowcroft and Britto, 2002).

1.5 Genotypic typing of bacteria

1.5.1 Caveats to typing bacteria

Microbiologists and epidemiologists use various methods to characterise bacteria from infection in a single host to global bacterial populations, and these can involve either short-term or long-term investigations (Struelens *et al.*, 1996). Fields such as bacterial population genetics, pathogenesis of infection, epidemiological surveillance and outbreak investigations can employ genotypic typing to discriminate between genomes. The study of population genetics on a large data-set of isolates from various backgrounds can generate information on the clonality of an organism. This can then form a basis for the study of pathogenesis, as it can highlight differences between groups of bacteria and their virulence properties. Epidemiological surveillance entails combining clinical and laboratory data over time as an ongoing process, which then enables the identification of trends such as disease frequency, epidemic clones, drug resistance and risk factors in a target population (Struelens *et al.*, 1996; Struelens, 1998). An outbreak can be defined as;

> "an increased incidence of an infectious disease in a specific place during a given period that is above the baseline rate for that place and time frame".
>
> (Tenover *et al.*, 1995)

In an outbreak investigation, the use of genotypic methods can allow the tracking of important clones (a group of isolates within the same type) and the identification of reservoirs of infection and vehicles of transmission. If performed within a rapid time frame, it allows health-care personnel the potential to control infection by containing or interrupting the spread of the organism (Struelens *et al.*, 1996).

Several recommendations or criteria have been proposed to evaluate bacterial typing systems for microbial epidemiology (Maslow *et al.*, 1993) including typability, reproducibility, discriminatory power, stability and epidemiological concordance. The typability is the proportion of isolates that can be assigned a type, within the test population. Reproducibility is the ability of the typing system to correctly assign a strain to the same type repeatedly. This criterion is thought to be the most difficult to ascertain for many typing techniques and, as the reproducibility affects the discriminatory power, it is also one of the most important (Struelens *et al.*, 1996). The discriminatory power is the average probability of two unrelated strains being assigned different types, as shown by Simpson's index of diversity (Hunter, 1990). The stability of markers can be affected by recombination events on the bacterial genome, and variations can occur after prolonged storage or in large-scale epidemics (Struelens *et al.*, 1996). This variation may be acceptable

if the typing system can still show the clonal relatedness of epidemiologically-linked strains. The epidemiological concordance is the probability that presumed single-clone outbreaks result in epidemiologically-related strains that can be assigned to the same type or clone (Struelens *et al.*, 1996).

Most of the above criteria: typability, reproducibility, discriminatory power, stability and epidemiological concordance can be calculated definitively with a formula, and their values vary from zero to one (Table 1.3, page 32) (Struelens *et al.*, 1996). Other criteria that are useful in selecting an appropriate typing system include rapidity, availability, cost, ease of performance and interpretation.

The typing methods used to differentiate bacterial strains and their clones can be sorted into two groups, phenotypic and genotypic (Maslow *et al.*, 1993). Phenotypic methods rely on the ability of an organism to express the characteristic chosen to be typed, and are dependent on the genes controlling these characteristics. This can produce low typability, reproducibility or discriminatory values and therefore, genotypic methods have been developed for typing bacteria, as they test the genome or genes directly (Maslow *et al.*, 1993). For the purpose of this thesis, only the genotypic methods and some suggested applications are described.

1.5.2 Genotypic methods

Genotypic methods can involve cleavage by restriction endonucleases, amplification via the polymerase chain reaction (PCR), and direct amplification and sequencing of the bacterial genes on chromosomal or plasmid DNA. Many of them are detailed below.

1.5.2.1 Plasmid typing

Plasmid typing was one of the first genotypic methods to be used for epidemiological typing and entails the isolation of plasmids from the strain and determining the number and sizes via agarose gel electrophoresis (Maslow *et al.*, 1993). Although the technique is easy to perform and is moderately priced (Table 1.3, page 32), plasmid typing can vary in typability and discriminatory power, according to the presence and number of plasmids and the bacteria studied (Maslow and Mulligan, 1996). Reproducibility can also be poor, as supercoiled and open circular forms of the same plasmid can migrate differently during electrophoresis (Maslow and Mulligan, 1996), and due to their extrachromosomal nature, they can be lost or acquired spontaneously (Maslow *et al.*, 1993).

Table 1.3. Characteristics of genotypic bacterial typing systems

Typing method[1]	Typability	Reproducibility	Discriminatory power	Ease of use	Ease of interpretation	Cost
Plasmid typing	Variable[2]	***	Variable	***	****	***
Genomic REA	****	Variable	Variable	**	***	***
Ribotyping/ IS-RFLP	****	*****	****	**	****	****
PFGE	****	*****	*****	**	****	****
AP-PCR, RAPD, rep-PCR	****	***	*****	**	***	***
AFLP	****	****	*****	**	***	****
DNA sequencing	*****	*****	*****	*	****	*****

Adapted from Maslow and Mulligan (1996) and van Belkum *et al.* (2001). Asterix scoring ranges from 1=low to 5=high.

[1]REA, restriction endonuclease analysis; IS-RFLP, insertion sequence restriction fragment length polymorphism; PFGE, pulsed-field gel electrophoresis; AP-PCR, arbitrarily primed-polymerase chain reaction; RAPD, randomly amplified polymorphic DNA; rep-PCR, inter-repeat element PCR; AFLP, amplified fragment length polymorphism.
[2]Scoring of method is dependent on the organism tested.

1.5.2.2 Restriction endonuclease analysis

Restriction endonuclease analysis (REA), a restriction fragment length polymorphism (RFLP) method, uses enzymes that will cleave the chromosomal and/or the plasmid DNA (if present) at specific sites, and potentially this can generate several hundred small fragments, which are then separated by conventional gel electrophoresis (Struelens *et al.*, 1996). Different strains of the same species can have different REA profiles due to the variations in their genomes and the subsequent changes in restriction sites (Maslow *et al.*, 1993). This simple technique can be used for any bacterial species and, after standardising conditions, such as enzyme selection, separation conditions and band size range, REA can be very reproducible and discriminatory (Struelens *et al.*, 1996). *Streptococcus pyogenes*, *Clostridium difficile* and *Legionella pneumophila* have been typed successfully using REA (Cleary *et al.*, 1988; Clabots *et al.*, 1992; Fry *et al.*, 1999), however, the interpretation of the complex profiles is difficult. Also, any plasmid DNA present may affect the pattern generated, so that differences may be influenced by the plasmid and not due to any genomic variation (Maslow *et al.*, 1993).

1.5.2.3 Insertion sequence- restriction fragment length polymorphism analysis and ribotyping

Another RFLP approach is to transfer the REA profiles from the gel to a membrane, otherwise known as Southern blotting (Southern, 1975). Nucleic acid probes designed to target genes, insertion sequence (IS) elements or rRNA genes can then be hybridised to specific fragments which contain a homologous sequence to the probe (Maslow *et al.*, 1993). Insertion sequences and transposons can be found in various bacterial chromosomes and may be used as probes. Generally, all IS elements are under 2.5 Kb and can insert multiple copies in a bacterial genome affecting the expression of neighbouring genes, but are sufficiently stable to be applied to RFLP studies (Mahillon and Chandler, 1998). One such established method is the IS*6110*-based DNA fingerprinting for *Mycobacterium tuberculosis*, and recommendations for standardisation are published (van Embden *et al.*, 1993). A number of tuberculosis investigations, including confirmation of an outbreak in institutional settings, and tracking geographical spread of *M. tuberculosis* clones have used IS*6110*-RFLP (Foxman and Riley, 2001). Ribotyping uses the ribosomal gene, *rrn* as the probe and, as all bacteria contain this gene, the typability is high (Table 1.3, page 32) (Maslow *et al.*, 1993). These Southern blot approaches are however, relatively laborious and time-consuming (Struelens *et al.*, 1996).

1.5.2.4 Pulsed-field gel electrophoresis

Pulsed-field gel electrophoresis (PFGE) is similar to REA, but generally uses rare-cutting restriction enzymes to produce less than 30 large genomic fragments, which are between 10 Kb to 800 Kb in size (Maslow *et al.*, 1993). The conditions for separating these fragments differ from conventional electrophoresis, in that the orientation of the electric field is pulsed periodically (Maslow and Mulligan, 1996). Field inversion gel electrophoresis (FIGE), is a simple form of PFGE, and uses forward and reverse current at 180°, although another form, named counter-clamped homogenous electric field (CHEF), has electrodes in a hexagonal array which are capable of generating an unlimited variety of field strengths and angles to further separate similarly-sized fragments (Sambrook and Russell, 2001). This method is technically demanding, but can be used on all bacteria, and can be highly discriminatory and reproducible (Table 1.3, page 32) (Struelens *et al.*, 1996). Like the Southern blot methods, PFGE is also a lengthy process, typically taking two to six days to obtain results, and the initial outlay of equipment is expensive (Maslow and Mulligan, 1996). There has also been a lack of inter-laboratory standardisation of PFGE, which makes comparing data difficult between laboratories (Struelens *et al.*, 1996). Despite these drawbacks however, many bacterial pathogens, including *Escherichia coli*, *Staphylococcus aureus*, and *Mycobacterium avium* have been epidemiologically-typed successfully using PFGE (Arbeit *et al.*, 1990; Prevost *et al.*, 1992; Arbeit *et al.*, 1993).

1.5.2.5 PCR amplification methods

A further genotypic approach exploits DNA amplification with PCR, using either; (i) arbitrary primers and/or conditions of low stringency, known as arbitrarily primed-PCR (AP-PCR) (Welsh and McClelland, 1990) or randomly amplified polymorphic DNA (RAPD) (Williams *et al.*, 1990); or (ii) primers directed outward from repetitive elements to amplify short spacer sequences, namely inter-repeat element PCR (rep-PCR) (Versalovic *et al.*, 1991). By separating the resulting different-sized amplicons using gel electrophoresis, a DNA fingerprint can be visualised and, due to the speed of PCR, a result can be generated within a day (Struelens *et al.*, 1996). The methods AP-PCR and RAPD can be used for a variety of bacteria, although identifying suitable primers and conditions that give good discrimination for a particular organism may prove difficult (Maslow *et al.*, 1993). Inter-laboratory reproducibility has also been demonstrated to be problematic (Table 1.3, page 32), and is documented by van Belkum and colleagues (1995) for *Staphylococcus aureus*. In their study, seven laboratories typing 59 isolates by AP-PCR generated 16 to 30 different genetic types (van Belkum *et al.*, 1995). The third approach, rep-PCR, uses various

repetitive elements, but these are mostly organism-specific, and therefore this technique may not be universally applicable to all bacteria. These elements include repetitive extragenic palindromes (REP), found in *E. coli* and *Salmonella typhimurium* (Gilson *et al.*, 1984); enterobacterial repetitive intergenic consensus (ERIC) sequences, present in many enterobacteria (Hulton *et al.*, 1991); and the BOX elements found in *Streptococcus pneumoniae* (Martin *et al.*, 1992).

1.5.2.6 Amplified fragment length polymorphism analysis

The PCR methods described above can be processed further using restriction enzymes, which may increase the discrimination of the technique (Maslow and Mulligan, 1996). One method, amplified fragment length polymorphism (AFLP) analysis, is essentially a three step procedure; an initial enzyme digest-ligation of genomic DNA, then the selective amplification of the restriction fragments using specific primers, and finally the gel analysis of the amplified fragments (Vos *et al.*, 1995). The AFLP technique thus combines the reliability of the RFLP technique with the power of PCR to produce a reproducible, discriminatory, rigorous, and quick method (Table 1.3, page 32) (Vos *et al.*, 1995). For most bacteria, a double enzyme digest is performed for AFLP (Janssen *et al.*, 1996). However, single-enzyme AFLP has been applied to some bacteria, including *L. pneumophila*, and AFLP has been chosen by the European Working Group on Legionella Infections (EWGLI) to standardise and use as an epidemiological typing scheme across 11 European countries (Fry *et al.*, 2000). The AFLP method has been further developed by amplifying restriction fragments using fluorescently-labelled primers, (fluorescent-AFLP) and utilises automated DNA sequencers to determine the size of the labelled amplified fragments. Fluorescent-AFLP also permits the inclusion of internal size standards and has been shown to be accurate, discriminatory, reproducible and may be capable of standardisation for *E. coli* (Arnold *et al.*, 1999). There are however some drawbacks in this and other PCR typing methods, which include: optimisation and choice of reagents/protocols; inter-run and inter-laboratory reproducibility; and interpretation of results, making these methods difficult to standardise (Struelens *et al.*, 1996).

1.5.2.7 DNA sequencing

Another genotypic approach to type bacteria is through gene variation, detected by sequencing the gene investigated. Sequencing of 16S rDNA is well established for phylogenetic classification, but the relative stability of the 16S rRNA gene makes it unsuitable for typing below the species level, so other, more variable loci are used. Multi-

locus sequence typing (MLST) was developed from the phenotypic technique, multi-locus enzyme electrophoresis (MLEE) (Maiden *et al.*, 1998). Briefly, MLEE involves the preparation of between 15 to 25 enzyme extracts, and their subsequent electrophoresis using horizontal starch gels. The relative mobility of each enzyme locus varies in different strains of a particular species, and can be given a specific allele number, which correlates with amino acid substitutions (Maslow *et al.*, 1993). When these multiple loci are combined as an electrophoretic type (ET), a potentially large number of different ETs can be generated for a particular organism. Due to its moderate discriminatory value, and its slow and technically demanding nature, MLEE is used more for population genetics than epidemiological typing (Maslow *et al.*, 1993), and has also been used to estimate genetic diversity in several bacteria (Selander *et al.*, 1986). The MLST scheme requires the DNA sequencing of some of the genes encoding the enzymes used in MLEE, where each variant of the corresponding genes are given an allele number, and when combined, will produce an allelic profile or sequence type (ST) for each strain tested. Schemes for *Neisseria meningitidis*, *Streptococcus pneumoniae*, and *Staphylococcus aureus* have been published to study global epidemiology, and have reported on the population structure, serotype correlation, and important MRSA clones (Maiden *et al.*, 1998; Enright and Spratt, 1998; Enright *et al.*, 2000), respectively. Multi-locus sequence typing, combining optimal typability, excellent discrimination, and a simple, portable way of comparing data between laboratories, however can be labour intensive, and the initial expenditure and further costs are very high (Table 1.3, page 32) (Maiden *et al.*, 1998). Nevertheless, MLST has been successful in typing meningococci directly from clinical specimens where the isolate has not been recovered, therefore generating a typing result for enhancing surveillance (Diggle *et al.*, 2003). Virulence genes can also be targeted for typing, such as *porA* and *porB* in *N. meningitidis*, which encode the serotyping antigen (Feavers *et al.*, 1999). This single locus approach can enable the study of virulence factors under environmental pressure, but these genes can be highly variable and may obscure true relationships between species (Gurtler and Mayall, 2001). Nevertheless, this approach may be necessary for species with limited genetic diversity, such as *B. pertussis*, which is described later (1.6, page 39).

1.5.3 Genetic events in bacteria visualised using genotypic methods

When using genotypic methods, the variations observed are caused by essentially four evolutionary genetic mechanisms: mutation, hypermutation, genetic recombination and natural selection (van Belkum *et al.*, 2001). Mutations are thought to be random events and are heritable via vertical transmission. They are caused either by DNA replication errors

such as base pair substitutions, insertions and deletions, or by the movement of transposons and IS elements within the genome. Hypermutation is probably caused by a lack of DNA mismatch repair, which can enhance the "mutability" of a cell and the resulting organisms are called "mutators" (van Belkum *et al.*, 2001). Rapid mutation and a high rate of DNA exchange are thought to be advantageous for mutators to adapt to a new environment. Slipped-strand mispairing can occur when these hypermutable genes contain short sequence repeats, such as those present in IS elements. Thus, the IS elements can control cassette switching and affect the expression of genes, which can then be identified as a different phenotypic trait (van Belkum *et al.*, 2001).

Genetic recombination and horizontal gene transfer can include the acquisition of genes into genomes and make the estimates of clonal evolution difficult to calculate (van Belkum *et al.*, 2001). The rate of recombination influences the clonality of the organism, where bacteria with low recombination are highly clonal and stable (e.g., *Salmonella enterica*) and those with high recombination are non-clonal and panmictic (where gene exchange occurs randomly and at a high rate in a population) (e.g., *N. meningitidis*) (Maynard Smith *et al.*, 1993). Researchers have thus found difficulties when deciphering an ancestral lineage for panmictic species such as *Helicobacter pylori*, due to their high genetic variability.

Natural selection has an overall role of filtering out any new genotypes produced by mutations or recombination that have detrimental fitness effects. Two variables affect natural selection: the fitness of the new genotype (whether it is neutral, detrimental or beneficial), and the population size. Understandably, variant genotypes are more likely to be present in larger populations, due to the mutation and efficient selection of the organism in question.

These genetic events, detailed above, have been interpreted for PFGE as changes in band patterns (Tenover *et al.*, 1995). As PFGE patterns rely on the presence of specific sites at which a restriction endonuclease will cleave the genome, mutations and recombinations affecting these sites can be visualised as changes. This interpretation for PFGE is based on the number of fragment differences between the profiles, to which the authors have defined particular genetic events (Tenover *et al.*, 1995). As typing methods are used for comparing relatedness, in an outbreak setting these PFGE interpretations can help in characterising strains. Tenover *et al.* (1995) state that if there are two to three fragment differences between the outbreak pattern and that produced from the test isolate, this corresponds to either a point mutation or an insertion or deletion, and is therefore closely

related and probably part of the outbreak. Two genetic events in an isolate generating four to six fragment differences are interpreted as possibly part of the outbreak. Unrelated isolates that are not part of the outbreak are indicated by seven or more fragment differences, where there have been three or more genetic events. This is based on observations of isolates collected over a six-month period or from large, extended outbreaks (Tenover *et al.*, 1995).

Tenover *et al.* (1995) do not recommend their criteria for studies of large populations or organisms that have been collected over periods of one year or longer, as they are too stringent, yet these recommendations have been applied universally for a number of bacteria. When evaluating a particular method therefore, its appropriateness, the genetic variation it reveals, and its interpretation are major factors to consider, as well as the clonality of the organism investigated.

1.5.4 The "ideal" bacterial typing system

By applying the recommended criteria to a particular typing system (1.5.1, page 30), the appropriateness of a method can be evaluated for each species studied. An ideal typing system should possess;

> "... optimal typability, a high degree of reproducibility, adequate stability, and unprecedented resolving power ...".

(van Belkum *et al.*, 2001)

The system chosen must also fit the purpose of the study, whether for genetic diversity, population dynamics or epidemiological typing. Standardisation of bacterial typing has been attempted for *E. coli* O157:H7, *M. tuberculosis*, and *L. pneumophila*, using PFGE, IS*6110*-RFLP and AFLP, respectively, and has enabled working groups in separate laboratories to compare results (Swaminathan *et al.*, 2001; van Embden *et al.*, 1993; Fry *et al.*, 2000).

Many typing methods have been applied to *B. pertussis* for studies of genotypic diversity, population dynamics and epidemiological typing, and have revealed seemingly contradictory evidence concerning the stability of the *B. pertussis* genome. However, to date, standardisation of these methods has not been established.

1.6 Genetic diversity and epidemiological typing of *Bordetella pertussis*

1.6.1 Genome plasticity of *Bordetella pertussis*

In order to study evolution, diversity and epidemiology of *B. pertussis*, determining the stability or plasticity of the genome helps in the choice of typing method and interpretation of the data. Genome plasticity is the accumulation of changes in a genome, based on molecular processes such as transposition (transposable elements, insertion sequences, or transposons, which can translocate to another site of the chromosome), transformation (the transfer of DNA by a bacterium from the surrounding environment), or mutation (van Belkum *et al.*, 2001).

Stibitz and Yang (1999) characterised 14 *B. pertussis* strains representative of 14 different PFGE patterns obtained from 70 strains isolated from a whooping cough outbreak in Alberta, Canada from December 1989 to May 1991. Using a 'chromosomal surveying' approach, they measured the distances between 19 *B. pertussis* genes within each genome, using vectors, the intron-encoded restriction enzyme I-*Sce*I, and PFGE (using *Xba*I) to size the fragments. Ten chromosome maps were deduced from the 14 strains and these revealed large chromosomal inversions. The molecular mechanism for these inversions is unknown, but Stibitz and Yang suggested that insertion elements could catalyse such an inversion in the *B. pertussis* chromosome (Stibitz and Yang, 1999). They also proposed that changes in PFGE patterns could be affected by these large chromosomal rearrangements, and not solely by the loss or gain of restriction sites as described by Tenover *et al.* (1995). This implies a dynamic *B. pertussis* population structure and genetic diversity studies are detailed below.

1.6.2 Genotypic diversity of *Bordetella pertussis*

Musser *et al.* (1986) studied genetic diversity in *Bordetella* species using MLEE, and included 23 strains of *B. pertussis*. They deduced that the genetic diversity of *B. pertussis* is lower than that of other bacteria tested, such as *H. influenzae*, *L. pneumophila* and *E. coli* (Musser *et al.*, 1986), with only 3 ETs identified using 15 enzymes. This could be due to; pertussis vaccination, which has reduced the population size; the niche specialisation of *B. pertussis*, whose host is exclusively human; or a recent evolutionary origin, as the earliest recorded description of whooping cough was in 1578 (Musser *et al.*, 1986). They also reported that the type strain of *B. pertussis*, 18323^T (NCTC 10739^T), used in mouse potency tests for

WCVs, was more genetically similar to *B. bronchiseptica* and *B. parapertussis* than *B. pertussis*, supporting the view that the type strain is an unusual strain of *B. pertussis* (Musser *et al.*, 1986).

This work was extended later in an analysis of 18 *B. pertussis* strains (188 *Bordetella* strains in total) by MLEE, and the presence of any of the three IS elements, IS*481*, IS*1001*, and IS*1002* (Table 1.4, page 41) (van der Zee *et al.*, 1997). This study revealed only four different ETs, resulting from only four of the 15 enzymes, were polymorphic for the *B. pertussis* isolates. These authors deduced that *B. pertussis* strains cluster separately from *B. parapertussis* and *B. bronchiseptica* due to their ETs and the presence of IS*481* and IS*1002*.

Khattak *et al.* (1992) tested 105 *B. pertussis* isolates from the UK and Germany, including circulating strains, vaccine strains and control isolates, by PFGE. They found 17 PFGE types in total, with one type accounting for 25 of the 105 isolates tested. The authors concluded that *B. pertussis* is a highly genetically variable organism, but no statistical calculations were presented. The work of Stibitz and Yang (1999), showing genomic plasticity in *B. pertussis*, supported this conclusion, but conflicted with the results of the MLEE studies of Musser *et al.* (1986) and van der Zee *et al.* (1997).

IS*1002*-restriction fragment length polymorphism (IS*1002*-RFLP) analysis, a method developed for *B. pertussis* by van der Zee *et al.* (1996b), demonstrates a sufficiently high degree of variation in the profiles to allow typing of *B. pertussis*. Van der Zee *et al.* (1996b) studied 106 *B. pertussis* isolates from The Netherlands, Germany, Canada, the USA, China, Japan and the UK. The resulting profiles were sorted into 36 different RFLP types, which could be clustered into five clades (A to E) of related profiles. Results indicated a partial geographic isolation for some of the RFLP types. In particular, nine of the ten isolates from Japan were clustered into one clade (clade E) that contained only two other isolates, both from Canada (van der Zee *et al.*, 1996b). Also, 15 isolates from Canada which were previously resolved as 15 different PFGE types (de Moissac *et al.*, 1994), were differentiated into 12 IS*1002*-RFLP types, and the authors suggested that the discrimination of IS*1002*-RFLP, if a little lower, is comparable to PFGE for *B. pertussis* (van der Zee *et al.*, 1996b). One particular RFLP type, named D1, was present in five of the six countries analysed, indicating that this is a successful clone (group of isolates sharing the same RFLP type) that has spread globally.

Table 1.4. Insertion elements present in *Bordetella*

Insertion element	Present (number of copies in genome sequenced[1])			
	B. pertussis Tohama-I	*B. parapertussis* 12822	*B. bronchiseptica* RB50	*B. holmesii*
IS*481*	Yes (238)	No	Yes[2] (0)	Yes[3]
IS*1001*	No	Yes (22)	Yes[2] (0)	NK
IS*1002*	Yes (6)	Yes (9)	Yes[2] (0)	NK
IS*1663*[4]	Yes (17)	No	No	NK

[1]Numbers in parentheses indicate the number of copies of the particular insertion sequences in the genomes sequenced (not *B. holmesii*) at the Sanger Institute (Parkhill *et al.*, 2003). [2]The *B. bronchiseptica* strain chosen to be sequenced by the Sanger Institute contained no copies of the stated insertion elements, however, other reports suggest there are copies present in other strains (van der Zee *et al.*, 1996a; Gladbach *et al.*, 2002). [3]IS*481* has been detected in *B. holmesii* using PCR-based methods (Reischl *et al.*, 2001; Sloan *et al.*, 2002). [4]A previously unknown insertion element in *B. pertussis*, discovered after genomic sequencing of the Tohama-I strain at the Sanger Institute (Parkhill *et al.*, 2003).

NK=not known.

The authors also suggested that, after analysing isolates from The Netherlands, IS*1002*-RFLP analysis could be used in an epidemiological setting, as two clusters containing 3 and 4 strains revealed two RFLP patterns, type D1 and C2, respectively (van der Zee *et al.*, 1996b). Strains from an epidemic in 1994 revealed that 15/21 (71%) of them belonged to two RFLP types, B3 or C5, and suggested that clonal expansion cause pertussis epidemics. The RFLP type D1, found in several countries, was also present in all four periods between 1954 and 1994, indicating this RFLP type is persistent through time as well as space (van der Zee *et al.*, 1996b).

The two strains in the Dutch WCV exhibited RFLP types A1 and D2. These RFLP types were only found a few times between 1950 and 1954 and not in later periods. The authors hypothesise that;

> "… over 40 years of vaccination has resulted in the expansion of strains which are less affected by vaccine-induced immunity due to antigenic differences with the vaccine strains …".
>
> (van der Zee *et al.*, 1996b)

Another explanation is that the differences reflect the normal dynamics of *B. pertussis* populations.

Further work from The Netherlands using IS*1002*-RFLP analysis also supported the hypothesis of *B. pertussis* adaptation due to vaccine-induced immunity (van Loo *et al.*, 1999). A total of 213 Dutch isolates from 1949 to 1996 were collected and grouped into five time periods, and 45 RFLP types were observed. Prior to vaccination (1949 to 1954), there were 12 RFLP types present in the *B. pertussis* population, including the two vaccine strains. Between 1965 and 1972, the two vaccine RFLP types disappeared, and only five RFLP types were observed. From 1965 onwards, three RFLP types dominated the *B. pertussis* population (DNA types dt-12, dt-29, and dt-35). Furthermore, one of these types, dt-12, was found in all five time periods from 1949 to 1996. A smaller study in Italy, analysed 76 *B. pertussis* strains from an Italian vaccine field trial which started in 1993, and found similar results in that dt-12, dt-29, and dt-35 dominated in 75% of the strains tested (Mastrantonio *et al.*, 1999).

Genotypic diversity (GD) values were also calculated for each time period of these Dutch data (van Loo *et al.*, 1999), based on DNA polymorphisms detected by restriction endonuclease analysis and assuming that evolution mainly occurs by mutation and genetic drift (Nei and Tajima, 1981). These values are therefore dependent on the bacterial

population size and frequency of each RFLP type, with values approaching one being of a high genotypic diversity, and values approaching zero having a low genotypic diversity. The highest GD of 0.89 was observed in 1949 to 1954, and the lowest at 0.60 was seen in the next time period studied, 1965 to 1972 (van Loo *et al.*, 1999). This significant drop in genotypic diversity is thought by the authors to have been caused by a decrease in the bacterial population size and/or by clonal expansion of "more fit" strains after vaccination was introduced (van Loo *et al.*, 1999). This clonal expansion theory was further supported by the appearance of dt-29 in 11/18 (61%) of the isolates from 1965 to 1972 and the persistence of this RFLP type in subsequent years. The genotypic diversity then increased, suggesting a gradual expansion of other RFLP types. Most of the strains 120/213 (56%) were collected between 1991 and 1996 possibly skewing the results. However, the authors suggest that;

> "... vaccination has selected for strains that are adapted to a highly vaccinated population ...",
>
> (van Loo *et al.*, 1999)

and that this may be the cause of pertussis re-emergence in other countries.

Pulsed-field gel electrophoresis and IS*1002*-RFLP analysis are more preferable methods of studying genotypic diversity than the phenotypic technique, MLEE. Data from MLEE analysis show *B. pertussis* as a clonal species within a highly-related genus (Musser *et al.*, 1986). However, results from PFGE and IS*1002*-RFLP analysis suggest a more dynamic species (Khattak *et al.*, 1992; van Loo *et al.*, 1999). These observation using PFGE and IS*1002*-RFLP analysis are supported by the demonstraion of the genomic plasticity of *B. pertussis* (1.6.1, page 39) (Stibitz and Yang, 1999).

1.6.3 Epidemiological typing of *Bordetella pertussis*

Serotyping of *B. pertussis* using the fimbriae/agglutinogens has been used for many decades for epidemiological investigations of pertussis (Eldering *et al.*, 1957). In Sweden, clinical isolates collected between 1970 and 1995 were investigated by serotyping, and it was found that serotype 1,2,3 isolates predominated from 1970 to 1978, with serotype 1,3 being related to outbreaks of pertussis (Tiru *et al.*, 1997). However, following the cessation of general pertussis vaccination in 1979 (Therre and Baron, 2000), serotype distribution changed, so that in 1995, serotype 1,2 was present in over 80% of isolates (Tiru *et al.*, 1997). In the UK, during the two large pertussis epidemics of 1977 to 1979 and 1981 to 1983, serotyping data and vaccination status were collected for 855 cases (Preston and Carter,

1992). This study reported that serotype 1,3, as in Sweden, predominated in both epidemics, and serotypes 1,2 and 1,2,3 were found in 243/769 (32%) of unvaccinated children compared to only 9/86 (10%) in vaccinated subjects, which was statistically significant. These authors also stated that WCVs had a higher efficacy against serotypes 1,2 and 1,2,3 than against serotype 1,3 (Preston and Carter, 1992). Therefore, it was advised that ACVs should contain agglutinogens for each serotype to ensure adequate efficacy and protection against pertussis disease (Preston and Carter, 1992).

Due to the low discrimination of serotyping, it is not suitable for short-term outbreak investigations and PFGE has been used for more recent epidemiology studies. An outbreak in Alberta, Canada, from December 1989 to May 1991, initiated a study to determine whether the movement of virulent *B. pertussis* strains within and/or between towns and communities caused the outbreak (de Moissac *et al.*, 1994). A total of 70 strains were typed by PFGE using *Xba*I, and various epidemiological grouping analyses were performed. Fourteen isolates from the city of Edmonton and one from each of 14 northern Alberta towns (n=28) generated 13 PFGE profiles. A highly heterogeneous population was initially observed in Edmonton, thus revealing that PFGE could discriminate amongst known epidemiologically-linked isolates, which could further complicate the analysis of prevalence and spread of the isolates tested. A further six pairs of intrafamilial isolates (n=12) were tested. Each isolate from each pair shared identical PFGE types, suggesting direct transmission within families was occurring in the outbreak. Isolates from Fort Smith (n=18), a town north of Edmonton, all exhibited an identical PFGE type, type a, indicating the outbreak in Fort Smith was due to one strain. Another town shared another PFGE type, type b, in 11 out of 12 isolates, and both types a and b accounted for 44 of the 70 isolates. The authors concluded that a single, highly infectious strain of *B. pertussis* did not cause the whooping cough outbreak and by studying the vaccination status of the population, they suggested that vaccination was less protective in periods of peak pertussis activity (de Moissac *et al.*, 1994).

Another rise in whooping cough cases occurred in 1993 in the UK and was investigated by PFGE and serotyping of isolates received by the Pertussis Reference Laboratory in Manchester (Syedabubakar *et al.*, 1995). Of the 4000 notifications of pertussis in this peak year, 280 *B. pertussis* isolates were received at their laboratory, of which 180 were typed and 29 different PFGE types produced. DNA type 1 predominated in 72/180 (40%) of the isolates tested, and DNA type 3 encompassed 42/180 (23%). There were 11 unique DNA types containing one isolate each. Studying the temporal distribution, 17 DNA types

appeared in the upswing phase of the epidemic compared to just 8 DNA types in the downswing phase, suggesting greater genetic plasticity in the upswing phase (Syedabubakar *et al.*, 1995).

In a separate study, no significant correlation was found between PFGE type and serotype. DNA type 1 occurred in strains from nine fully vaccinated children, seven of which were serotype 1,2. DNA type 2 accounted for 10/67 (15%) of isolates from September 1990 to March 1991 (Khattak *et al.*, 1992), but was markedly absent a couple of years later in 1993. Comparison of the data from Khattak *et al.* (1992), with that of Syedabubakar *et al.* (1995), revealed that only eight of the 15 DNA types were seen in both 1990 to 1991 and 1993, and 21 new types were identified from the 1993 isolates. Syedabubakar and colleagues (1995) proposed that the 1993 pertussis epidemic was caused, in part, by the emergence of genetic variants after the "failure" or mis-match of the WCV, though they did not report the DNA types of the vaccine strains. Nevertheless, this theory has been supported by studies in The Netherlands (van der Zee *et al.*, 1996b; Mooi *et al.*, 1998; van Loo *et al.*, 1999).

The studies in Canada and the UK mentioned above, reported that pertussis epidemics may not be caused by a single, highly virulent strain (de Moissac *et al.*, 1994; Syedabubakar *et al.*, 1995). In contrast, Beall *et al.* (1995) analysed 78 isolates from the USA and Senegal, and reported indistinguishable PFGE profiles from epidemic strains. The authors used three different enzymes (*Xba*I, *Spe*I, and *Dra*I), and 16 PFGE profiles were generated. All of the 54 isolates from an outbreak in Delaware in 1986 were of one PFGE type. Only one of the 25 sporadic isolates matched the epidemic PFGE type, but clinical data revealed that this isolate was cultured from a laboratory worker who had been working on the epidemic strains. The authors therefore concluded that these data;

> "… justified an association between a single PFGE type and epidemic pertussis …".
>
> (Beall *et al.*, 1995)

They also explained that although the 1989 to 1991 epidemic in Canada was province-wide and lasted 18 months, the three month outbreak which occurred in Fort Smith generated only one PFGE type (de Moissac *et al.*, 1994), which is more comparable to the smaller, three month situation in Delaware. Beall and colleagues concluded that PFGE of *B. pertussis* could enable the;

> "… characterisation of endemic and epidemic patterns, secondary attack rates, vaccine efficacy, and the role of atypical illness in transmission …".

(Beall *et al.*, 1995)

In order to correctly identify virulent strains or clones that cause severe disease or outbreaks, the collection of clinical and epidemiological data are important and, combined with the equivalent typing data, associations can be tested. Clinical data, such as isolation date, age of subject, vaccination status, and geographical origin are essential in deducing whether particular isolates are indeed from the same strain. Unfortunately, the standardisation of both the methodology and nomenclature of PFGE for *B. pertussis* has not been established and is necessary if direct comparison is required. It would be useful to know whether there are any prevalent strains causing most of the pertussis cases, or more severe disease on a global scale, such as the successful IS*1002*-RFLP type D1 (van der Zee *et al.*, 1996b).

1.6.4 Gene sequence variation of *Bordetella pertussis*

Following the pertussis epidemic in The Netherlands in 1996, and the *B. pertussis* population study using IS*1002*-RFLP analysis by van der Zee *et al.* (1996b), Mooi *et al.* (1998) investigated whether this resurgence was due to the emergence of different *B. pertussis* genetic variants to that of the WCV strains. Mooi *et al.* (1998) chose two *B. pertussis* virulence factors, the adhesin pertactin (1.2.3.3, page 15) and pertussis toxin (1.2.2.1, page 8), and sequenced the genes for these factors from isolates obtained before and after vaccination was introduced, to investigate whether any antigenic variation may have occurred in the Dutch *B. pertussis* population. After sequencing approximately 90% of the pertactin (*prnA*[2]) gene of six *B. pertussis* isolates, two repeat regions encoding the amino acids GGxxP and PQP in regions 1 and 2, respectively, were found. Polymorphism was found in region 1, which differed in the number of repeats present, and/or the amino acid composition. Three *prnA* types were revealed in the isolates from 1949 to 1996, namely *prnA*(1), *prnA*(2), and *prnA*(3) (Figure 1.2, page 16) (Mooi *et al.*, 1998). When sorted by year of isolation, 35 isolates between 1949 and 1980 were all *prnA*(1), which was also found in the two strains used to produce the Dutch WCV. However, since 1981, types *prnA*(2) and *prnA*(3) have appeared and percentages of both these types has gradually risen temporally.

---

[2] As the genetic nomenclature for pertactin and pertussis toxin used by Mooi and colleagues has varied from publication to publication, and is nonstandard according to the Nomenclature Committee of the American Society for Microbiology (ASM) Publications Board, the following will be used throughout the thesis: P.69A=*prn1*, P.69B=*prn2*, P.69C=*prn3*, etc, will become *prnA*(1), *prnA*(2), *prnA*(3), etc, and S1A=*ptxS1*, S1B=*ptxS2*, S1D=*ptxS4*, etc, will become *ptxA*(1), *ptxA*(2), *ptxA*(3), etc.

Since 1993 these two novel types have comprised approximately 90% (123/133 isolates) of the Dutch *B. pertussis* population.

The pertactin sequence data from isolates between 1989 and 1996, were also analysed using the age groups of the pertussis cases, to determine whether the vaccination status of an individual was associated with a particular pertactin type. Mooi and colleagues (1998) thus assumed that at 0 to 3, 4 to 11, 12 to 48, and over 48 months, children had no, partial, optimal, and waning immunity, respectively. The authors found that the percentage of the vaccine type *prnA*(1) revealed a downward trend in the no, partial and optimal immunity groups (from 22% to 8%) (Mooi *et al.*, 1998). This showed that infection by the non-vaccine types, *prnA*(2) and *prnA*(3), was more likely in vaccinated subjects, and the authors stated that;

> "… vaccine-induced immunity against *prnA*(1) strains is stronger than those against *prnA*(2) and *prnA*(3) …".
>
> (Mooi *et al.*, 1998)

No information regarding the vaccination status of the individuals were available, so confidence in the immunity levels cannot be determined, although from 1993 to 1996, 85% of pertussis patients older than 12 months were vaccinated (de Melker *et al.*, 1997).

Polymorphism was also found in the gene encoding the S1 subunit of pertussis toxin (*ptxA*) for 49 *B. pertussis* isolates (Mooi *et al.*, 1998), as single nucleotide polymorphisms (SNPs) in two regions that were previously identified as T cell epitopes (Peppoloni *et al.*, 1995). Three *ptxA* types were observed, *ptxA*(1), *ptxA*(2), and *ptxA*(3), (Figure 1.4, page 48). When grouped into three time periods, isolates from the pre-vaccination era, 1949 to 1954, either contained *ptxA*(2) or *ptxA*(3), which were found in the two strains of the Dutch WCV. In 1978 to 1985, a novel *ptxA* type, *ptxA*(1) appeared and was present in 12/15 (80%) of the isolates. This predominant *ptxA* type increased further, to 15/17 (88%) of isolates, between 1990 and 1996.

Mooi *et al.* (1998) concluded that these new pertactin and pertussis toxin types not found in the Dutch WCV appeared in the *B. pertussis* population through time due to vaccine-driven evolution. They also suggested that this antigenic shift may be contributing to the resurgence of pertussis in other countries with high WCV coverage (1.4.3, page 26), and that further work was needed to study the population structure of *B. pertussis* in other countries (Mooi *et al.*, 1998).

Figure 1.4. Alignment of the pertussis toxin subunit 1 (*ptxA*) gene of *Bordetella pertussis*, showing the polymorphic regions

```
GenBank
accession   ptxA
number      type      106                                  196                                  247                        580
                       |                                    |                                    |                          |
AJ006155    ptxA(1)   GAT CCT CCC GCC ACC -//- GTG CTC GAC CAT CTG -//- GCT TTC GTC TCC ACC -//- GAG TAT TCC AAC GCT
                       D   P   P   A   T        V   L   D   H   L        A   F   V   S   T        E   Y   S   N   A
AJ006157    ptxA(2)   ... ... ... ... ... -//- ... ... ... ... ... -//- ... ... ... ... ... -//- ... ... ... ... ...

AJ245368    ptxA(3)   ... ... ... ... ... -//- ... ... GAA ... ... -//- ... ... ... ... ... -//- ... ... ... ... ...
                                                         E
AJ006159    ptxA(4)   ... ... ... ... ... -//- ... ... GAA ... ... -//- ... ... ... ... ... -//- ... ... CCC ... ...
                                                         E                                                P
AJ506994    ptxA(5)   ... ... CCA ... ... -//- ... ... ... ... ... -//- ... ... ... ... ... -//- ... ... ... ... ...
                               P
AJ506995    ptxA(6)   ... ... ... ... ... -//- ... ... ... ... ... -//- ... ... GTT ... ... -//- ... ... ... ... ...
                                                                                  V

                                 676
                                  |
AJ006155    ptxA(1)   -//- GTC CGC ATA CCG CCG GTG ATA GGC
                            V   R   I   A   P   V   I   G
AJ006157    ptxA(2)   -//- ... ... ATG ... ... ... ... ...
                                    M
AJ245368    ptxA(3)   -//- ... ... ATG ... ... ... GTG ...
                                    M               V
AJ006159    ptxA(4)   -//- ... ... ATG ... ... ... ATG ...
                                    M               M
AJ506994    ptxA(5)   -//- ... ... ... ... ... ... ... ...

AJ506995    ptxA(6)   -//- ... ... ... ... ... ... ... ...
```

This figure shows six pertussis toxin types. Only *ptxA*(1) through to *ptxA*(3) were described before work on this thesis began. Non-silent mutations are shown in bold and associated amino acid changes are indicated beneath the relevant codons. Yellow shaded amino acids indicate silent mutations. Dots indicate sequence identity to *ptxA*(1) and numbers indicate the position of nucleotides relative to the start codon of AJ006155. Please refer to Appendix 1 for amino acid abbreviations.

Following the studies in The Netherlands, Mooi collaborated with workers in Finland and Italy to determine the polymorphism of the genes coding for pertactin and pertussis toxin in isolates from these countries (Mooi *et al.*, 1999; Mastrantonio *et al.*, 1999). In Finland, a genetic shift was seen in *B. pertussis* isolates from 1953 to 1996. The Finnish WCV also contains strains possessing *prnA*(1), *ptxA*(2) and *ptxA*(3). Five isolates tested from 1953 to 1964 were *prnA*(1) and *ptxA*(2), but six isolates from 1982 demonstrated a shift to *prnA*(2) and *ptxA*(1). The 43 most recent Finnish isolates from 1990 to 1996 were all *ptxA*(1), but revealed four pertactin types, uncovering a new type, *prnA*(4), with only 3 (7%) of the isolates containing *prnA*(1), whilst *prnA*(2) predominated in 31 (72%) of the isolates. The numbers of isolates in this study in Finland are low however and, although antigenic shifts may be occurring, there have not been any major outbreaks or resurgence of pertussis in Finland compared to The Netherlands. The authors suggested that this lack of resurgence may be due to differing levels of protection by the WCV when challenged with pertactin variants. Furthermore, they stated that the frequency of *prnA*(3) in The Netherlands was higher than in Finland; 51%, compared to 5/43 (12%), and that the frequency of *prnA*(3) may have caused the 1996 epidemic in The Netherlands. Experiments using mouse challenge models may elucidate this matter (Mooi *et al.*, 1999).

The study in Italy focused on *B. pertussis* isolates from vaccinated and unvaccinated children to investigate any association between pertactin and pertussis toxin variants and vaccination status (Mastrantonio *et al.*, 1999). These isolates were collected during an Italian pertussis vaccine trial between 1993 and 1995 and, therefore, a temporal analysis over a larger time scale cannot be made due to the unavailability of isolates. In 129 isolates, four pertactin variants were found, *prnA*(1), (2), (3) and *prnA*(5), another novel type. The vaccines tested all contained *prnA*(1), and this type was present in only 8/129 (6%) of the isolates tested. The prevalent *prnA* types in Italy were of similar frequencies to those of the recent isolates from The Netherlands: *prnA*(2), 53/129 (41%) and *prnA*(3), 65/129 (51%) in Italy, compared to 36% and 51%, respectively. Interestingly, there was no significant association between the distribution of pertactin variants among the vaccinated and unvaccinated group. However, the authors stated that;

> "... isolates with *prnA*(1) are more affected by vaccine-induced immunity than isolates with *prnA*(2) or *prnA*(3) ...".
>
> (Mastrantonio *et al.*, 1999)

Only 30 Italian isolates were investigated for pertussis toxin gene polymorphism. All isolates tested were of type *ptxA*(1), reflecting similar recent distributions of *B. pertussis* in The Netherlands and Finland, and the vaccines used in the Italian trials also contained

either *ptxA*(2) and/or *ptxA*(3). The authors concluded that, although vaccine coverage was lower in Italy than in The Netherlands, the population structure of *B. pertussis* was similar (Mastrantonio *et al.*, 1999). Their findings did not support the theory of vaccine-driven evolution, but suggested that either the population structure may have been influenced by neighbouring countries with high vaccine coverage, or that even a relatively low vaccine coverage may have had a significant effect on the *B. pertussis* population (Mastrantonio *et al.*, 1999).

1.7 Aims of thesis

Bordetella pertussis research has had a recent revival after the resurgence of pertussis in many countries with high vaccine coverage. The Netherlands produced data suggesting a theory that circulating strains had emerged which were different to the vaccine strains and was thus causing a pertussis resurgence (Mooi *et al.*, 1998). Although there has been no apparent resurgence in the UK, determining the genotypic diversity and establishing epidemiological typing of *B. pertussis* isolates in the UK would be invaluable. Thus, an evaluation of the current *B. pertussis* situation will be made, together with the impact of vaccination policy and the resulting pertussis incidence. Direct comparisons with other countries such as The Netherlands and the USA will also be performed (Mooi *et al.*, 1998; van Loo *et al.*, 1999). Therefore the aims of this thesis are described below.

1.7.1 To characterise the range and distribution of sequence variation in pertactin and pertussis toxin genes of *Bordetella pertussis* isolates from the United Kingdom

Bordetella pertussis isolates will be obtained from archive collections from various sources in the UK, and recent isolates selected from those submitted to the Pertussis Reference Laboratory at Manchester University. These isolates will be investigated for pertactin and pertussis toxin S1 gene variation, and analysed for any temporal trends that have occurred since widespread vaccination against pertussis was introduced in the UK in 1957 (Chapter 3).

1.7.2 To investigate the genotypic diversity of a representative sample of historical, through to current, UK *Bordetella pertussis* isolates

The genotypic diversity of UK isolates will be calculated by selecting isolates from the initial data-set collected for aim 1.7.1 (see above) at random, and applying the IS*1002*-RFLP typing method of van Loo *et al.* (1999). Genotypic diversity values will then be directly compared with those from The Netherlands, and to elucidate whether;

> "… vaccination has selected for strains that are adapted to a highly vaccinated population …".
>
> (van Loo *et al.*, 1999)

This investigation is covered in Chapter 3.

1.7.3 To examine current *Bordetella pertussis* isolates for any association between clinical, epidemiological, and laboratory data

For isolates obtained from the Pertussis Reference Laboratory, Manchester, there are extensive epidemiological data stored at the Immunisation Division, CDSC, London. These include age, sex, vaccination history, severity and outcome of illness of the patients. Laboratory data such as serotyping, pertactin and pertussis toxin sequence types, and IS*1002*-RFLP types, will be merged with the epidemiological data and interrogated to identify any association between strain characteristics and disease patterns (Chapter 3).

1.7.4 To characterise *Bordetella pertussis* strains collected by the Missouri State Public Health Laboratory from Missouri, USA

A small number of *B. pertussis* strains from Kansas City, Missouri were collected when there was an increase in pertussis incidence, and these were tested alongside some earlier and contemporary strains for their relatedness. Phenotypic and genotypic methods to be used included; serotyping, *prnA* and *ptxA* gene sequencing, PFGE and IS*1002*-RFLP analysis. These data will then be discussed and compared to the UK data in Chapter 4.

1.7.5 To evaluate molecular typing techniques for their indices of discrimination by using a panel of *Bordetella pertussis* isolates

From the main set of *B. pertussis* isolates collected for this study, a smaller panel of isolates will be carefully selected according to their isolation date, serotype (if known), *prnA* and *ptxA* type, patient vaccination status, date of birth and severity of disease. These will then

be randomised into a final panel, including epidemiologically-related strains and replicates of strains, and then blinded so as not to bias experiments. This panel will be analysed using various methods including PFGE and IS*1002*-RFLP analysis, so that direct comparisons can be made between methods (Chapter 5).

1.7.6 To expand work on a DNA sequence-based approach to the genotypic characterisation of *Bordetella pertussis*

An MLST scheme based on the MLEE technique used on *B. pertussis* (Musser *et al.*, 1986), will be devised to deduce whether this approach gives enough discrimination for genotypic characterisation. Virulence genes will be tested for their variability, and inclusion of the aforementioned targets pertactin and pertussis toxin S1 subunit genes will also be considered in a sequence-based typing scheme[3]. Sequence-based typing will enable direct comparisons between laboratories for any *B. pertussis* strain. This work is featured in Chapter 6.

1.7.7 To identify novel molecular methods to type *Bordetella pertussis* isolates

Molecular methods such as PFGE and IS*1002*-RFLP analysis have been used with some success in discriminating *B. pertussis* isolates (Beall *et al.*, 1995; van Loo *et al.*, 1999). Another study comparing RAPD, ERIC-PCR and PFGE, concluded that RAPD and ERIC-PCR are simple and rapid methods, yet the poor discrimination level in both these techniques makes them unsuitable for studying the epidemiology of *B. pertussis* (Moissenet *et al.*, 1996). Single-enzyme AFLP analysis has been applied to several bacteria as a flexible, simple, quick and relatively inexpensive technique, but there has been no evidence of any prior investigations using AFLP for *B. pertussis* epidemiology. A selection of enzymes and primers will therefore be tested for their ability to type *Bordetella pertussis* isolates. Results are described in Chapter 7.

---

[3]*During the course of this thesis,* Van Loo *et al.* (2002) *reported an MLST scheme based on surface protein genes, which has shown some further discrimination and this has affected the approach of this aim.*

1.7.8 To investigate a number of epidemiologically-significant markers by using clinical material obtained from *Bordetella pertussis*-positive patients

Although primary culture/isolation of *B. pertussis* is the "gold standard" for pertussis diagnosis, it is relatively insensitive and can take up to four days for colonies to appear. The detection of pertussis infection has been successful using other methods such as serology and PCR. However, subsequent typing techniques such as PFGE and IS*1002*-RFLP analysis require a large amount of bacterial DNA to obtain adequate results. Direct sequence typing will therefore be attempted on nasopharyngeal aspirates (NPAs) and pernasal swabs (PNSs) received by RSIL that were previously culture and/or PCR-positive for *B. pertussis*. Targets such as pertactin and the pertussis toxin S1 subunit genes will be amplified by PCR directly on clinical samples and, if successful, the amplicons sequenced (Chapter 8).

Before this project began, the sequencing of the B. pertussis genome was still at the contiguous sequences (contig) assembly stage at the Wellcome Trust Sanger Institute in Cambridge. When the genome project was at the final annotation stage, the principal investigator, Dr Julian Parkhill gave me special permission to access the unpublished genome. The annotated genome was eventually published in September 2003 (Parkhill et al., 2003).

2 MATERIALS AND METHODS

All methods and procedures used at the Respiratory and Systemic Infection Laboratory (RSIL), Central Public Health Laboratory (CPHL), London, were performed using Good Laboratory Practice and with due regard to safety. In addition, Control of Substances Hazardous to Health (COSHH) regulations, appropriate Risk Assessments (RAs) and Public Health Laboratory Service (PHLS) Standard Operating Procedures (SOPs) were consulted. Patient data were handled according to Caldicott guidelines and the Data Protection Act of 1998.

The addresses of suppliers of equipment and chemicals referred to in this thesis are listed in Appendix 2. Plastic consumables were obtained from the Stores department at CPHL, unless otherwise stated.

2.1 *Bordetella* isolates

All 495 *Bordetella* strains available for this thesis are listed in Appendix 3 and summarised in Table 2.1 (page 56).

2.1.1 Historical *Bordetella pertussis* isolates

Prior to 2002, no systematic collection of *Bordetella pertussis* isolates from the United Kingdom was available for examination. Therefore, stored *B. pertussis* isolates were identified and obtained from various sites in the UK, including the Wellcome Bacterial Collection, held by the National Collection of Type Cultures (NCTC); NCTC's own reference collection, PHLS CPHL, London; the University of Glasgow, Glasgow, Scotland; Pertussis Reference Laboratory (PRL), Manchester University, Manchester; and the Centre for Applied Microbiology and Research (CAMR), Salisbury. Isolates are detailed in Appendix 3, Part A (page 297). Date of isolation, or where this was not known, the year deposited, was used for all analyses. Information concerning the depositor and source was known mostly for the isolates from the Wellcome Bacterial Collection. A total of 294 isolates were collected from 1920 to 1986, although not all were used in this thesis (see "Reasons for exclusion" in Appendix 3, Part A, page 297).

Table 2.1. Summary of bacterial isolates available for this thesis

Source of isolates (Appendix 3, part)	Organism	Encompassing period	Number of isolates
Wellcome Bacterial Collection (Part A)	*B. pertussis*	1941-1985	116
National Collection of Type Cultures (Part A)	*B. pertussis*	1920-1967	14
University of Glasgow (Part A)	*B. pertussis*	1977	11
Manchester University (Parts A and B)	*B. pertussis*	1978-1979; 1998-1999	5; 103
Centre for Applied Microbiology and Research (Part A)	*B. pertussis*	1981-1985	148
Respiratory and Systemic Infection Laboratory (Part B)	*B. pertussis*	1999-2002	35
Wellcome Bacterial Collection, National Institute for Biological Standards and Control, Aventis Pasteur, National Collection of Type Cultures, National Institute of Public Health and the Environment in The Netherlands, and Institut Pasteur (Part C)	*B. pertussis*	n/a	16
Missouri PHL, USA (Part D)	*B. pertussis*	1984-2001	45
National Collection of Type Cultures and National Institute of Public Health and the Environment in The Netherlands, (Part E)	*B. parapertussis*	n/a	2
Total available		1920-2002	495

n/a=not applicable, these strains were either used for reference, or were vaccine strains.

2.1.2 Contemporary UK *Bordetella pertussis* isolates

Bordetella pertussis isolates were collected between September 1998 and April 1999 from Manchester University (Appendix 3, Part B, page 307). This sub-set of isolates represents every *B. pertussis* isolate in England and Wales in this time period referred to Manchester University for serotyping. A selection of isolates which were submitted directly to RSIL, PHLS CPHL, London were either from the Paediatric Intensive Care Unit (PICU) study, which ran from November 1998 to November 2000 (Crowcroft *et al.*, 2003) or from the enhanced laboratory diagnosis of pertussis for England & Wales, formally established on the first of April 2002 (Anon, 2002). This latter scheme was introduced to confirm pertussis cases using PCR (for infants under six months of age with suspected pertussis) and ELISA serology (for adults with a cough persisting for over 21 days and children with a cough persisting for over 14 days), where both tests are available through RSIL. The use of the additional tests, PCR and serology has increased the diagnostic yield over culture (Fry *et al.*, 2004). Clinical and epidemiological data were available for the isolates from Manchester University, and were obtained from the enhanced surveillance database, held at the Communicable Diseases Surveillance Centre (CDSC), London, see Appendix 3, Part B (page 307). A total of 138 recent UK isolates were available for this thesis.

2.1.3 *Bordetella pertussis* vaccine and reference strains

Enquiries made to various vaccine manufacturers and the National Institute for Biological Standards and Control (NIBSC) identified three strains used in the UK whole-cell vaccine (WCV) made by Burroughs Wellcome (referred to in text as Wellcome) from 1957 to 1996. The strains CN2992B and CN3099, which are identical to CN2992 and CN4132, respectively, were obtained from the Wellcome Bacterial Collection, and strain CN5476, was obtained from NIBSC, Potters Bar (Appendix 3, Part C, page 315). From 1996, a WCV manufactured by Aventis Pasteur has been used in the UK, although unfortunately, the two strains IM1414 and IM1416, were not obtained for further evaluation. Information of their serotype, pertactin and pertussis toxin S1 genotype are, however, known (Njamkepo *et al.*, 2002). The strains used for two pertussis acellular vaccines (ACVs) used in the UK since 1999, the Tohama-I strain and strain 10536 ("Kendrick" strain), were obtained from NIBSC and Aventis Pasteur, Maidenhead, respectively.

The *B. pertussis* type strain 18323^{T} (NCTC 10739^{T}) was obtained from NCTC, CPHL (Appendix 3, Part C, page 315). Reference strains for two typing methods IS*1002*-based restriction fragment length polymorphism (IS*1002*-RFLP) analysis, and pulsed-field gel

electrophoresis (PFGE), were kindly donated by Han van der Heide from the National Institute of Public Health and the Environment in The Netherlands (RIVM) or Elisabeth Njamkepo from the Institut Pasteur in Paris.

2.1.4 *Bordetella pertussis* isolates from Missouri, USA

Forty-five clinical isolates of *B. pertussis* from Missouri PHL were sent to RSIL, for analysis, one of which was later identified as *B. parapertussis* and therefore excluded from further analysis (Appendix 3, Part D, page 316). Ten isolates were from Kansas City, which were geographically and temporally clustered within seven weeks between June and July 1999. The other isolates were epidemiologically-unrelated, with 15 collected between 1984 and 1989 and 19 between 1996 and 2001. Clinical data were limited to isolates from May 1989 to January 2001 and are detailed in Appendix 3, Part D (page 316).

2.1.5 *Bordetella parapertussis* strains

A strain of *Bordetella parapertussis*, NCTC 10853, was obtained from the NCTC and listed in Appendix 3, Part E (page 318). The *B. parapertussis* strain B24, used to produce the probe for the IS*1002*-RFLP analyses, was obtained from Han van der Heide of the RIVM, The Netherlands.

2.2 Revival, growth and preservation of *Bordetella* species

2.2.1 Revival of *Bordetella* from freeze-dried ampoules

Strains from the Wellcome Bacterial Collection, NCTC, CAMR, and the University of Glasgow, were received as freeze-dried cultures in ampoules. These were recovered as recommended by the NCTC and PHLS (CPHL and PHLS, 2000a). Ampoules were etched with a diamond cutter and a heated glass pipette applied for one to two seconds to produce an audible and encircling crack. The ampoule was left for approximately five minutes in a rack to allow the pressure inside the ampoule to equilibrate with that outside.

The following was performed in a Class 1 Safety cabinet. Approximately 0.5 ml of nutrient broth (Media Services, CPHL) was added to the ampoule and then incubated at room temperature for five to ten minutes. Two to three drops of mixed bacterial suspension were inoculated onto a pre-warmed charcoal blood agar (CHAR) plate, containing 10% volume to volume ratio (v/v) horse blood (Media Services), and left to dry for three to five

minutes. The inoculum was then streaked around the entire plate using a sterile loop. Plates were incubated at 37°C and checked for growth after three days, and incubated for at least seven days. To ensure that strains were pure, a single colony was picked and subcultured onto a fresh CHAR plate for a further two to four days.

The remaining bacterial suspension was transferred to a labelled sterile 1.5 ml microcentrifuge tube. This was centrifuged at 16,000 x*g* for five minutes (14,000 rpm on a model 5415D, Eppendorf UK Limited). The supernate was aspirated using a fresh pipette tip and the pellet stored at –80°C. These samples were used for further revival or DNA extraction attempts.

2.2.2 Culture conditions

All *Bordetella pertussis* strains were grown on CHAR plates. An addition of 40 mg/l of cephalexin was used to inhibit growth of other bacteria (CHC), where necessary. Inoculated plates were incubated at 37°C in a humid environment for three to seven days. Small, silver colonies were accepted as the morpholoigcal identification of *B. pertussis*. *Bordetella parapertussis* isolates were also grown on CHAR plates (Media Services). These were incubated using the same conditions as for *B. pertussis*, but for a shorter time interval of two to four days.

2.2.3 Preservation of *Bordetella pertussis* isolates

A Standard Operating Procedure (SOP) produced by the CPHL was followed for preservation (CPHL and PHLS, 2000b). Sterile beads (2 mm embroidery beads, Creative Beadcraft Ltd) were prepared for use by placing approximately 0.5 ml of the beads into 2 ml Sarstedt tubes, which were then autoclaved. Using aseptic techniques, bacterial culture was emulsified in approximately 1-2 ml of sterile nutrient broth with 5% (v/v) glycerol (Media Services) in a bijou, to make a heavy suspension. The suspension was mixed using a pipette, and then split into two labelled, sterile Sarstedt tubes containing beads. The tubes were capped then agitated to release any air bubbles around the beads, and then left for 5-10 minutes. The excess suspension was removed using a fresh pipette, and transferred into two to four labelled sterile 1.5 ml microcentrifuge tubes to a volume of approximately 0.5 ml per tube. The tubes were then centrifuged at 16,000 x*g* (14,000 rpm) for five minutes. The supernate was aspirated using a fresh pipette and both the cell pellets and the beads were stored at –80°C, splitting the aliquots between freezers. DNA extraction was performed using the cell pellets (2.4.1, page 60).

2.3 Serotyping of *Bordetella pertussis*

Serotyping was performed at RSIL or at Manchester University (one isolate from 1964, three from 1978, two from 1979, 37 from 1998, and 61 from 1999). The method below was used at RSIL.

Polyclonal rabbit antisera to agglutinogens 1, 2, and 3 (*B. pertussis* anti-agglutinogen 1 [no. 89/596], *B. pertussis* anti-agglutinogen 2 [no. 89/598], and *B. pertussis* anti-agglutinogen 3 [no. 89/600], NIBSC) were used in a slide-agglutination assay. Briefly, a loopful of B. *pertussis* previously grown on CHAR plates was suspended in approximately 200 µl of phosphate buffered saline (PBS) (Media services). A glass microscope slide was wiped with 70% ethanol and 5 µl each of cell suspension, and anti-agglutinogen 1, 2, or 3 were mixed and then rotated for approximately four minutes, to test for agglutination. A positive or negative reaction was recorded for each agglutinogen tested, and a negative control (PBS) was performed for each isolate on the slide. Controls NCTC 10908 (serotype 1,3) and NCTC 10907 (serotype 1,2) were tested for agglutinogens 1, 2, and 3, each time serotyping was performed.

At Manchester University, they follow a similar slide-agglutination assay, although they produce rabbit antisera in-house, which would require further standardisation and testing.

2.4 DNA extraction

2.4.1 DNA purification from pure culture

To ensure recovery of high molecular weight DNA, the Nucleon blood and cell culture (BACC) 2 Kit was used (Amersham Biosciences UK Ltd). The BACC1 protocol for small volumes were followed from the manufacturer's instructions (revision date 1997), with some minor adjustments:

Cell preparation, cell lysis and deproteinisation were performed on the bench. A *B. pertussis* pellet was re-suspended in 700 µl of Reagent B (400 mM Tris-HCl (pH 8.0), 60 mM EDTA, 150 mM NaCl, 1% w/v SDS) (350 µl if the pellet was small) by gentle mixing using a pipette, and split into two 1.5 ml microcentrifuge tubes, to approximately 350 µl per tube. One of the two aliquots was then stored at –80°C for archive. To each working aliquot, 2.5

μl of a 50 μg/ml RNaseA solution (R-4875, Sigma) were added and incubated at 37°C for 30 minutes. Deproteinisation was acheived by adding 100 μl of the 5M sodium perchlorate solution. Tubes were gently inverted several times for 10 to 20 seconds to ensure thorough mixing.

The stages where chloroform was handled were carried out in a fume cupboard. The DNA was extracted by adding 600 μl of chloroform to the suspension, and then inverting several times for 10 to 20 seconds to produce a milky-white emulsion. To produce a physical layer between the aqueous and the organic phase, 150 μl of the Nucleon silica suspension was added and inverted as before. Tubes were centrifuged at 350 x*g* for 1 minute to separate the aqueous and organic phase. Without disturbing the middle silica layer, the upper aqueous phase containing DNA in solution, was transferred to a fresh 1.5 ml microcentrifuge tube (approximately 450 μl).

The DNA precipitation and washing was performed on the bench. Approximately two volumes of cold absolute ethanol were added to the aqueous phase (900 μl). Tubes were then inverted a few times to precipitate the DNA, which appears as tiny white strands. To pellet the DNA, the tubes were centrifuged at 4,500 x*g* (7000 rpm) for 5 minutes. The supernate was pipetted or poured carefully into a suitable waste ethanol container, taking care not to disrupt the DNA pellet. One ml of 70% cold ethanol was added to the pellet and the tube inverted several times to wash the DNA. The tubes were re-centrifuged as before, and the supernate discarded.

The DNA pellet was air-dried for 15 to 30 minutes with the tube caps open and tubes inverted to remove any remaining ethanol. DNA was re-suspended in 50 μl of 1 x TE buffer (10 mM Tris (T-6066, Sigma), 1 mM EDTA (E-5134, Sigma), pH 8.00 ± 0.05) to initially dissolve the pellet. The DNA suspension was either placed on a rotary shaker (Mikroshaker 20 E, Camlab Limited) for two hours or stored overnight at 4°C to fully dissolve the DNA. The DNA extracts were then either quantified as below (2.4.2, page 61) or stored at −40°C until use.

2.4.2 **Quantification of DNA**

The concentration of extracted DNA was calculated using the GeneQuant spectrophotometer (Amersham). The DNA was measured in triplicate at A_{260nm}, with the concentration automatically calculated, based on the assumption that double-stranded

DNA has a concentration of 50 μg/ml in a 10 mm pathlength cell at an optical density (OD) of 1 (± 1% of the reading). Mean concentration values were calculated and recorded. Measurements by this method were reproducible and reliable.

2.4.3 DNA purification from clinical samples

Bacterial DNA from frozen nasopharyngeal aspirates (NPAs) and pernasal swabs (PNSs) was isolated using the Qiagen Blood & Cell Culture DNA Mini Kit (Qiagen). This kit was chosen because of its success in purifying high molecular weight DNA from various types of samples. The 'Blood and Body Fluid Spin Protocol' (revision date January 1999) was followed and all reagents were provided in the kit, unless stated otherwise.

The NPAs/PNSs samples were taken from –70°C and allowed to thaw at room temperature. Into each labelled 1.5 ml microcentrifuge tube, 40 μl of QIAGEN Proteinase K was added. To this, 400 μl each of the sample and Buffer AL was pipetted into the tube and pulse-vortexed immediately for 15 seconds. (NOTE: If the NPA/PNS was less than 400 μl, then the other reagents were adjusted accordingly, e.g., a 200 μl sample requires 20 μl QIAGEN Protease and 200 μl Buffer AL). Tubes were then incubated for 10 minutes at 56°C for the cells to lyse.

After a brief centrifuge, 400 μl (or appropriate volume) of 100% ethanol was added to the tubes to precipitate the DNA. This was briefly vortexed, then centrifuged again, before the mixture was transferred to a labelled QIAamp spin column (within a 2 ml collection tube). The columns were centrifuged at 6,000 x*g* for 1 minute to collect the DNA in the filter and remove the lysate. The collection tube was discarded and replaced with a clean tube. Buffer AW1 (500 μl) was added to the column and centrifuged at 6,000 x*g* for 1 minute to wash the DNA. The collection tube was discarded and replaced as before. A final wash with 500 μl of Buffer AW2 was performed and centrifuged at 20,000 x*g* for 3 minutes. To prevent possible Buffer AW2 carryover, the spin column was placed in a clean collection tube and centrifuged for a further minute at 20,000 x*g*. The spin column was then transferred to a sterile, labelled 1.5 ml microcentrifuge tube and 200 μl of Buffer AE, which was pre-heated to 56°C, was added. Tubes were incubated for 1 minute at room temperature to increase DNA yield and then centrifuged at 6,000 x*g* for 1 minute.

The eluate was then divided into two aliquots and stored at -80°C until required.

2.5 DNA sequencing

All the genes studied in this thesis are listed in Table 2.2 (page 65). Targets were sequenced either to screen for variation as possible epidemiological markers, or to determine their genetic stability as molecular probes (for example, the IS*1002* probe).

2.5.1 Design of oligonucleotide primers for PCR amplification and DNA sequencing

Primers were designed for the following genes: pertactin (*prnA*), adenylate kinase (*adk*), and filamentous haemagglutinin (*fhaB*). Appropriate *B. pertussis* sequences were downloaded from GenBank, National Centre for Biotechnology Information (NCBI), Bethesda, MD. (http://www.ncbi.nlm.nih.gov). For pertactin, all the variants of this gene were aligned in GeneBase (Applied Maths) (Table 2.2, page 65). Primer design was based on the complete pertactin gene sequence available (GenBank accession number, J04560), and the constructed alignment was used to check for suitable primer sites. Primers for *adk* and *fhaB* were designed using the sequences Z29715 and X53405, respectively (GenBank).

Suitable primers were chosen using GeneBase with the following considerations: (i) due to the high G+C content of the *B. pertussis* genome (67.72%), short oligonucleotides (17-20 nt) were chosen to keep the predicted temperature of dissociation (T_d) under 65°C; (ii) to eliminate hairpin loops and primer dimer formation, the "maximum internal complementary units" function in GeneBase was set to 8 to stop structures of more than 4 complementary bases within the primer(s) from occurring; (iii) the internal stability of primers was checked, such that primers that had GC-rich 5'-ends were chosen in preference to those primers with GC-rich 3'-ends; and (iv) when designing primers for PCR reactions, the primers should not complement each other, but should have similar T_d values, +/- 5°C. Candidate primers were then synthesised by MWG Biotech, purified using their High-Purity Salt-Free (HPSF®) method, see Table 2.3. (Page 66).

2.5.2 General methodology for PCR

Each gene target is described separately in 2.5.4, 2.6.1, 6.2 and 8.2 (pages 68, 81, 218, and 250), with all PCR parameters tested. All PCR reactions were prepared in a clean environment, by using a PCR cabinet decontaminated with ultra-violet light, and dedicated pipettes and tips. Latex or nitrile gloves were worn throughout the PCR preparation and all

plastic disposables were certified DNase-free. To validate all PCRs, positive and negative controls were prepared for each PCR assay, adding genomic DNA from *B. pertussis* NCTC 10739[1] and nuclease-free water (Promega), respectively. Reactions were all at a final volume of 50 µl, with the *Taq* DNA polymerase, the 10 x PCR buffer and 50 mM $MgCl_2$ obtained from Invitrogen, the dNTPs from Roche, the primers synthesised by MWG Biotech and nuclease-free water from Promega. Further details of the primers are given in Table 2.3, page 66. The amplifications were performed using the PTC-200 DNA Engine, or the PTC-225 DNA Engine Tetrad (MJ Research) in 0.2 ml PCR tubes, or the Hybaid Touchdown™ Thermal Cycler (Hybaid Ltd) using 0.5 ml PCR tubes.

Some optimisation or adaption of a published method was required. This involved either: (i) the addition of dimethyl sulphoxide (DMSO), to facilitate DNA strand separation and prevent secondary structures, (ii) the adjustment of $MgCl_2$ concentration and/or the annealing temperature to maximise the specificity of the PCR reaction, and (iii) the adjustment of DNA template concentration.

2.5.3 Separation and visualisation of PCR products

PCR products (5 µl) were loaded on a 1.25-2% (weight to volume ratio, w/v) agarose gel (Invitrogen), with 3 µl of 6 x gel-loading buffer type II, 0.25 % (w/v) bromophenol blue, 0.25 % (w/v) xylene cyanol FF, 15 % (w/v) Ficoll in water (Sambrook and Russell, 2001). Standard ladders of either 100 bp or 1 Kb (Invitrogen) were also loaded at least twice on each gel row, to allow accurate size estimation of the amplicons produced. Fragments were analysed on a horizontal gel electrophoresis apparatus in 1 x TBE (89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8.3, Sigma) at either 5.3 V/cm for 45-60 minutes in a GNA-100 tank or at 4.1 V/cm for 1-2 hours in a GNA-200 tank, (Amersham). Gels were then transferred to 1 litre of distilled water containing 1 µg/ml ethidium bromide (Sigma), and left to stain for approximately 30 minutes. Gels were then photographed using the Polaroid MP-4 system with black and white 667 film (Polaroid), under UV light at 305 nm.

Table 2.2. ***Bordetella pertussis*** **gene targets selected and sequences used to design primers**

Genes	Protein transcribed	GenBank accession numbers (to design primers, if performed)	Allele type (if any)
prnA	Pertactin	J04560	*prnA*(1)[1]
		AJ011091	*prnA*(1)
		AJ011092	*prnA*(2)
		AJ011093	*prnA*(3)
		AJ011015	*prnA*(4)
		AJ011016	*prnA*(5)
		AJ132095	*prnA*(6)
		AJ133784	*prnA*(7)
		AJ133245	*prnA*(8)
ptxA	Pertussis toxin	not applicable	
Insertion element IS*1002*		not applicable	
adk	Adenylate kinase	Z29715	
fhaB	Filamentous haemagglutinin	X53405	

[1]J04560 contains minor point mutations outside the major polymorphic regions (region 1 and 2), but has also been designated as *prnA*(1).
Primers were not designed for the targets pertussis toxin and insertion element IS*1002* for this thesis.

Table 2.3. Oligonucleotide primers used for amplification and sequencing of the *Bordetella pertussis* genes

Primer name	Sequence (5' – 3')	Gene	Position	GenBank accession number	Reference
PR8F	CCCATTCTTCCCTGTTCCAT	*prnA*	81-100	J04560	Mooi *et al.* (2000)
PF	TGTCTCTGTCACGCATTGTC		152-171		Mooi *et al.* (1998)
AF	GCCAATGTCACGGTCCAA		649-666		Mooi *et al.* (1998)
PRN707R	AGGGCGCCGATATGCAAG		707-690		this thesis
PRN1157F	CACCGCACGGCAATGTCAT		1157-1175		this thesis
AR	GCAAGGTGATCGACAGGG		1234-1217		Mooi *et al.* (1998)
BF	AGCTGGGCGGTTCAAGGT		1542-1559		Mooi *et al.* (1998)
PRN1618R	GGTCCGCGAAGACATTCAT		1618-1600		this thesis
PRN8	AGGGTAAAGGTCGCCGCGCT		1763-1744		Boursaux-Eude *et al.* (1999)
PRN1976F	ACGCGGCGGTCAACACG		1976-1992		this thesis
BR	CGGATTCAGCGCAACTC		2076-2059		Mooi *et al.* (1998)
PRN2413F	GGCAAGTACCGCACCCAT		2413-2430		this thesis
PRNF1	CAGTTCGATGCGCTTGCC		2628-2611		Boursaux-Eude *et al.* (1999)
PR	ATGCCGTTGGTGTGTACCGT		2714-2695		Mooi *et al.* (1998)
PR5R	GCCTGAGCCTGGAGACTGG		2931-2949		Mooi *et al.* (2000)
S1F	TAGGCACCATCAAAACGCAG	*ptxA*	474-493	M14378	Mooi *et al.* (1998)
S1FM	ACAATGCCGGCCGTATCCTC		946-965		Mooi *et al.* (1998)
S1R	TCAATTACCGGAGTTGGGCG		1350-1331		Mooi *et al.* (1998)
Hg1	GCCGATGCGTTCCATATA	IS*1002*	423-406	Z54268	van der Zee *et al.* (1996b)
Hg2	AGCCCTTCTTGTTAACAGGG		132-151		van der Zee *et al.* (1996b)
Bp-adkF	CGCCCCGATCACAAAAA	*adk*	119-135	Z29715	this thesis
Bp-adkM	GAGATCGAAGTGCCGGAA		478-495		this thesis
Bp-adkR	GGCGCGGTTTAGCTCTC		821-805		this thesis

Table 2.3. Oligonucleotide primers used for amplification and sequencing of the *Bordetella pertussis* genes (cont.)

Primer name	Sequence (5' – 3')	Gene	Position	GenBank accession number	Reference
fha-224F	GCCGATTACTTCACTTCGCT	*fhaB*	224-2430	X53405	this thesis
fha-525R	GGGTCGGCGATATTGACA		542-525		this thesis
fha-572R	CGACGTTGAACTGCTGGA		589-572		this thesis
fha-820F	GCCAGCAACCTGACGCT		820-836		this thesis
fha-1219F	GCCATCACCGTGTCGTC		1219-1235		this thesis
fha-1366R	GCCGACGCGATCTTCAC		1382-1366		this thesis
fha-1705F	GTCACGCTGGGTTCGGT		1705-1721		this thesis
fha-1889R	GCCCTTGGACTGCAGGT		1905-1888		this thesis
fha-2223F	CCTGGGCGATGTCCAGT		2223-2239		this thesis
fha-2361R	CGTTCACGGTCATGGCA		2377-2361		this thesis
fha-2729F	CGGTCAGGGTCGATCAA		2729-2745		this thesis
fha-2884R	GCATCCTTGCCGCTGAC		2900-2884		this thesis
fha-3182F	GCTCCATGACGCTGGGTAT		3182-3200		this thesis
fha-3482R	CGGGCGTGAAGAAATCCT		3498-3482		this thesis

The positions of the primers are according to the numbering of the reference sequence used (GenBank accession numbers given).

2.5.4 PCR amplification of the pertactin and the pertussis toxin S1 subunit genes

2.5.4.1 Pertactin gene

This PCR protocol was adapted from Mooi *et al.* (1998). To encompass the two polymorphic regions (see 1.6.4, page 46), primers AF and BR were used to produce an amplicon of approximately 1428 bp (Figure 2.1, page 69). When the entire pertactin gene was sequenced, two further amplicons were required, a 626 bp amplicon at the 5'-end generated with primers PR8F and PRN707R, and a 973 bp 3'-end fragment using PRN1976F and PR5R. Optimisation reactions (50 µl) contained 10% v/v DMSO, 1.5-3.5 mM $MgCl_2$, 10 mM Tris-HCl (pH 8.3), 2.5 mM KCl, 200 µM dNTP, 10 pmol each of forward and reverse primers, 2.5 U *Taq* DNA polymerase, and 10-100 ng template DNA.

Cycling conditions were as follows, initial denaturation at 95°C for 3-5 minutes, then 30 or 40 cycles consisting of denaturation at 95°C for either 10, 15, 20 or 30 seconds, annealing at 47-71°C for 10, 15, 20 or 30 seconds, and extension at 72°C for 45 seconds or 1 minute, followed by a final extension of 7 minutes at 72°C, with a final 4°C hold.

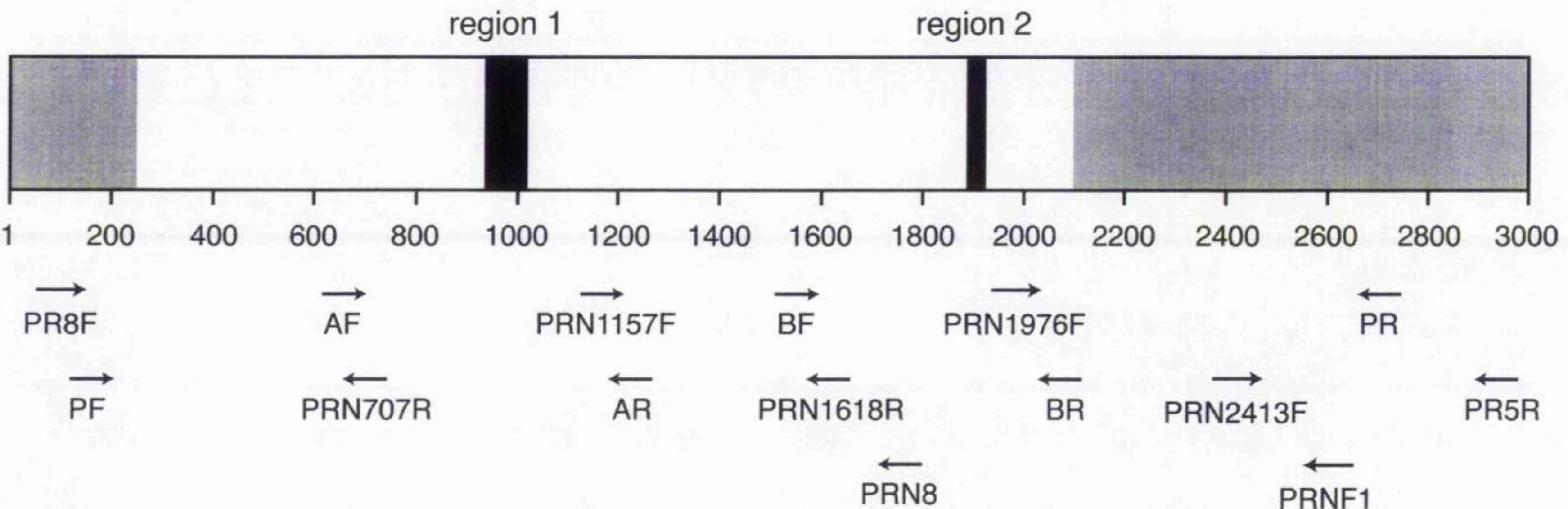
Figure 2.2 (page 70) illustrates the optimisation experiments using the primer pair AF and BR, adjusting the DNA template concentration and annealing temperature. Amplification of the 5'-amplicon (Figure 2.3, page 71) did not require optimisation, but amplification of the 3'-amplicon using PRN1976F and PR5R was optimised by adjusting the annealing temperature (Figure 2.4, page 72).

<u>Optimised conditions</u>

For amplification of the pertactin gene using any of the forward and reverse primer combinations above, the reaction mixture contained 10% v/v DMSO, 2.5 mM $MgCl_2$, 10 mM Tris-HCl (pH 8.3), 2.5 mM KCl, 200 µM dNTP, 10 pmol each of primers, 2.5 U *Taq* DNA polymerase, and 100 ng template DNA from pure culture extraction.

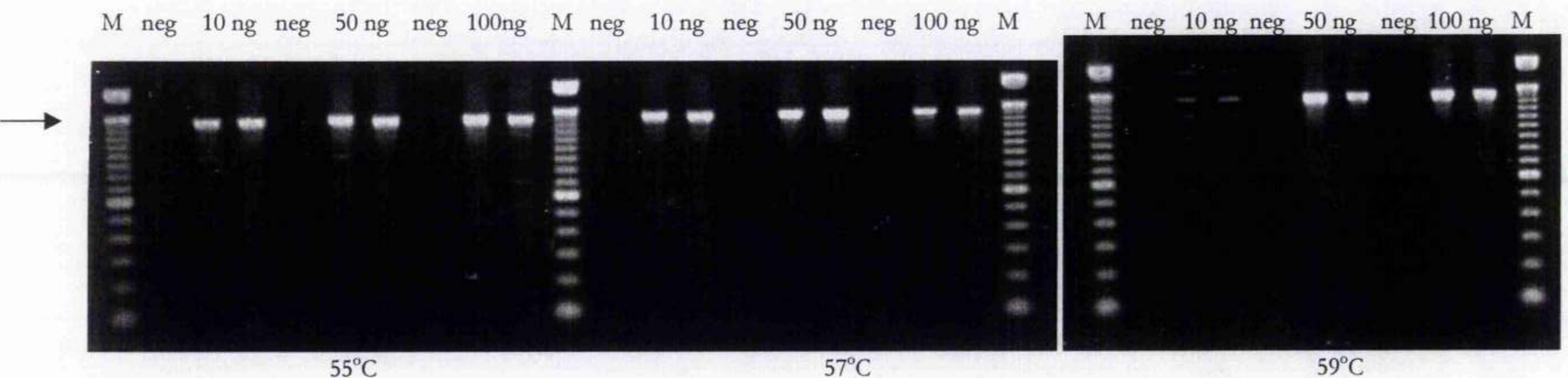
The cycling conditions used with primers AF and BR comprised an initial denaturation at 95°C for 5 minutes, then 30 cycles consisting of denaturation at 95°C for 15 seconds, annealing at 57°C for 15 seconds, and extension at 72°C for 45 seconds, followed by a final extension of 7 minutes at 72°C, with a final 4°C hold.

Figure 2.1. Schematic of the *Bordetella pertussis* pertactin gene (*prnA*) showing regions of polymorphism



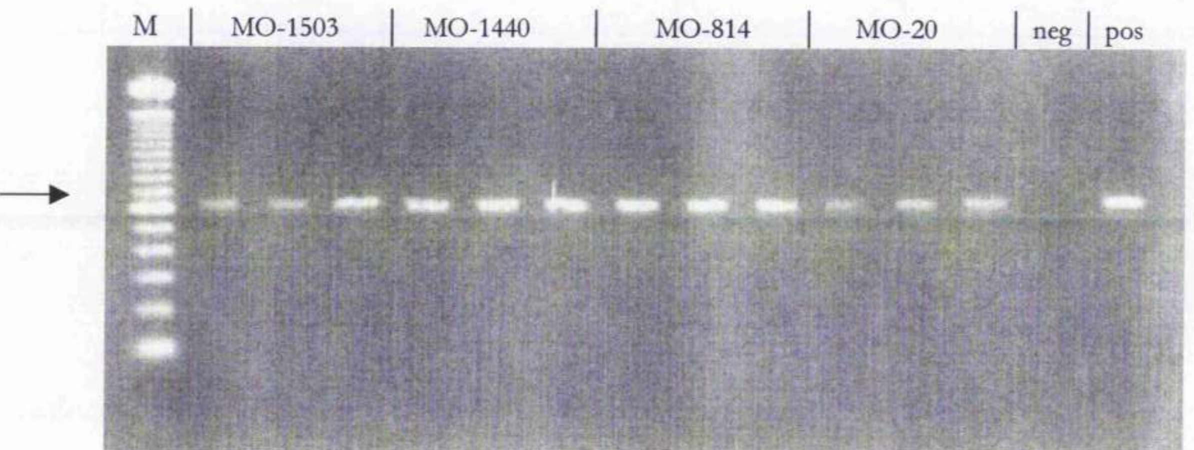
Regions 1 and 2, which code for the repeats GGxxP and PQP, respectively, are shown in black. The region including the mature protein is shown in white. The regions removed for the precursor protein are shown in grey. Numbers correspond to the base pair position as found in the GenBank accession sequence J04560. The arrows show the approximate position and direction of primers used for PCR and sequencing. Details of primers can be found in Table 2.3, page 66.

Figure 2.2. PCR amplification of the *Bordetella pertussis* pertactin gene with primers AF and BR, at various DNA concentrations and annealing temperatures



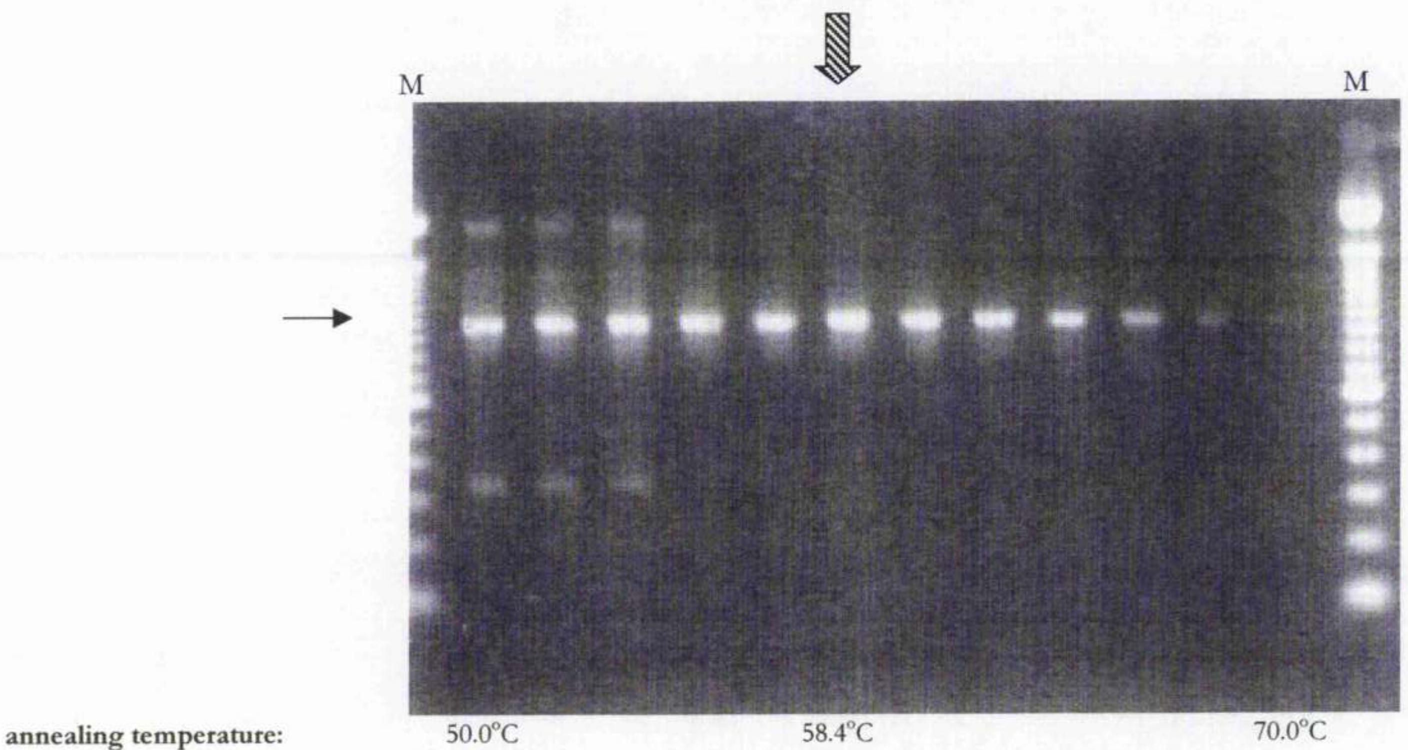
M=100 bp ladder (Invitrogen). Lanes marked neg are negative controls and paired wells correspond to 10 ng, 50 ng and 100 ng template DNA in the PCR reaction. DNA was derived from the type strain of *Bordetella pertussis*, NCTC 10739[T]. The arrow highlights the predicted amplicon size of 1428 bp, and the temperatures marked below each gel are the annealing temperatures.

Figure 2.3. PCR amplification of the 5'-end of the *Bordetella pertussis* pertactin gene with primers PR-8F and PRN707R



M=100 bp ladder (Invitrogen), neg=negative control and pos=100 ng of NCTC 10739^{T} template DNA. Triplicate DNA preparations from *Bordetella pertussis* strains from Missouri, USA were tested. The annealing temperature was 55 °C. The arrow highlights the predicted amplicon size of 626 bp.

Figure 2.4. PCR amplification of the 3'-end of the *Bordetella pertussis* pertactin gene with primers PRN1976F and PR5R, at various annealing temperatures



M=100 bp ladder (Invitrogen). 100 ng template DNA from the type strain of *Bordetella pertussis*, NCTC 10739[T] was used. The annealing temperature was varied from 50 to 70°C, from left to right on the gel. The black arrow highlights the predicted amplicon size of 973 bp and the hatched arrow indicates the optimal reaction used for subsequent experiments.

Amplification of the 5'-end of the pertactin gene needed no optimisation, and used primers PR8F and PRN707R. The cycling conditions were as follows: initial denaturation at 95°C for 5 minutes, then 30 cycles consisting of denaturation at 95°C for 30 seconds, annealing at 55°C for 30 seconds, and extension at 72°C for 45 seconds, followed by a final extension of 10 minutes at 72°C, with a final 4°C hold (Figure 2.3, page 71).

To amplify the 3'-end of the pertactin gene, primers PRN1976F and PR-5R were used. The cycling conditions were as follows: initial denaturation at 95°C for 5 minutes, then 30 cycles consisting of denaturation at 95°C for 30 seconds, annealing at 58.4°C for 30 seconds, and extension at 72°C for 45 seconds, followed by a final extension of 10 minutes at 72°C, with a final 4°C hold (Figure 2.4, page 72).

2.5.4.2 Pertussis toxin S1 subunit gene

The PCR to amplify the *ptxA* gene was also adapted from Mooi *et al.* (1998). The same primers (SF and SR) were used to generate a product of approximately 876 bp (Figure 2.5, page 74). Reactions of 50 μl were optimised with or without 5% v/v DMSO, and contained 1.5-3.0 mM $MgCl_2$, 10 mM Tris-HCl (pH 8.3), 2.5 mM KCl, 200 μM dNTP, 10 pmol each of forward and reverse primers, 2.5 U *Taq* DNA polymerase, and 10 or 100 ng template DNA.

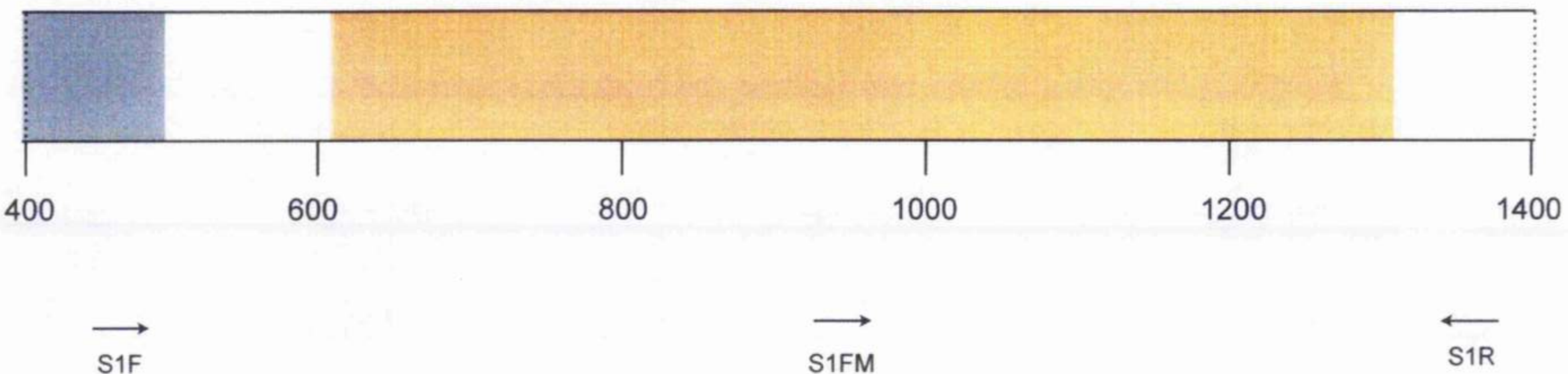
Amplification parameters for optimisation consisted of initial denaturation at 95°C for 3 or 5 minutes, then 30 cycles of 95°C for 15 seconds, annealing between 48-68°C for 15 seconds, and 72°C for 45 seconds or 1 minute, with a final extension of 72°C for 7 or 10 minutes, and a final 4°C hold.

Figure 2.6 (page 75) illustrates the optimisation experiment used with the primer pair SF and SR, and adjusting the $MgCl_2$ concentration in the presence or absence of DMSO.

<u>Optimised conditions</u>

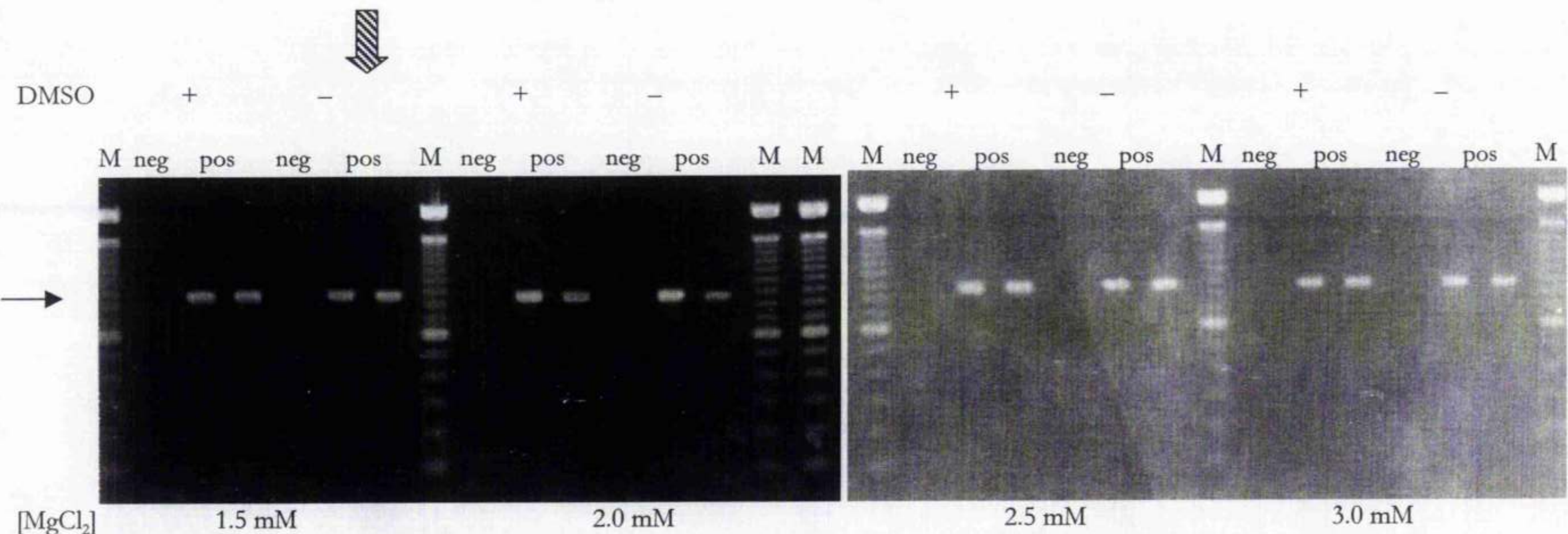
Reaction mixtures for the amplification of the pertussis toxin gene using the primer combination SF and SR contained 1.5 mM $MgCl_2$, 10 mM Tris-HCl (pH 8.3), 2.5 mM KCl, 200 μM dNTP, 10 pmol each of primers, 2.5 U *Taq* DNA polymerase and 100 ng of template DNA.

Figure 2.5. Schematic of the *Bordetella pertussis* pertussis toxin subunit 1 gene (*ptxA*)



The region encoding the mature protein is shown in orange, and the grey region indicates the bases before the start codon. The numbering corresponds to the GenBank sequence M14378, which includes the pertussis toxin operon encoding subunits S1 to S5, and is approximately 5 Kb. The arrows show the approximate position and direction of the primers used for PCR and sequencing. Details of primers can be found in Table 2.3, page 66.

Figure 2.6. PCR amplification of the *Bordetella pertussis* pertussis toxin S1 gene with primers SF and SR, at various magnesium chloride concentrations and in the presence or absence of dimethyl sulphoxide



M=100 bp ladder (Invitrogen), neg=negative controls and pos=100 ng of NCTC 10739^{T} template DNA. Duplicate reactions were performed. The $MgCl_2$ concentration ranged from 1.5 mM to 3.0 mM, with either the presence (+) or absence (–) of 5% DMSO. The annealing temperature was 59 °C. The black arrow highlights the predicted amplicon size of 876 bp and the hatched arrow indicates the optimal reaction used for subsequent experiments.

The cycling conditions used with primers SF and SR comprised an initial denaturation at 95°C for 3 minutes, then 30 cycles consisting of denaturation at 95°C for 15 seconds, annealing at 59°C for 15 seconds, and extension at 72°C for 1 minute, followed by a final extension of 10 minutes at 72°C, with a final 4°C hold.

2.5.5 **PCR product purification**

Excess dNTPs, primers and *Taq* DNA polymerase were removed from the PCR amplicon with the Wizard® PCR Preps DNA Purification System Kit (Promega). This kit was chosen for its ease of use, speed and ability to recover large amplicons (>1.5 Kb) at a high yield. Manufacturer's instructions were followed (revision date 5/01), with some alterations detailed below.

One hundred microlitres of the Wizard® PCR Preps DNA Purification Buffer were added to labelled sterile 1.5 ml microcentrifuge tubes, the number of tubes dependent on the number of amplicons purified. Triplicate PCR reactions were usually pooled to give a better yield, thus, approximately 135 µl PCR reaction was transferred to the respective labelled tube for purification. These were agitated and then 1 ml of Wizard® PCR Preps DNA Purification Resin was added. Tubes were vortexed for 3 x 5 second bursts over a 1 minute period.

If seven or more samples were processed at the same time, then the Vac-Man® Laboratory vacuum manifold (Promega) was used (maximum of 20 samples), following section V.A. of the manufacturer's instructions (2.5.5.1, page 76). Section V.B. was followed, a method using 2 ml syringes, when six or fewer samples were processed (2.5.5.2, page 77).

2.5.5.1 Vacuum manifold method for PCR product purification

The manifold was set up with a vacuum pump (BDH) and moisture trap. Labelled Wizard® Minicolumns were placed on to the required number of inlets, and a syringe barrel (3 ml, provided in kit) screwed into the Luer-Lok® extension of each minicolumn. The resin-DNA mixture was then transferred into the syringe barrel and vacuum was applied to draw the mixture into the minicolumn. This captures the resin, to which the DNA is attached.

The vacuum was stopped and 2 ml of 80% isopropanol added to the syringe barrel to wash the DNA. The vacuum was then re-applied to draw the isopropanol through the column.

When the minicolumn was clear of liquid, the vacuum was stopped and the column was transferred to a fresh, labelled 1.5 ml microcentrifuge tube. Tubes were then centrifuged at 10,000 x*g* for 2 minutes to remove residual liquid. The minicolumn was transferred to a fresh, labelled 1.5 ml microcentrifuge tube and 50 µl of nuclease-free water was added to the column to elute the DNA. Tubes were centrifuged again at 10,000 x*g* for 30 seconds to recover the DNA. The minicolumn was removed and discarded, the waste from the manifold was discarded and the purified PCR product was stored at 4°C awaiting DNA quantification (2.4.2, page 61).

2.5.5.2 Syringe method for PCR product purification

Fresh microcentrifuge tubes were labelled, and a minicolumn placed onto each tube. Disposable, sterile 2 ml syringes were dismantled, with the barrel being placed firmly onto the minicolumn, and the plunger set aside. The resin-DNA mixture was then transferred into the syringe barrel and, over a waste jar, the plunger was inserted back into the barrel. The plunger was then gently pushed to force the slurry through the minicolumn and into the waste jar.

The plunger was removed from the barrel and 2 ml of 80% isopropanol added to the resin. As before, the plunger was inserted into the barrel and gently pushed to wash the DNA and to clear the barrel of liquid. The syringe was then discarded and the tube with the minicolumn was centrifuged as detailed above in 2.5.5.1 (page 76).

Purified PCR products were stored at 4°C, until DNA quantification (2.4.2, page 61) was performed, and then stored at -30°C until required.

2.5.6 Sequencing reaction and precipitation protocols

Two systems were used to generate sequence data, both based on the dideoxynucleotide method (Sanger *et al.*, 1977) either; (i) using an ABI Prism® 377 DNA Sequencer (Applied Biosystems) at the University of Durham; or (ii) the CEQ™ 2000 (Beckman Coulter) at RSIL from April 2000, which was upgraded to a CEQ™ 2000XL in April 2001, then to a CEQ™ 8000 in September 2002. Manufacturer's instructions were followed, using 75-150 ng of purified PCR product as template.

All sequence reaction mixtures were prepared in a PCR cabinet (described in 2.5.2, page 63), and all plastic consumables and other reagents not included in the kits were previously sterilised or certified DNase-free. There were differences between the two systems in the preparation of the sequence reactions and the precipitation, and these are described below.

2.5.6.1 ABI Prism® BigDye™ Terminator Cycle Sequencing Ready Reaction Kit

This kit was used primarily for sequencing the pertactin and pertussis toxin S1 amplicons of most UK and Missouri *B. pertussis* isolates (see Chapters 3 and 4). The revised version C (1998) protocol was followed, using the "Cycle Sequencing on the DNA Thermal Cycler (TC1) or DNA Thermal Cycler 480" section for the cycling conditions. Reactions were performed in 0.5 ml thin-walled PCR tubes on the Hybaid Touchdown™ Thermal Cycler (Hybaid Ltd).

The following were added to a 0.5 ml thin-walled PCR tube; 8 µl Terminator Ready Reaction Mix (ABI), 3.2 pmol of sequencing primer, 90 ng of PCR product template and nuclease-free water to a final volume of 20 µl. A drop of mineral oil overlaid the sequence reaction to reduce evaporation during the PCR cycling. The cycling conditions were as follows; initial denaturation at 95°C for 1 minute, then 25 cycles of denaturation at 96°C for 30 seconds, annealing 50°C for 15 seconds and extension at 60°C for 4 minutes, followed by a 20°C hold.

To separate the sequence reaction mixture from the mineral oil, the entire mixture (approximately 50 µl) was transferred onto a length of Parafilm 'M' (15 cm x 5 cm). The Parafilm was then gently rocked lengthways to move and separate the mineral oil from the aqueous phase, as the oil adheres to the Parafilm. The clean sequence reaction mixture (approximately 20 µl) was then transferred to a sterile 1.5 ml Sarstedt tube and the isopropanol precipitation method was followed.

To each reaction, 80 µl of 75% isopropanol was added. The tubes were vortexed briefly to mix the isopropanol and reaction, and then left at room temperature for 15 minutes to precipitate the extension products. To separate the product from the supernate, the tubes were centrifuged at 14,000 x*g* for 20 minutes. The supernate was aspirated and discarded, taking care not to dislodge or touch the pellet, where the tubes were marked to identify the location of the pellet. The pellet was then washed with 250 µl of 75% isopropanol, to remove any residual salts, and then centrifuged for 5 minutes at 14,000 x*g*. The supernate

was again aspirated carefully and discarded. A small piece of Parafilm was stretched over the mouth of each tube and a hole was made using a pipette tip. These were then vacuum dried using a vacuum chamber for approximately 30 minutes, re-capped, packaged and sent to University of Durham for further processing on an ABI Prism® 377 DNA Sequencer.

2.5.6.2 CEQ™ 2000 Dye Terminator Cycle Sequencing with Quick Start Kit

The manufacturer's protocol, version BCI P/N 608118:AA (2000) was used with the Beckman Coulter platform. A pUC18 control template (provided in kit) was included each time a batch of sequencing reactions was performed. The PTC DNA Engines Models 200 or 225 (MJ Research) were used with 0.2 ml thin-walled PCR tubes.

Sequence reactions were set up as in 2.5.6.1 (page 78), but in labelled 0.2 ml tubes; 8 µl DTCS Quick Start Master Mix (Beckman Coulter), 3.2 picomoles of sequencing primer, 75-150 ng of PCR product template and nuclease-free water to a final volume of 20 µl. The cycling conditions were as follows; 30 cycles of denaturation at 96°C for 20 seconds, annealing at 50°C for 20 seconds and extension at 60°C for 4 minutes, followed by a 4° C hold.

For precipitating the dye terminators, a Stop Solution was prepared by mixing 3 M NaOAc (Sigma), 100 mM EDTA (Sigma) and 20 mg/ml glycerol (provided in kit) at a 2:2:1 ratio. This Stop Solution was prepared fresh on the day, corresponding to the number of reactions processed (e.g., for 8 reactions, 16 µl EDTA, 16 µl NaOAc and 8 µl glycerol is required).

The 20 µl reaction mixture was transferred into a labelled, sterile 1.5 ml microcentrifuge tube. Five microlitres of Stop Solution was added to each tube and gently agitated. Ice-cold ethanol (95%, 60 µl) was added and mixed thoroughly. The tubes were then centrifuged at 4°C for 15 minutes at 19,000 x*g* (14,000 rpm). The supernate was gently aspirated (approx. 90-95 µl), leaving a small white pellet at the bottom of the tube. The pellet was then rinsed twice with 200 µl 70% ice cold ethanol, with each rinse requiring a centrifuge at 4°C for 2 minutes at 19,000 x*g* (14,000 rpm). After the second rinse, the tube lids were closed and a hole was made in each lid to allow the pellet to dry. The tubes were then placed in the vacuum chamber for about 20 minutes, or until the pellets were dry.

The pellets were then prepared to run on the CEQ™ 2000 analyser by dissolving them in 30 µl of Sample Loading Solution (SLS, containing formamide). Care was taken to locate the pellet before opening the tube, as the pellet may be statically charged and can be lost or not properly dissolved in the SLS. The SLS was initially mixed up and down using a filter-free tip (formamide in the SLS can react with the filter polymer to produce aerosols that can degrade the deoxy terminators) until dissolved and left on the bench for 5 minutes. The sample was then transferred to the 96-well sequencing plate and overlaid with a drop of mineral oil to stop the sample from degrading and the formamide from oxidising. The sample plate could then be loaded straight onto the CEQ™ 2000 analyser to determine the sequence, or sealed with a Seal & Sample Aluminum Foil Lid (Beckman Coulter) and stored at –40°C.

2.5.7 Sequence analysis

The sequence data were analysed using both the chromatogram file, a hard copy of the chromatogram and the nucleotide sequence. All sequences from either the ABI Prism® 377 DNA Sequencer or the CEQ™ 2000 analyser, were checked manually by studying the chromatogram. The pUC18 control sequence was compared with the reference sequence for quality control purposes. The contigs were constructed by aligning the novel sequences with known sequences from GenBank. The software packages used were either GeneBase or Kodon (Applied Maths).

Alignment of sequence data was performed using GeneBase version 1.0 with a reference sequence. The hard copy chromatogram and the unknown sequence were edited manually and simultaneously. The edited sequence file was then saved in an appropriate folder and given a unique name according to its strain designation and the primer used. This sequence was clustered with other known variants of the appropriate gene and then identified to a known type, e.g., *prnA*(1), or given a new designation, if novel.

GeneBase was updated in February 2002 by Applied Maths, who introduced Kodon as a sequence and whole genome analysis software package. The chromatogram and the unknown sequence were edited semi-automatically, with both the chromatogram and sequence saved in a Kodon database. The contig was then aligned with a cluster of known variants and designated as above.

Where a sequence was of poor quality (e.g., too many ambiguities, low signal, insertion or deletion of bases due to mis-calling) the sequence reaction was repeated (2.5.6, page 77) and, in some circumstances, the target gene was re-amplified (2.5.4, page 68) and re-sequenced.

2.6 IS*1002*-Restriction Fragment Length Polymorphism analysis

This method was adapted from van Loo *et al.* (1999). One major change to this protocol was in the labelling and detection, from peroxidase and X-ray film with the Enhanced Chemiluminescence Gene Detection System (Amersham), to the digoxigenin (DIG) system (Roche), using colorimetric detection. Key steps in optimising this technique were; (i) the PCR amplification of the DIG-labelled IS*1002* probe (2.6.1, page 81), (ii) the concentration of probe (2.6.5, page 85), (iii) the temperature required for hybridisation (2.6.5, page 85), and (iv) the subsequent washing conditions (2.6.6, page 86).

2.6.1 Digoxigenin-labelling of the IS*1002* probe via PCR amplification

The 293 bp IS*1002* probe was produced by PCR using the PCR DIG Labelling Mix from Roche. This kit was chosen, as the DIG-dUTP is incorporated into the amplicon where there is a thymine base, and thus the probe becomes highly labelled and very sensitive. *Bordetella parapertussis* DNA from strain B24 was used for the amplification of the probe. This particular strain of *B. parapertussis* is used to minimise cross-hybridisation with the similar insertion elements IS*481* and IS*1001*, as strain B24 does not contain these other elements compared to other *B. parapertussis* isolates (pers comm, Han van der Heide).

Optimisation

Optimisation of this PCR reaction was performed with the following parameters; reaction mixtures (100 µl) were prepared in a PCR cabinet and transferred into 0.2 ml thin-walled tubes with or without 10% v/v DMSO containing, 2.5 mM $MgCl_2$, 10 mM Tris-HCl (pH 8.3), 2.5 mM KCl, 200 µM dNTP (containing 70 µM DIG-dUTP), 10 pmol each of Hg1 and Hg2 primers (Table 2.3, page 66), 2.5 U *Taq* DNA polymerase, and 10 ng of template DNA.

Tubes containing mixtures were transferred to a thermal cycler and amplified as follows; initial denaturation of 95°C for 5 minutes, then 30 or 35 cycles of denaturation 95°C for 45 seconds, annealing 40-64°C for 45 seconds and elongation 72°C for 1 minute, with a final

extension of 72°C for 7 minutes, and a 4°C hold. Figure 2.7 (page 83) shows the optimisation when 10% v/v DMSO is included and 35 amplification cycles are performed.

Optimised conditions

The IS*1002* element was amplified successfully when the reaction contained 10% v/v DMSO, 2.5 mM $MgCl_2$, 10 mM Tris-HCl (pH 8.3), 2.5 mM KCl, 200 μM dNTP (containing 70 μM DIG-dUTP), 10 pmol each of primers, 2.5 U *Taq* DNA polymerase and 10 ng of template DNA. The optimum cycling conditions had an initial denaturation of 95°C for 5 minutes, then 35 cycles consisting of denaturation at 95°C for 45 seconds, annealing at 50°C for 45 seconds, and extension at 72°C for 1 minute, followed by a final extension of 7 minutes at 72°C, with a final 4°C hold (Figure 2.7, page 83).

PCR products were then visualised as in Section 2.5.3 (page 64), with a 100 bp ladder (Invitrogen) to identify fragments of 293 bp. If the IS*1002* fragment was present, 3 x 100 μl mixes were pooled and purified using the Wizard® PCR Preps DNA Purification System Kit (2.5.5, page 76). Purified IS*1002* fragments were then quantified (2.4.2, page 61) and stored at –30°C until use (2.6.5, page 85).

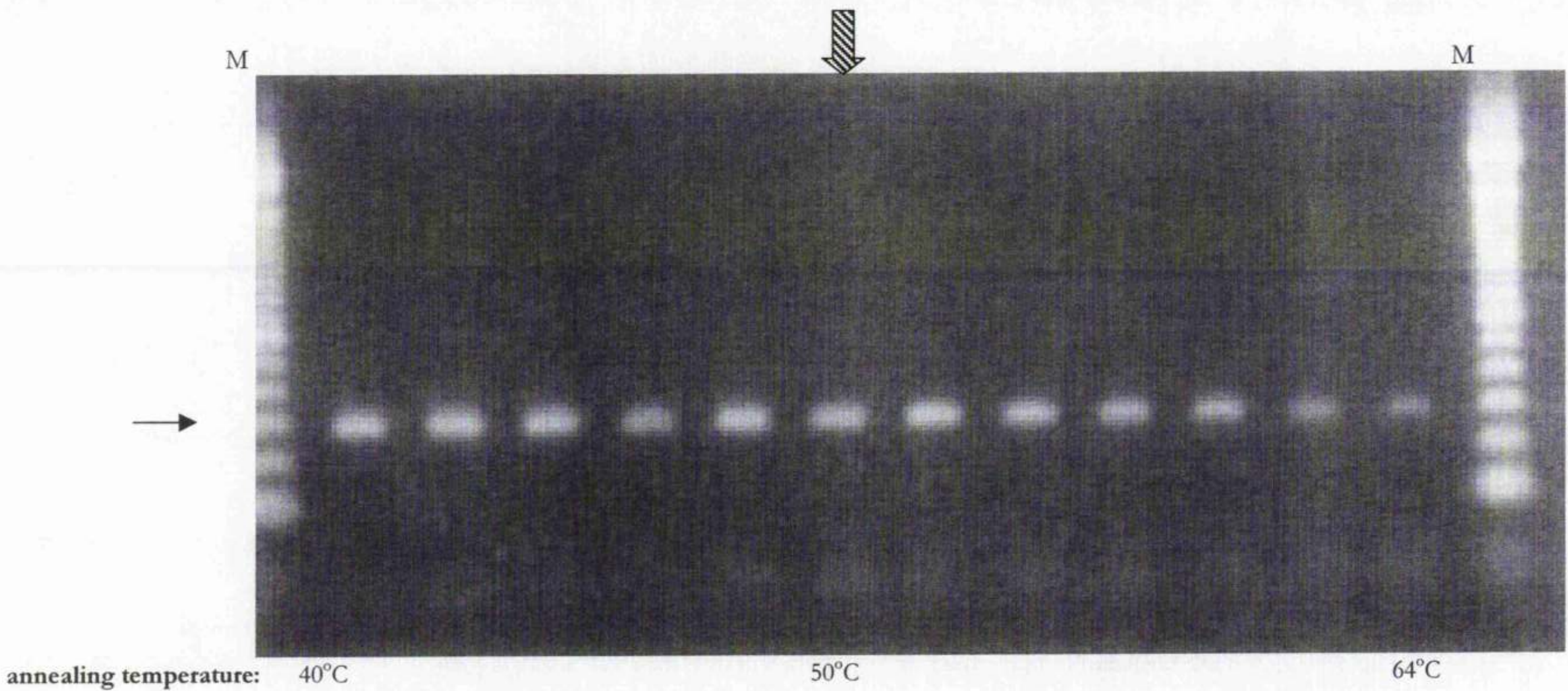
2.6.2 Restriction endonuclease digestion with *Sma*I

Five μg of extracted genomic DNA (see 2.4.1, page 60) were digested for 4 hours at 37°C with 10 U of the restriction endonuclease *Sma*I (recognition sequence CCC↓GGG, Roche) in 1 x Buffer A (provided with *Sma*I) in a final volume of 50 μl. Reactions were stopped by the addition of 15 μl of 6 x gel-loading buffer type II (Sambrook and Russell, 2001), mixed thoroughly, and products stored at –30°C.

2.6.3 Separation of DNA fragments by Field Inversion Gel Electrophoresis

The same CHEF-Mapper XA System (Bio-Rad) apparatus was used for every gel to increase reproducibility and to allow comparison of the reference strain across different blots. Electrophoresis was performed overnight for 20 hours.

Figure 2.7. PCR amplification of the *Bordetella pertussis* insertion element IS*1002*, at various annealing temperatures



M=100 bp ladder (Invitrogen). Reactions were performed in the presence of 10% v/v DMSO and 10 ng of template DNA from the B24 strain of *Bordetella parapertussis*. The annealing temperature was varied from 40°C to 64°C, from left to right on the gel. The black arrow highlights the predicted amplicon size of 293 bp and the hatched arrow indicates the optimal reaction used for subsequent experiments.

A 1% agarose gel was prepared using 200 ml of 0.5 x TBE and 2 g of pulsed field certified agarose (Bio-Rad). A gel tray of 25 cm x 16 cm was wiped with 70% ethanol and a 25-tooth comb was placed near the top of the tray. The CHEF-Mapper tank was filled with 1800 ml of 0.5 x TBE, the pump switched on at speed 7 and the chiller system set to 20°C. After cooling and casting, the gel was placed in the tank and secured well with pipette tips. The gel was left to equilibrate in the TBE buffer for at least 15 minutes.

The digests were prepared as described in Section 2.6.2. (page 82) and the molecular-weight ladders were taken from –30°C and allowed to thaw at room temperature. Two of the three ladders were applied as a mixture by initially pre-heating 1.5 µl of a 100 ng/µl dilution of CHEF DNA 8-48 Kb ladder (Bio-Rad) at 65°C for 5 minutes. After equilibrating to room temperature, 1.5 µl of a 100 ng/µl dilution of a 1 Kb ladder (Invitrogen) was added, and made up to 15 µl with 12 µl of 6 x gel-loading buffer type II. The DIG-labelled DNA molecular weight marker II (Roche) was prepared by adding 5 µl of 6 x gel-loading buffer type II to every 10 µl of marker. Twelve *Sma*I DNA digests were run on each gel along with the *B. pertussis* type strain (NCTC 10739^T), to determine the intra-gel and inter-gel reproducibility. The pump was switched off before loading the samples. The combined CHEF/1 Kb ladder mix was loaded in the first and last wells, the DIG marker in every fourth well, and the reference strain NCTC 10739^T next to the outer two DIG markers, and in the middle well (well numbers 4, 13, and 22), using a total of 23 wells. Each gel was run for 20 hours at 180 V (equivalent to 5.5 V/cm, forward voltage), and reverse voltage of 120 V (~ 3.6 V/cm), with forward and reverse switch times of 0.1 seconds, and a linear ramp. To ensure that the DNA had run into the gel, the pump was left off for 5 minutes after starting electrophoresis, then switched back on to re-circulate the TBE buffer and prevent the tank from over-heating. On completion of the run, the gel was stained in 1 µg/ml ethidium bromide for 30 minutes to visualise the restricted DNA, and a photograph was taken. The gel was then processed immediately to transfer the DNA to a nylon membrane.

2.6.4 Vacuum transfer of DNA to a nylon membrane

The gel was prepared before transfer of DNA to Hybond N^+ nylon membranes (Amersham), as follows; the gel was denatured for 2 x 15 minutes in 200 ml denaturing solution, 1.5 M NaCl, 0.5 M NaOH (Sigma), and then rinsed in sterile distilled water briefly. The gel was then neutralised in 200 ml neutralising solution, 1.5 M NaCl, 0.5 M Tris-HCl, pH 7.6 (Sigma) for 2 x 15 minutes, then rinsed again in sterile distilled water. A

piece of nylon membrane was prepared by cutting a size that encompasses the bands from the gel, with at least a 5 mm overlap, and labelled in the upper right-hand corner with the date transferred. The membrane was then equilibrated for at least 5 minutes in 100 ml of 2 x SSC solution, (0.3 M NaCl, 30 mM sodium citrate) (Invitrogen), and the solution was kept for use later.

The VacuGene XL vacuum blotting system (Amersham) was prepared following the manufacturer's instructions. A vacuum of approximately 50 cm^3 H_2O was applied and 50 ml of transfer buffer (10 x SSC buffer) were poured on to the surface of the gel. Transfer buffer was topped up when the gel surface was dry and approximately 100 ml was used in total. The vacuum was applied for over an hour to ensure adequate transfer of DNA fragments to the nylon membrane.

To prevent the DNA smearing on the membrane, the gel was first removed, and then the pump was switched off. The membrane was rinsed for two minutes in the 2 x SSC solution retained from earlier and then dried at 37°C for approximately 20 minutes between two pieces of clean filter paper. The gel was re-stained in 1 μg/ml ethidium bromide for 30 minutes and re-examined under UV illumination to ensure that DNA transfer was successful. The DNA was fixed to the nylon membrane by UV cross-linking. Each membrane was placed DNA side down on the UV transilluminator, and then exposed to UV at 305 nm for 1-2 minutes. The membrane was then either stored in a dark, dry place at room temperature indefinitely, or immediately processed following the section below.

2.6.5 Pre-hybridisation and hybridisation of membranes

The membrane was transferred to a hybridisation bag (Invitrogen) and 20 ml of DIG Easy Hyb buffer (Roche) added. Air bubbles were expelled before the bag was heat-sealed, and then incubated in a water-bath at 47°C for 4 hours. The corner of the bag was cut and the DIG Easy Hyb buffer removed.

Optimisation

An experiment to establish the optimal amount of probe, and the temperature of hybridisation was performed with eight replicate membrane strips (approximately 30 mm x 155 mm). The initial gel was run and blotted as in 2.6.3 and 2.6.4, respectively, where each strip had the DIG ladder and a *B. pertussis* strain from Missouri (MO-234, MO-83, MO-73, or MO-9). Membrane strips were labelled 1 through to 8, and treated in separate

hybridisation bags (Table 2.4.a, page 88). Concentrations of IS*1002* probe used were 3.3, 1.3, 0.7, and 0.3 pmol probe/cm^2 membrane, and two water baths were set at 42°C and 52°C. In smaller, separate bags with the individual strips, the probe was added to 0.5 ml DIG Easy Hyb buffer, the air bubbles were removed, the bag was re-sealed, and then incubated in the relevant water-bath overnight. Subsequent washing and blocking of the strips were performed together. Detection of bands were carried out in separate bags, and the volume of Buffer 3, 5-bromo-4-chloro-3-indolylphosphate p-toluidine salt (BCIP) and nitroblue tetrazolium chloride (NBT) (Invitrogen) was adjusted accordingly (see 2.6.6, below). Figure 2.8 (page 89) shows the developed strips, numbered 1 to 8.

Optimised conditions

A hybridisation temperature of 47°C and a IS*1002* probe concentration of 3.3 pmol/cm^2 in 1 ml of DIG Easy Hyb buffer was used for all subsequent experiments. The next step (2.6.6, page 86) followed immediately.

2.6.6 Washing and detection of digoxigenin-labelled bands

All washes and incubations were performed on a shaking platform in plastic boxes, to give even, thorough results for each membrane treated. The corner of the hybridisation bag was cut and the hybridisation buffer was collected using a 1 ml Gilson pipette. The hybridisation buffer was frozen at –30°C in a sterile plastic bijou, for one further use on an unhybridised membrane. The membrane was rinsed in a few mls of Wash solution A (2 x SSC, 0.1% w/v SDS), pre-warmed to 37°C, to remove any excess hybridisation buffer. A further two washes, each for 5 minutes with 200 ml Wash solution A, were carried out on the shaking platform. The membrane was then washed for 2 x 5 minutes with 200 ml Wash solution B (0.1 x SSC, 0.1% w/v SDS) at 37°C.

Optimisation

An optimisation experiment was used to test the stringency of the washing conditions by varying the concentration of SSC and the temperature of the second wash, in Wash solution B. Six replicate membrane strips (approximately 30 mm x 155 mm) were produced as above (2.6.5, page 85), and each included the DIG ladder, a *B. pertussis* strain (CN 2055) and a *B. parapertussis* strain (UK 38 or NCTC 10853). Membrane strips were labelled 1 through to 6, and were washed in separate plastic boxes (see Table 2.4.b, page 88). Volumes of Wash solution A were adjusted according to membrane size and the concentrations of SSC in Wash solution B were tested at 0.1 x and 0.01 x. Aliquots of

Wash solution B were incubated either at room temperature on a shaking platform, or in two water baths at 52°C and 68°C with gentle rocking. Washing times were the same as above, and subsequent washing and blocking of the strips and the detection of bands were performed together (see below). Figure 2.9 (page 90) shows the developed strips, numbered 1 to 6.

To prevent non-specific binding of the probe, the membrane was incubated for 30 to 45 minutes with 100 ml Buffer 2, 1% w/v blocking reagent (Roche), 0.1 M Tris-HCl, 0.15 M NaCl, pH 7.6. The anti-DIG-AP Fab fragments (Roche) were centrifuged at 13,000 xg for 5 minutes prior to initial use, and 1 minute before each use thereafter. This was to remove any small antibody aggregates that could lead to background noise (spots) at the detection stage. The antibody was diluted 1:4000 in Buffer 1 (0.1 M Tris-HCl, 0.15 M NaCl, pH 7.6) (20 μl antibody, 80 ml Buffer 1). The membrane was incubated in the antibody solution for 30 minutes, and the reaction stopped by washing the membrane twice for 5 minutes in 200 ml Buffer 1.

The membrane was equilibrated in 100 ml of Buffer 3 (0.1 M Tris-HCl, 0.1 M NaCl, 50 mM $MgCl_2$, pH 9.5) for 5 minutes. The membrane was then drained and transferred in to a plastic bag (24 cm x 18 cm). To detect the DIG-labelled bands, 25 ml of substrate solution (0.1 M Tris-HCl, 0.1 M NaCl, 50 mM $MgCl_2$, 175 μg/ml BCIP, 337 μg/ml NBT) was pipetted into the bag, the air bubbles carefully expelled, and the bag heat-sealed. The developing membrane was then incubated in the dark at room temperature. The labelled bands usually appeared after an hour but the membranes were typically incubated for 6-16 hours, depending on the background level. To stop the reaction the membrane was removed from the bag and washed in distilled sterile water. The membrane was then dried between two sheets of clean filter paper. Membranes were stored in a dark, cool environment, until the next step (2.6.7, see below).

Table 2.4. Optimisation of IS*1002*-RFLP with membrane strips

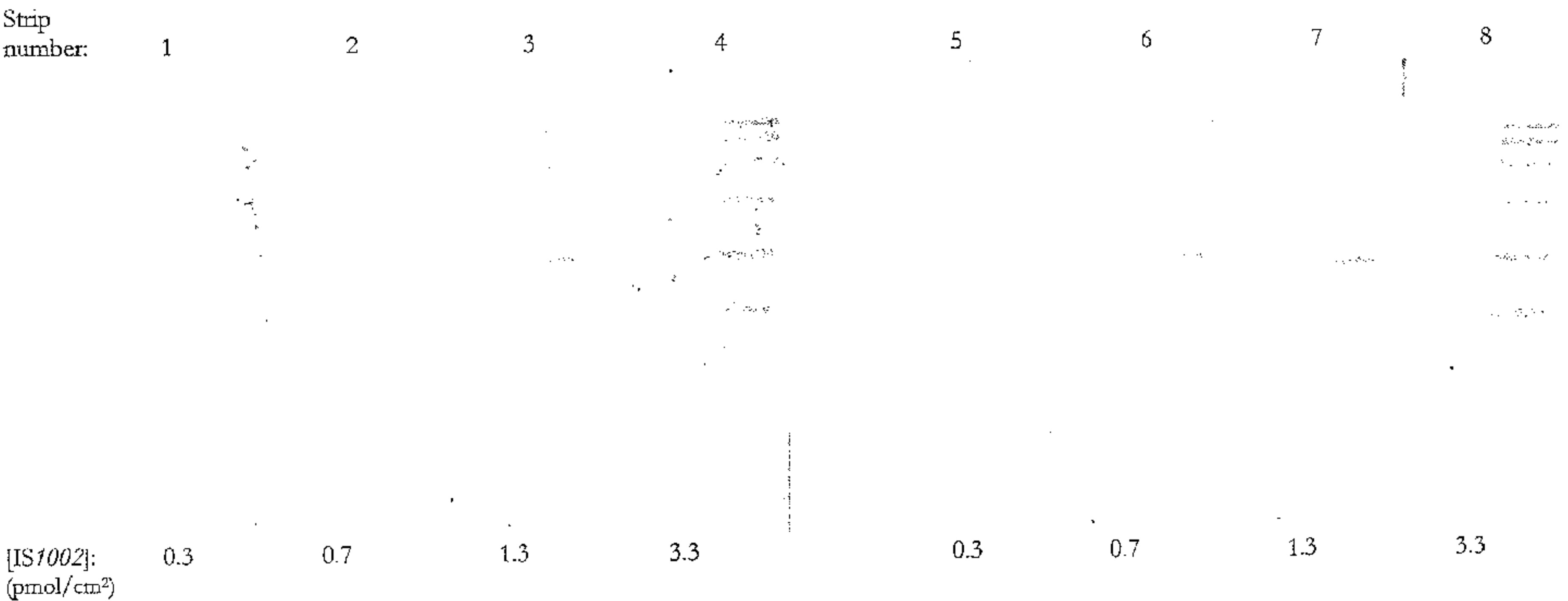
a. Membrane strips used to optimise the hybridisation temperature and concentration of IS*1002* probe.

Strip number	*Bordetella pertussis* strain	Hybridisation temperature	IS*1002* probe concentration (pmol/cm^2)
1	MO-234	42°C	0.3
2	MO-83	42°C	0.7
3	MO-73	42°C	1.3
4	MO-9	42°C	3.3
5	MO-234	52°C	0.3
6	MO-83	52°C	0.7
7	MO-73	52°C	1.3
8	MO-9	52°C	3.3

b. Membrane strips used to optimise the stringency of the washing conditions, using Wash solution B.

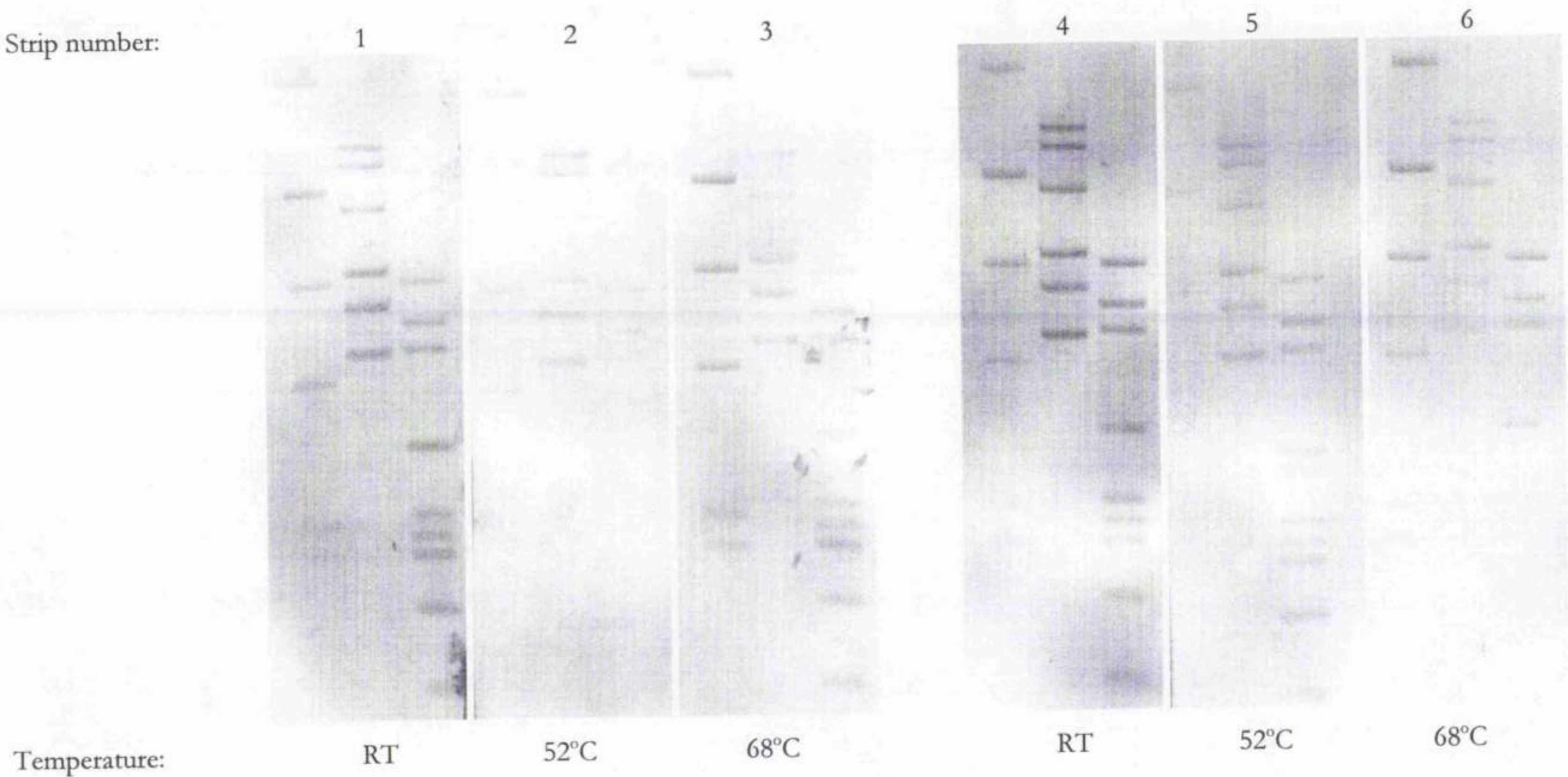
Strip number	*Bordetella pertussis* strain	*Bordetella parapertussis* strain	SSC buffer concentration	Washing temperature
1	CN 2055	UK 38	0.1 x SSC	Room temperature
2	CN 2055	UK 38	0.1 x SSC	52°C
3	CN 2055	UK 38	0.1 x SSC	68°C
4	CN 2055	UK 38	0.01 x SSC	Room temperature
5	CN 2055	NCTC 10853	0.01 x SSC	52°C
6	CN 2055	NCTC 10853	0.01 x SSC	68°C

Figure 2.8. IS*1002*-RFLP of *Bordetella pertussis*, with various hybridisation temperatures and concentrations of the IS*1002* probe



Strip number and concentration of probe per cm^2 membrane are as indicated above. Strips 1 to 4 underwent hybridisation at 42°C, strips 5 to 8 at 52°C. See Table 2.4.a (page 88) for strains used on each strip.

Figure 2.9. IS*1002*-RFLP of *Bordetella pertussis* and *Bordetella parapertussis*, with various washing conditions



Strip number and temperature of wash are as indicated above. Wash solution for strips 1-3, 0.1 x SSC, and for strips 4-6, 0.01 x SSC. RT=room temperature. See Table 2.4.b (page 88) for strains used on each strip.

2.6.7 Analysis of IS*1002*-Restriction Fragment Length Polymorphism profiles

Developed membranes were scanned at 200 dpi (HP Scan Jet ADF; Hewlett Packard), encompassing approximately the same area for each membrane, to minimise variation in analysis. The tagged information file format (TIFF) files were then imported into the BioNumerics software (version 2.5; Applied Maths), for the normalisation and band assignment of the IS*1002*-RFLP profiles. This was done in four steps according to the BioNumerics manual; i) the 21 lanes were defined using a thickness setting of 31 points with three nodes, and the background subtraction and spot removal was set at 20 points and 6 points, respectively; ii) the densitometric curves (taken from the intensity of the hybridisation on the membrane) were defined using lanes with a thickness of 19 points and two nodes. For each membrane, a spectral analysis was performed, and values from this were used for the background subtraction and the arithmetic average filtering; iii) the DIG-labelled DNA marker contained six visible bands after detection (23130 bp, 9416 bp, 6557 bp, 4361 bp, 2322 bp, and 2027 bp), and these bands were selected to normalise the profiles across the membrane. The reference system used to estimate the band sizes is dependent on the six bands in the DIG-labelled DNA marker, and the BioNumerics manual recommends using a logarithmic-dependent cubic spline fit; and iv) the initial band assignment used the following band search filter settings; minimum profiling, 5.0%; "gray zone", 5.0% (relative to the maximum value); minimum area, 0.00%; and shoulder sensitivity, 5.

After this automatic selection, the assignment of bands were checked as a quality exercise, according to the relative intensities of hybridised bands and the curve intensities. These varying intensities were due to: the amount of restricted DNA loaded on the gel; the efficiency of the vacuum transfer of DNA to the nylon membrane; the hybridisation of probe and; the washing and detection times. However, weakly hybridised bands were not assigned at the author's discretion.

Optimisation for identification of IS*1002*-RFLP types

To identify IS*1002*-RFLP types, normalised IS*1002*-RFLP profiles were compared with the Dice coefficient of similarity, and clustered using the unweighted pair group method with arithmetic averages (UPGMA). Other similarity coefficients (Pearson, Jaccard) and dendrogram types (Ward) were used to analyse the data, but these were sub-optimal when

analysed visually, and did not compare reliably with previous studies (van der Zee *et al.*, 1996b; van Loo *et al.*, 1999).

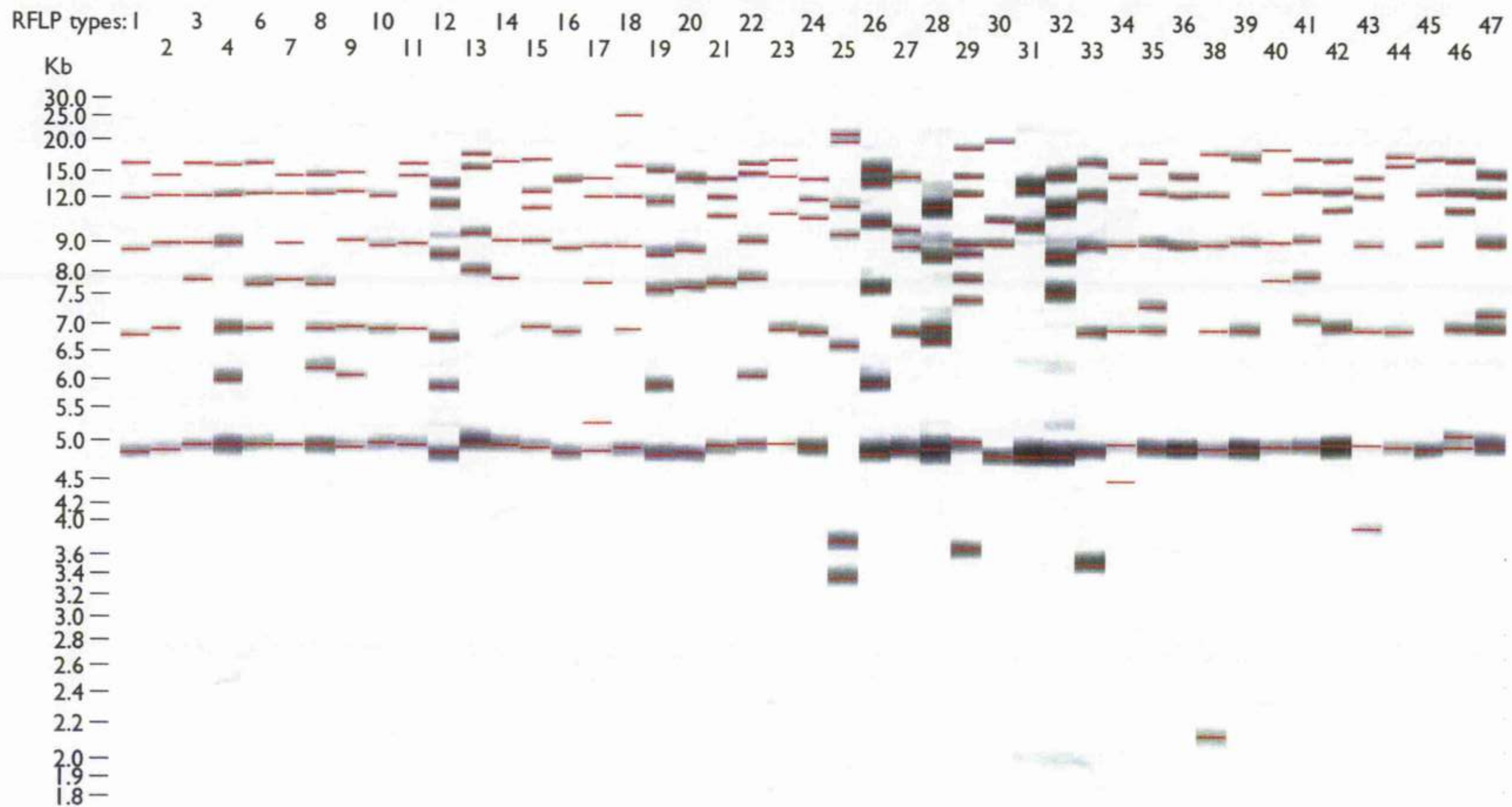
A range of band position tolerances (from 0.0 to 1.2%) and optimisation settings (from 0.0 to 1.0%) were used. A range of percentage similarity values (from 90 to 100%) was also assessed to determine the optimal settings for type allocation. In order to define optimal conditions, the data were examined visually to ensure that the percentage similarity threshold (at or above which the patterns were considered to belong to the same type) were consistent with their intra- and inter-gel appearance (Fry *et al.*, 1999). The type strain of *B. pertussis* (NCTC 10739^T) was also used as a reference strain to determine intra- and inter-gel variation. To optimise the definition of an IS*1002*-RFLP type effectively, all 291 isolates that were typed by this method were included in this optimisation stage.

Optimum conditions for identification of IS*1002*-RFLP types

The reproducibility (as measured by variation in the molecular sizes calculated with reference to the type strain) ranged from 92.1% to 100.0% (average 99.5%) for intra-gel and 91.8% for inter-gel variation. Optimal conditions for type allocation were with 0.6% optimisation and 1.0% tolerance, and using 91.8% similarity settings within the dendrogram produced. Using these criteria, a total of 45 different RFLP types were observed (Figure 2.10, page 93). The number of bands per RFLP type varied between 3 and 9, with DNA band sizes of between approximately 2096 bp to 23.3 Kb.

Initially, the type RFLP-5 was identified as a separate type to RFLP-1, because of the apparent size difference of the highest band hybridised (approximately 16.4 Kb in RFLP-5 compared to 15.4 Kb in RFLP-1). Upon further investigation, however, when clustered using the optimised conditions, RFLP-5 isolates clustered with over 91.8% similarity with RFLP-1, and therefore, isolates exhibiting RFLP-5 were re-classified as RFLP-1. This re-classification also occurred for RFLP-37 isolates to become RFLP-25 isolates, as they clustered at 93.3% similarity, which is over the 91.8% similarity cut-off.

Figure 2.10. Normalised IS*1002*-RFLP *Bordetella pertussis* patterns determined by Dice, UPGMA, and a 91.8% similarity cut-off



The thin red lines indicate bands included in the analysis. RFLP-5 and RFLP-37 have been re-classified and details of these are in 2.6.7, page 91.

2.7 Pulsed-Field Gel Electrophoresis

This method, using *Xba*I (recognition sequence T↓CTAGA), was adapted from a paper which recommended the procedure after a consensus meeting of pertussis epidemiologists from 10 countries (Mooi *et al.*, 2000). One main adaption was the removal of phenylmethylsulfonyl fluoride (PMSF), which is highly toxic and also requires dissolving in a volatile organic solvent. As PMSF stops the action of proteinase K, the omission of PMSF necessitated multiple washes to remove proteinase K activity (2.7.1, see below). The concentration of *B. pertussis* cells to be embedded in agarose, as measured by their absorbance at 650 nm, was also optimised.

2.7.1 Preparation of genomic DNA

Bordetella pertussis isolates were cultured on CHAR plates and incubated at 37°C for 48 hours. Plug moulds were prepared, with six moulds being used per isolate (each 1 x 0.6 x 0.1 cm). Approximately half a plate of *B. pertussis* colonies was suspended in 1 ml of 1 x TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA) and the absorbance was measured at A_{650nm} (Helios Epsilon Spectrometer, Unicam Ltd). Measurements of *B. pertussis* cells at A_{650nm} values of approximately 0.6, 1.0 and 1.3 were tested, and it was found that an absorbance of 1.3 at A_{650} provided the best result. A 3% solution of low melting-point agarose (SeaPlaque, BMA) was prepared in 0.5 x TBE (0.3 g in 10 ml) and cooled to 40°C. Equal volumes of the *B. pertussis* suspension and agarose (approximately 700 µl each) were then mixed carefully and pipetted in to the plug moulds. The plugs were left for 30 minutes at room temperature to harden, then transferred into 5 ml of proteolysis buffer, 0.5 M EDTA pH 8.0, 1% w/v sarkosyl (BDH), 0.5 mg/ml proteinase K (Sigma), in 25 ml glass universal tubes. Tubes were incubated overnight in a waterbath at 55°C.

The plugs were then washed five times in 5 ml of 1 x TE (pH 8.0) on a rocker for 30 minutes each at 37°C. They were stored in 10 ml of 1 x TE at 4°C, changing the TE at monthly intervals. To reduce the smearing of the PFGE profiles, the plugs were left in the fridge for at least two hours before starting the restriction step (2.7.2, see below).

2.7.2 Restriction endonuclease digestion with *Xba*I

A third of a plug (0.3 x 0.6 x 0.1 cm) was initially equilibrated at room temperature in 200 µl of 1 x buffer solution H (provided with restriction endonuclease *Xba*I, Roche) for 30 minutes. This was then replaced with 200 µl of fresh 1 x buffer solution H and 40 U of

*Xba*I (Roche), and incubated overnight at 37°C. The reaction was stopped by the addition of 500 µl of 0.5 M EDTA (pH 8.0), and stored at 4°C for several weeks, or until separation by electrophoresis.

2.7.3 Separation and detection of DNA fragments

A 1% agarose gel was prepared (SeaKem Gold, BMA), with 225 ml of 0.5 x TBE for a gel tray of 20 x 25 cm. The buffer of each restricted slice was changed to 200 µl of 0.5 x TBE and equilibrated at room temperature for 30 minutes. The melted agarose was poured in to the gel tray, retaining a few mls to seal the wells, and allowed to set for 30 minutes. Two litres of 0.5 x TBE was poured in to the electrophoresis tank of a CHEF DR II system (Bio-Rad), and the pump switched on with the chiller set to 16°C. The gel comb was carefully removed and the plugs loaded. At the first and every fourth well a lambda ladder PFG marker (~ 0.75 µg, New England BioLabs) was loaded, which comprises multiple concatemers of 48.5 Kb, increasing to 1018.5 Kb. The restricted DNA slices are also loaded in the appropriate wells, using a total of 22 wells. The wells were then sealed with the remaining agarose and left to set for 10 minutes. The gel was transferred to the tank, fixed firmly in place, and left to equilibrate for 30 minutes.

Two running programs were attempted to evaluate which method was better for separating the large restricted fragments. The first had a running time of 26 hours; ramping 4 to 30 seconds, 6 V/cm, 18 hours and ramping 30 to 50 seconds, 6 V/cm, 8 hours. However, this did not give adequate separation and, therefore, a 40 hour programme was used for all strains analysed; ramping 5 to 6 seconds, 5.5 V/cm, 16 hours and ramping 8 to 35 seconds, 5.5 V/cm, 24 hours.

Gels were transferred to 1 litre of distilled water containing 1 µg/ml ethidium bromide (Sigma), and left to stain for approximately 30 minutes and then photographed using the Polaroid MP-4 system with black and white 667 film (Polaroid), under UV light at 305 nm.

2.7.4 Analysis of Pulsed-Field Gel Electrophoresis profiles

The PFGE profiles were analysed by the same method as for the IS*1002*-RFLP membranes (2.6.7, page 91), using the same software package (BioNumerics). The inverted densiometric value was selected for the PFGE experiment type to automatically invert the image from white bands on a black background to black on white. The settings used in the

four steps when processing the membrane in BioNumerics, are as follows; i) 21 lanes were defined, each of a thickness of 23 points with five nodes, and the background subtraction and spot removal was set at 20 points and 4 points, respectively; ii) the densitometric curves were defined using lanes with a thickness of 13 points and two nodes. For each PFGE gel, a spectral analysis was performed, and values from this were used for the background subtraction and the arithmetic average filtering; iii) the lambda marker consisted of concatemers of 48.5 Kb fragments, seen visually as 10 separate fragments from 48.5 Kb to 485 Kb, with an additional high band assigned as 1018.5 Kb. These bands were selected to normalise the profiles across the gel and were used as the reference system using a logarithmic-dependent cubic spline fit; and iv) when the bands were assigned for each profile, partially-restricted fragments, seen as weak-intensity high-mass bands, (compared to their strong-intensity counterparts) were not assigned.

Optimisation for identification of PFGE types

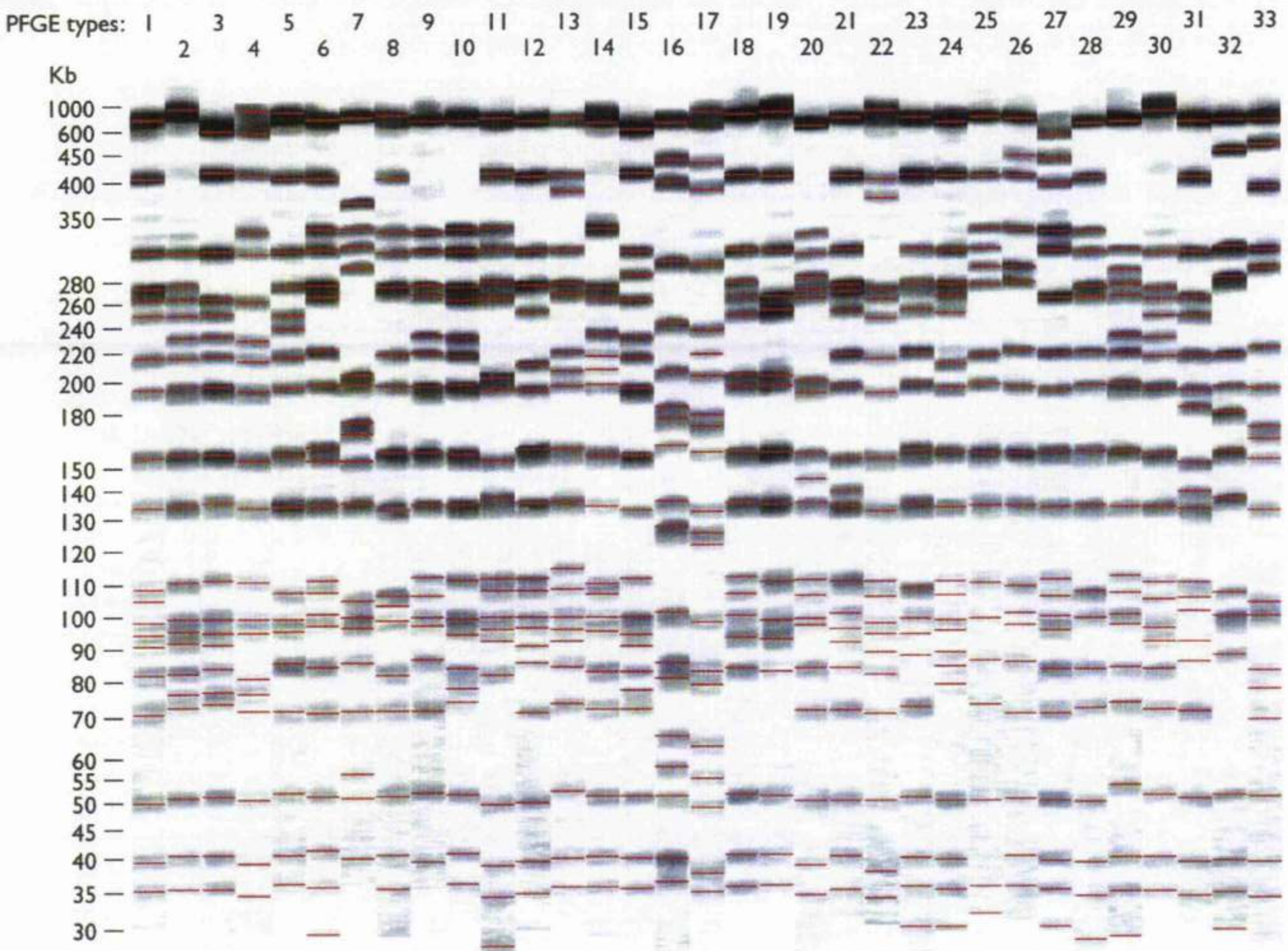
All PFGE profiles were defined in the same way as the IS*1002*-RFLP data but using only 113 isolates (2.6.7, page 91). Normalised PFGE profiles were compared with the Dice coefficient of similarity, and clustered using the unweighted pair group method with arithmetic averages (UPGMA). As for the IS*1002*-RFLP data, other similarity coefficients (Pearson, Jaccard) and dendrogram types (Ward) were used to analyse the data. However, when the profiles were checked manually after clustering, it was found that visually similar profiles were not being grouped accordingly, so these other coefficients were considered sub-optimal.

The same ranges of band position tolerances and optimisation settings used for the IS*1002*-RFLP optimisation, were attempted on the PFGE analyses (2.6.7, page 91). The type strain of *B. pertussis*, (NCTC 10739^{T}) was also used as a reference strain to determine intra- and inter-gel variation.

Optimum conditions for identification of PFGE types

The reproducibility (as measured by variation in the molecular sizes calculated with reference to the type strain) ranged from 93.4% to 100% (average 98.4%) for intra-gel and 91.5% for inter-gel variation. Optimal conditions for type allocation were with 0.65% optimisation and 1.0% tolerance, and using 94.0% similarity settings within the dendrogram produced. With these criteria, a total of 33 different PFGE types were observed (Figure 2.11, page 97). The DNA band sizes ranged between approximately 27.91 Kb and 895.27 Kb.

Figure 2.11. Normalised PFGE *Bordetella pertussis* patterns determined by Dice, UPGMA, and a 94% similarity cut-off



The thin red lines indicate bands included in the analysis.

2.8 Data analysis

2.8.1 Calculation of genotypic diversity values

Genotypic diversity (GD) values were calculated using IS*1002*-RFLP data from the UK, using the following formula (Nei and Tajima, 1981):

$$GD=(1-\sum_{i=1}^{n} x_i^2) \times n/n-1,$$

where x_i is the frequency of strains in each RFLP type and n=number of RFLP types per sub-set. Genotypic diversity values that are closer to 1 indicate a high diversity with many RFLP types, and a value approaching zero indicate few types and a low diversity. Genotypic diversity values were compared using the Pearson's Chi-squared (χ^2) test, and P values were based on 1000 computer simulations, by randomly shuffling the different RFLP types over the different sub-sets, without changing the number of RFLP types observed or the number of RFLP types found in each sub-set. The number of times the simulated squared GD exceeded the observed was counted, and then divided by 1000 to produce an estimate of the P value. This method has been described by van Loo *et al.* (1999). Confidence intervals of 95% were calculated by performing 1000 simulations from a multinomial distribution. These analyses were calculated using GLIM software (Numerical Algorithms Group Ltd) and Excel (Microsoft).

2.8.2 Identification of associations between clinical and laboratory data

The clinical and laboratory data were collected for the *B. pertussis* isolates from England and Wales, isolated between September 1998 and April 1999 (Appendix 3, Part B and C, pages 307 and 315). Relevant fields were chosen from the enhanced pertussis surveillance database, held at the PHLS Communicable Disease Surveillance Centre (CDSC), which included age, sex, geographical location, vaccination status, any detailed complications and outcome of disease. Unique identifiers (date of isolation, isolate number, name of patient and their date of birth, referral laboratory) were used to combine the clinical data with the laboratory results in Excel 2000 (Microsoft). Any discrepancies were rectified by referring to the original paperwork (see Appendix 4, page 319).

The data were then imported into STATA v 7.0 or 8.2 (Stata Corporation, Texas, USA) and analysed for any apparent significant associations. Laboratory data were compared against epidemiological data sequentially using χ^2 test. Any association with a P value under 0.200 (borderline significance) was interrogated further using multiple logistic regression analysis. Odds ratios values were then generated for each variable where possible, to compare their relative distributions and significance, with 95% confidence limits and their related P value. Linear trend analysis was also performed using STATA on the time-related aspects, age cohorts and the interval of isolation collection (temporal trend).

2.8.3 Evaluation of typing methods

The following formulae are taken from (Struelens *et al.*, 1996).

The typability (T) of a particular method can be expressed as a number between zero to one, where one represents a method that can type every isolate tested. The formula is:

$$T=N_t/N$$

where N_t is the number of isolates assigned a type and N the number of isolates tested.

Reproducibility (R) values were calculated using the formula:

$$R=N_r/N$$

where N_r is the number of isolates assigned the same type on repeat testing and N the number of isolates tested.

Indices of discrimination (D) were calculated, using Simpson's index of diversity:

$$D=1- 1/N(N-1) \times \sum_{j=1}^{S} n_j (n_j-1)$$

where N is the total number of strains in the sample population, S is the total number of types described, and n_j is the number of strains belonging to the jth type (Hunter, 1990).

Values of epidemiological concordance (E) were calculated using the formula:

$$E=N_e/N$$

where N_e is the number of strains assigned to epidemic clones and N the number of strains tested from well-defined outbreaks.

3 Population dynamics of *Bordetella pertussis* isolates in the UK

The national introduction of the pertussis WCV in the UK in 1957 helped reduce cases of pertussis to drop from 85,017 in 1957 to just 409 in 2003 (Immunisation Division, CDSC, HPA, 2004a). The UK vaccine coverage rate, at the time of writing, is 93%, although other countries with a high vaccine coverage have experienced a resurgence in pertussis. This study was designed therefore to evaluate the current situation of the *B. pertussis* population in the UK, both before and after widespread vaccination was introduced, as there has been no apparent resurgence. Recent epidemiological data was also used to identify any associations with the laboratory data generated, and whether there are any important clones of *B. pertussis* causing particular clinical outcomes. Direct comparisons with other countries such as The Netherlands and the USA were also performed and discussed.

3.1 Strategy

3.1.1 DNA sequencing of the pertactin and pertussis toxin genes

Bordetella pertussis isolates were obtained from archive collections from various sources in the UK, and recent isolates submitted to either Manchester University or the RSIL, CPHL, London (2.1, page 55). These isolates were investigated for pertactin and pertussis toxin gene variation (methods described in 2.5.4, page 68), and analysed for any temporal trends that have occurred since widespread vaccination against pertussis was introduced in 1957.

The pertactin gene types are based primarily on the variation in regions 1 and 2 (Figure 1.2, page 16), as designated by Mooi *et al.* (1998). To ensure that there were no nucleotide differences outside the two polymorphic regions 1 and 2 (Mooi *et al.*, 1998), over 90% of the pertactin gene was sequenced (3000 bp according to GenBank accession number J04560) for six *B. pertussis* UK clinical isolates and the three strains from the UK Wellcome WCV. This included the entire coding region, and required the three overlapping PCR amplicons described in 2.5.4.1 (page 68).

3.1.2 Genotypic diversity of UK *Bordetella pertussis* strains between 1940 and 1999

The genotypic diversity was calculated by following the typing method IS*1002*-RFLP analysis in section 2.6 (page 81) and using the formula detailed in section 2.8.1 (page 98). The *B. pertussis* insertion element IS*1002* is targeted, and its variation in copy number and position in the genome between different strains of *B. pertussis* makes it a suitable marker

for population studies. Genotypic diversity values could then be directly compared with those from The Netherlands, and used to elucidate whether;

> "... vaccination has selected for strains that are adapted to a highly vaccinated population ...".
>
> (van Loo *et al.*, 1999)

3.1.3 Identification of associations between clinical and laboratory data

To establish whether any associations or significant findings in a previous UK study between 1995 and 1997 (van Buynder *et al.*, 1999) were apparent in the current UK *B. pertussis* population, isolates from 1998 to 1999 from the UK were used. For these strains there are extensive epidemiological data stored at the Communicable Diseases Surveillance Centre (CDSC), London, including age, sex, vaccination history, severity and outcome of illness of the patients. Laboratory data such as serotyping, pertactin and pertussis toxin sequence types, and IS*1002*-RFLP types, were merged with the epidemiological data and interrogated to identify any association between strain characteristics and disease patterns.

3.2 Methods

The following techniques used in this chapter are described in Chapter 2; serotyping, pertactin and pertussis toxin S1 subunit gene typing (PCR amplification using primer pairs AF and BR, and SF and SR, respectively) and IS*1002*-RFLP analysis. The formula for calculating the genotypic diversity values is expressed in section 2.8.1 (page 98) and details of the association analysis are described in section 2.8.2 (page 98).

3.2.1 DNA sequencing of the pertactin and pertussis toxin genes

Before April 2000, the products of the pertactin and pertussis toxin sequencing reactions were sent to the University of Durham, and only sequencing using one sequencing primer (either forward or reverse) was done on an ABI Prism® 377 DNA Sequencer, to reduce processing costs. The sequence data received were of good quality in the regions of interest, although most contained a few unclassified nucleotides (ambiguous calls) at the beginning and the end of the generated sequence. After the Beckman CEQ™ 2000 analyser was installed in RSIL, sequencing of both forward and reverse strands for each target gene was performed. In general, the results from the pUC controls were of good quality, and therefore inadequate results were probably due to the individual sequence reaction. If a sequence reaction failed or a low signal was seen, then an increased amount of

purified PCR template was used in a new sequence reaction and re-run on the CEQ™ 2000 analyser. If the generated sequence contained mostly ambiguous calls, the initial amplification of the PCR template was repeated, the product purified, and new sequence reactions were prepared (see 2.5, page 63).

3.2.2 Random selection of UK *Bordetella pertussis* strains between 1940 and 1999 for IS*1002*-Restriction Fragment Length Polymorphism analysis

In order for the data to be within complete and successive decades, 291 *B. pertussis* isolates from 1940 to 1999, of previously determined pertactin and pertussis toxin gene types, were available for random sampling. The data were sorted initially into decades from the 1940s to 1990s, and then numbered sequentially from 1 to *x* within each decade. The "randbetween" function was used in Excel to select the isolates randomly. The selection was performed in three stages, with the first stage to deduce whether there was a relative change in RFLP types between the decades. The nine available isolates from the 1970s were tested and a further 12 isolates from each other decade were chosen. The second selection stage was performed to increase the number of isolates tested and, therefore, all 19, 13 and 9 isolates from the 1950s, 1960s and 1970s were tested, respectively, and 12 further isolates were chosen at random from each of the 1940s, 1980s and 1990s. At the third selection stage, a third set of 12 isolates were randomly selected from the 1940s, 1980s and 1990s. Therefore the following 149 isolates were investigated by IS*1002*-RFLP: 1940s, 36; 1950s, 19; 1960s, 13; 1970s, 9; 1980s, 36; and 1990s, 36. Although there were more isolates available from the 1940s, 1980s and 1990s, they were not chosen or tested to reduce the chance of skewing the data.

3.2.3 Identification of clinical and laboratory associations of pertussis cases from the UK between 1998 and 1999

Epidemiological data from the CDSC/Manchester University database were collected between September 1998 and April 1999 (n=194) and matched with the 99 *B. pertussis* isolates obtained from the PRL as in 2.8.2 (page 98). After merging the two data-sets, 90 of the 99 cases were available for further analyses. The information fields available for each case are listed in Table 3.1 (page 107). For age-related data, only the age group was used rather than age in weeks/months or years, for easier analysis and handling of data. The age groups chosen are also used in the CDSC and are designed to fit the incidence of a childhoood illness. The geographical regions were used rather than the laboratory or

GP/hospital address, due to the large number of different laboratories and GPs/hospitals, and these are shown in Figure 3.1, page 108.

Statistical analysis of the data was performed in two stages. Firstly, associations between all epidemiological variables were assessed by cross tabulating them and calculating Pearson's χ^2 probability to test for independence. This analysis helps identify variables that are highly related. Linear trend analysis using the score test for trends in odds was also performed to identify any trends by age or vaccine doses. Any associations which gave a probability value of <0.20 were used in the multiple logistic regression analyses to determine the independent effect of these factors.

The second stage entailed the analysis of associations between the epidemiological data and the laboratory typing data. This analysis was performed as above, using Pearson's χ^2 probability, linear trend analysis and multiple logistic regression analysis.

3.3 Results

3.3.1 Polymorphism in the pertactin gene

The pertactin sequence data of the six *B. pertussis* UK clinical isolates and the three strains from the UK Wellcome WCV were assembled into contigs (over 90% of the *prnA* gene, between 2768 and 2810 base pairs), and each sequence was deposited in GenBank via the BankIt submission entry (http://www.ncbi.nlm.nih.gov/BankIt/index.html), under unique accession numbers (Table 3.2, page 109). All nine complete sequences were identical to one of the pertactin types *prnA*(1), *prnA*(2) or *prnA*(3) described previously by Mooi *et al.* (1998).

Region 1 (containing the polymorphic GGxxP repeat motif) of the pertactin gene was sequenced in 318 UK *B. pertussis* isolates. To differentiate *prnA*(7) from *prnA*(1), as previously described by Mooi *et al.* (1998), where *prnA*(7) has a point mutation approximately 150 nucleotides upstream of region 2 (containing the PQP repeat motif), this was sequenced in 68 randomly-selected isolates of type *prnA*(1), four of type *prnA*(2), and one of type *prnA*(3). All isolates were identical in this region when compared to their respective pertactin types including the 68 *prnA*(1) isolates, and all 73 isolates contained five PQP repeat units in region 2 (see Figure 1.2, page 16). There is, therefore, 99%

confidence that <7.5% of the population described here as *prnA*(1), could actually have the single point mutation resulting in the designation *prnA*(7) proposed by Mooi *et al.* (1998).

Figure 3.2 (page 110) shows the temporal analysis of pertactin types of the 318 circulating UK *B. pertussis* isolates in the years 1920 to 2002. Of the isolates obtained between 1920 and 1979 all 97 (100%) (95% confidence intervals (CI): 96-100%) were of type *prnA*(1), which is the same pertactin type as the three strains used in the UK Wellcome WCV vaccine. In 1982, the novel pertactin types described by Mooi *et al.* (1998), *prnA*(2) and *prnA*(3), were seen. From 1982 to 1986, 32/88 (36%) (95% CI: 26-47%) were non-*prnA*(1), i.e., different from the *prnA*(1) type present in the UK Wellcome WCV. No clinical isolates from the UK were available for examination from 1987 to 1997. Of the available isolates analysed from recent years (1998 to 2002), 81/133 (61%) (95% CI: 52-69%) were non-*prnA*(1). The five results from 2000 were generated directly from previously *B. pertussis* PCR-positive clinical specimen extracts, and not from culture positive isolates (see Chapter 6). The decrease in the proportion of *prnA*(1) isolates collected between 1982 and 1986 compared to those isolated in 1998 to 2002, is statistically significant (χ^2=11.8, P=0.0006). Since 1996, the Aventis Pasteur WCV has also been used in the UK, where the pertactin types of the strains are also *prnA*(1) (Njamkepo *et al.*, 2002).

3.3.2 Polymorphism of the pertussis toxin S1 subunit gene

Only two variants, *ptxA*(1) and *ptxA*(2), were found in the 318 UK *B. pertussis* isolates analysed, and these two types also occur in the UK Wellcome WCV; *ptxA*(1), CN3099 and CN5476; and *ptxA*(2), CN2992 (CN2992B tested). Figure 3.3 (page 111) shows the temporal analysis of the *ptxA* types of the circulating UK *B. pertussis* isolates from 1920 to 2002. The frequency of *ptxA*(1) in isolates before vaccination was introduced, from 1920 to 1956, was 55% (40/73, 95% CI: 43-66%), compared to 100% (133/133, 95% CI: 97-100%) in isolates from 1998 to 2002. A linear trend analysis of the change in proportion of *ptxA*(1) out of the total number of isolates available was performed and found to be significant (χ^2 =104.225, P<0.0001). The pertussis toxin S1 subunit types of the Aventis Pasteur WCV, used since 1996, are *ptxA*(2) and *ptxA*(3) (Njamkepo *et al.*, 2002).

Table 3.1. Epidemiological information available for 90 cases from whom *Bordetella pertussis* isolates were recovered between September 1998 and April 1999

Factor	Parameter type	Data complete?	Included in analysis
Gender	Binary (male/female)	100%	Yes
Date of birth	Date	100%	No
Age (in years)	Number	100%	No
Age (in months)	Number	100%	No
Age (in weeks)	Number	62%	No
Age group (five groups)	<2 months 2-5 months 6-11 months 1-4 years 5+ years	100%	Yes
Laboratory	Variable	100%	No
General Practitioner/hospital	Variable	99%	No
Geographical region (nine regions)	Anglia and Oxford North Thames North West Northern and Yorkshire South Thames South and West Trent Wales West Midlands	100%	Yes
Date of specimen	Date	100%	Yes
Clinically typical symptoms seen	Binary (yes/no)	92%	Yes
Complications	Binary (yes/no)	90%	Yes
Conjunctivitis	Binary (yes/no)	90%	Yes
Pneumonia	Binary (yes/no)	90%	Yes
Apnoea	Binary (yes/no)	90%	Yes
Convulsions	Binary (yes/no)	90%	Yes
Hosptial admission	Binary (yes/no)	89%	Yes
Vaccination status	Binary (yes/no)	94%	Yes
Number of vaccine doses	0 to 4 doses	94%	Yes
Mortality	Binary (yes/no)	3%*	Yes
Other comments	Variable	21%	No

*The incomplete data concerning the death of patients were due to inadequate responses from the enhanced questionnaires (Appendix 3, Part B), where most were unknown.

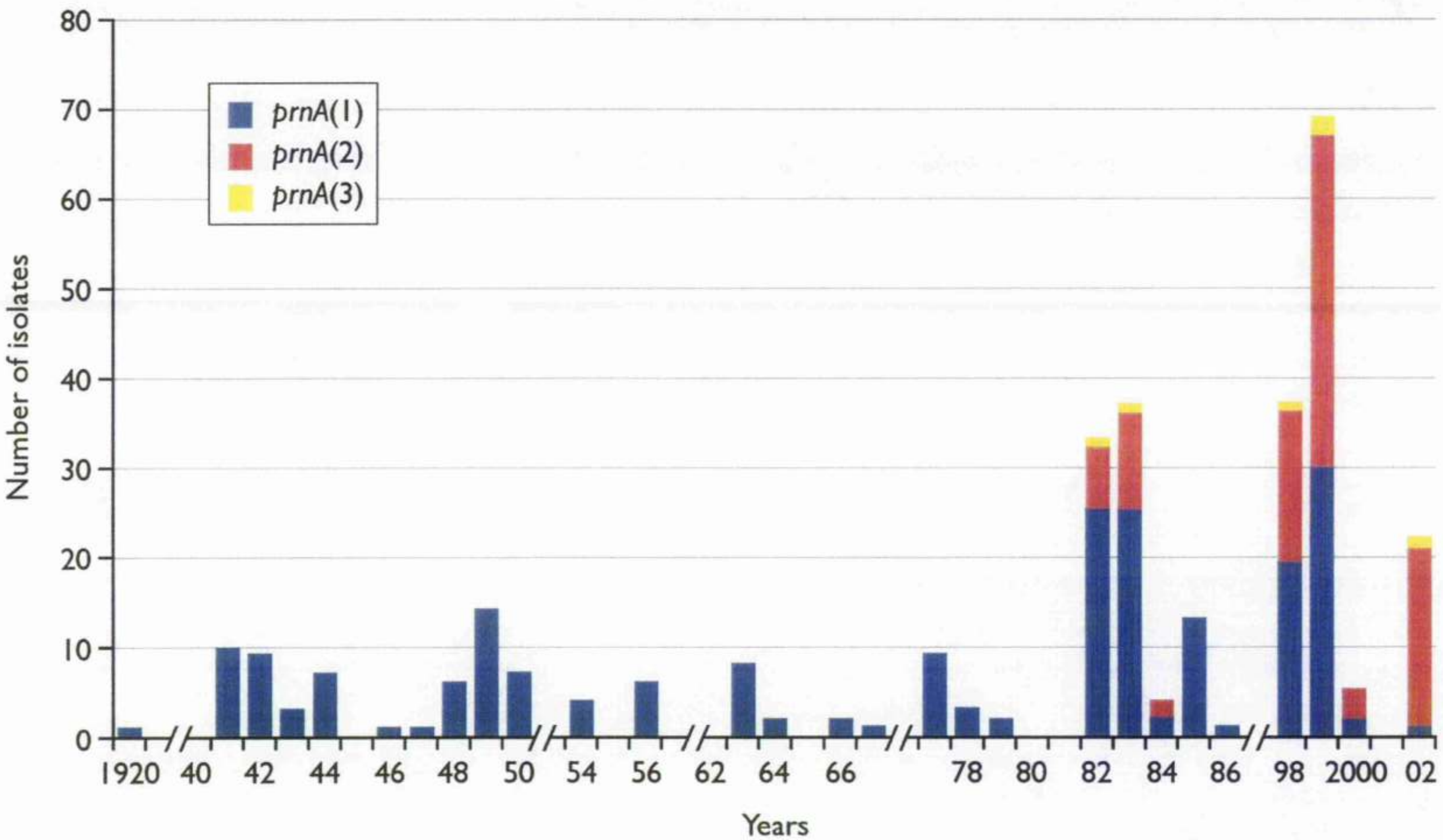
Figure 3.1. National Health Service regional boundaries of England and Wales before 1st April 1999

Table 3.2. UK *Bordetella pertussis* isolates investigated for nucleotide polymorphisms outside regions 1 and 2[1]

Isolate	Year deposited/ isolated	Origin	Pertactin type	GenBank Accession number
CN351	1942	Clinical isolate	*prnA*(1)	AF348480
DCH53	1982	Clinical isolate	*prnA*(1)	AF348481
DCH132	1983	Clinical isolate	*prnA*(2)	AF348482
98K300	1998	Clinical isolate	*prnA*(1)	AF348483
98K320	1998	Clinical isolate	*prnA*(2)	AF348484
99K45	1999	Clinical isolate	*prnA*(3)	AF348485
CN2992B[2]	1949	UK Wellcome WCV	*prnA*(1)	AF348486
CN3099	1950	UK Wellcome WCV	*prnA*(1)	AF348487
CN5476	1963	UK Wellcome WCV	*prnA*(1)	AF348488

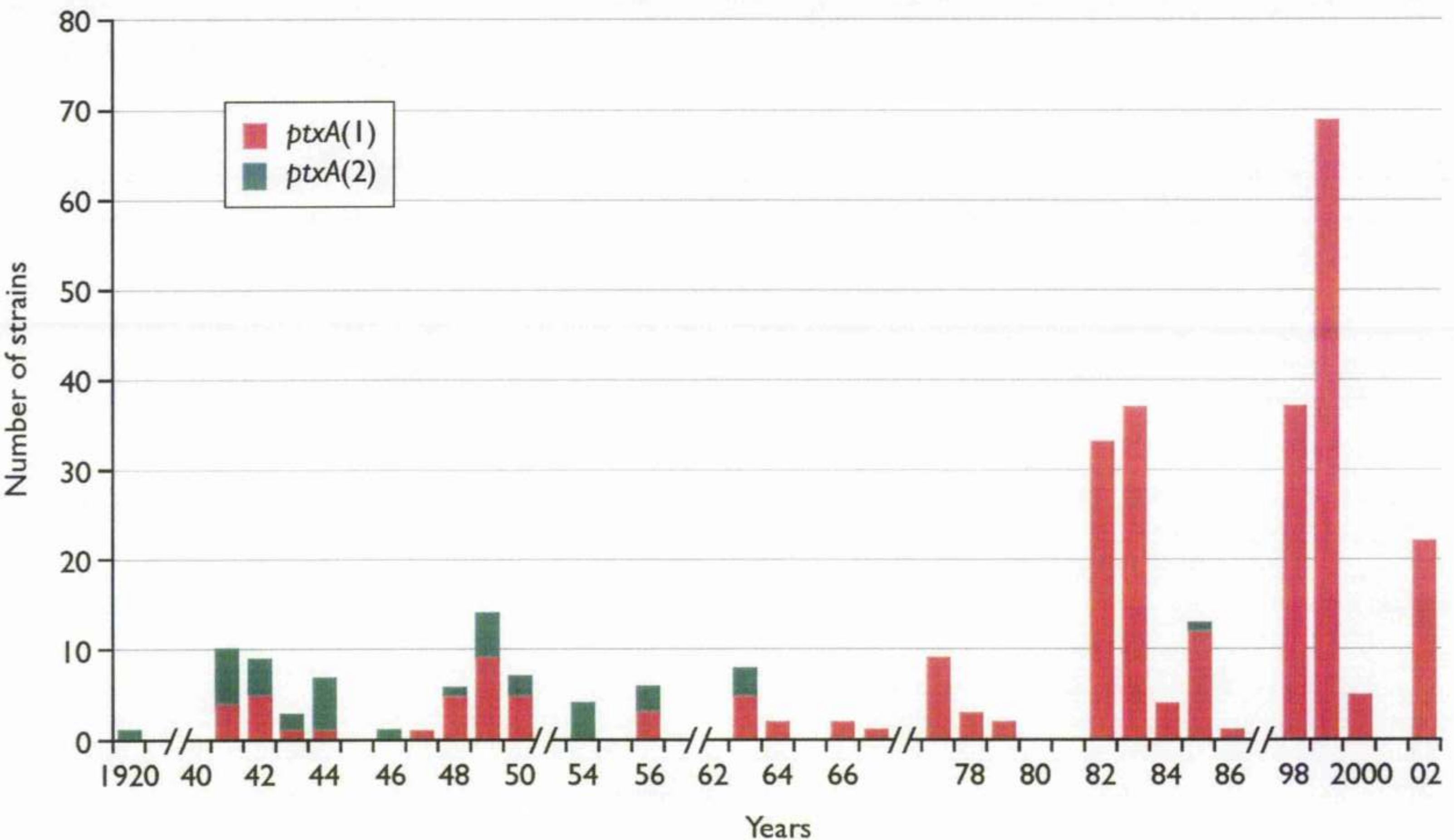
[1]This was accomplished by DNA sequencing 90% of the pertactin gene.
[2]CN2992B is equivalent to CN2992, a Wellcome WCV strain.

Figure 3.2. Pertactin gene variants between 1920 and 2002, in *Bordetella pertussis* isolates from the UK



Diagonal lines on the *x* axis indicate periods for which no isolates were available. All three strains from the UK Wellcome WCV contained *prnA*(1) and since 1996, a WCV from Aventis Pasteur has been used, containing strains of type *prnA*(1) (Njamkepo *et al.*, 2002).

Figure 3.3. Pertussis toxin S1 gene variants between 1920 and 2002 in *Bordetella pertussis* isolates from the UK



Diagonal lines on the *x* axis indicate periods for which no isolates were available. The UK Wellcome WCV contained *ptxA*(1), (n=2) or *ptxA*(2), (n=1), and since 1996, a WCV from Aventis Pasteur has been used, containing strains of type *ptxA*(2), (n=1) or *ptxA*(3), (n=1) (Njamkepo *et al.*, 2002).

3.3.3 Combined pertactin and pertussis toxin types

Data from 318 isolates were analysed by combining their pertactin and pertussis toxin S1 subunit types. As three *prnA* types and two *ptxA* types were seen in the UK, six theoretical combinations were possible. However, only four *prnA/ptxA* combinations were observed in the UK. Two of the UK Wellcome WCV strains, CN3099 and CN5476 are *prnA*(1)/*ptxA*(1), whilst the third vaccine strain, CN2992B is *prnA*(1)/*ptxA*(2). The two strains in the Aventis Pasteur WCV are *prnA*(1)/*ptxA*(2) and *prnA*(1)/*ptxA*(3) (Njamkepo *et al.*, 2002). Type *prnA*(1)/*ptxA*(1) was present in isolates from 1941 to 2002 (Figure 3.4, page 115). Type *prnA*(1)/*ptxA*(2) was present in 34/78 isolates (44%) from 1920 to 1963 and one isolate from 1985, but from 1982 to 2002, *prnA*(2)/*ptxA*(1) types appeared in 97/221 of isolates (44%), and to a lesser extent *prnA*(3)/*ptxA*(1) (6/221, 3%). This shift was very significant (Fisher's test, $P<0.001$, $n=318$). All four *prnA/ptxA* types seen in the UK were present in the 1980s.

3.3.4 Genotypic diversity of UK *Bordetella pertussis* between 1940 and 1999

Genetic diversity was determined by IS*1002*-RFLP analysis. The 149 UK isolates, chosen at random, generated 24 different RFLP types according to the presence and relative motility of bands containing the element IS*1002* when hybridised to the DIG-labelled probe (see 2.6.7, page 91 and Figure 3.5, page 116). The number of bands per RFLP type varied from four to six. Isolates with five bands represented 125/149 (84%) of the *B. pertussis* population tested and were present in all the decades (Figure 3.6, page 117). Restriction fragment length polymorphism types containing six bands were observed in all decades except the 1990s and, conversely, types with only four bands were present from the 1960s onwards, after mass vaccination was introduced. This shift from more bands in earlier decades to less bands in later decades, was significant (Fisher's test, $P=0.029$, $n=149$).

Data were sorted by decade and plotted on a bar chart to show the distribution of RFLP types (Figure 3.7, page 118). Of the 24 different RFLP types, 11 were seen in the 1940s, six in the 1950s, 1960s and 1970s, 11 in the 1980s, and six in the 1990s. The most abundant type (designated RFLP-1), occurred in 64/149 (43%) isolates, and was present in all decades. Nine other RFLP types were also present in more than one decade, including RFLP-2, which accounted for a further 25/149 (17%) of the *B. pertussis* population tested. The three Wellcome WCV strains exhibited one of each of the types RFLP-1, RFLP-2, and RFLP-3, and these types, when combined, were seen as a major proportion in isolates in the 1940s, 1980s, and 1990s at 24/36 (67%), 25/36 (69%), and 32/36 (89%), respectively.

Information regarding the RFLP types of the Aventis Pasteur WCV strains was not available.

When comparing the distribution of RFLP types between each decade, RFLP-3 through to RFLP-34 had to be grouped together because of the low numbers found in these types, in order to perform a valid chi-square test. Thus, RFLP-1, RFLP-2 and "other-RFLP" were compared for any significance between the decades. The 1940s isolates comprised 8/36 (22%) of type RFLP-1, and this distribution changed significantly to 10/17 (59%) RFLP-1 isolates in the 1950s (P=0.031). The proportion of RFLP-1 isolates then decreased from the 1950s to the 1980s (P=0.12). However, by the 1990s, RFLP-1 was the most common type and seen in 23 of the 36 (64%) isolates, with a significant change in the distribution from the 1980s to the 1990s (P=0.032).

Genotypic diversity (GD) values were calculated for each decade depending on the frequency and proportion of each RFLP type, and then plotted with 95% confidence intervals (Figure 3.7, page 118). Genotypic diversity values that are closer to 1 indicate a high diversity with many RFLP types, and a GD value approaching zero indicates few types and a low diversity. After the introduction of the pertussis WCV in 1957, the GD dropped from the 1940s (0.81) to the 1950s (0.64), (P=0.092). The GD of the UK *B. pertussis* population then increased gradually from the 1950s (0.64) to the 1980s (0.84), (P=0.06). However, in the 1990s the genotypic diversity significantly decreased from the level seen in the 1980s (0.84 compared to (cf.) 0.57, P<0.01).

Genotypic diversity values of RFLP types were also calculated with respect to pertactin types, pertussis toxin types, and combined *prnA/ptxA* types (Table 3.3, page 119). Isolates with *prnA*(1) (n=118) had a significantly higher GD value than *prnA*(2) isolates (0.85, cf. 0.39, respectively, P<0.001). Confidence of the *prnA*(3) genotypic diversity (GD=0.00) could not be made because only one isolate was present. Isolates displaying pertussis toxin type, *ptxA*(1) had a higher GD value than *ptxA*(2) (0.72 and 0.59, respectively), but was not significant (P=0.108). When the pertactin and pertussis toxin data were combined, because of the high GD values of *prnA*(1) and *ptxA*(1), the *prnA*(1)/*ptxA*(1) isolates exhibited a high GD (0.80), with respect to the RFLP type distribution. The three other *prnA/ptxA* types had lower GDs, at 0.59, *prnA*(1)/*ptxA*(2), 0.39, *prnA*(2)/*ptxA*(1), and 0.00, *prnA*(3)/*ptxA*(1) (latter type based on one strain only). All comparisons between the GD values of the combined types were significant; P=0.003 between *prnA*(1)/*ptxA*(1) and

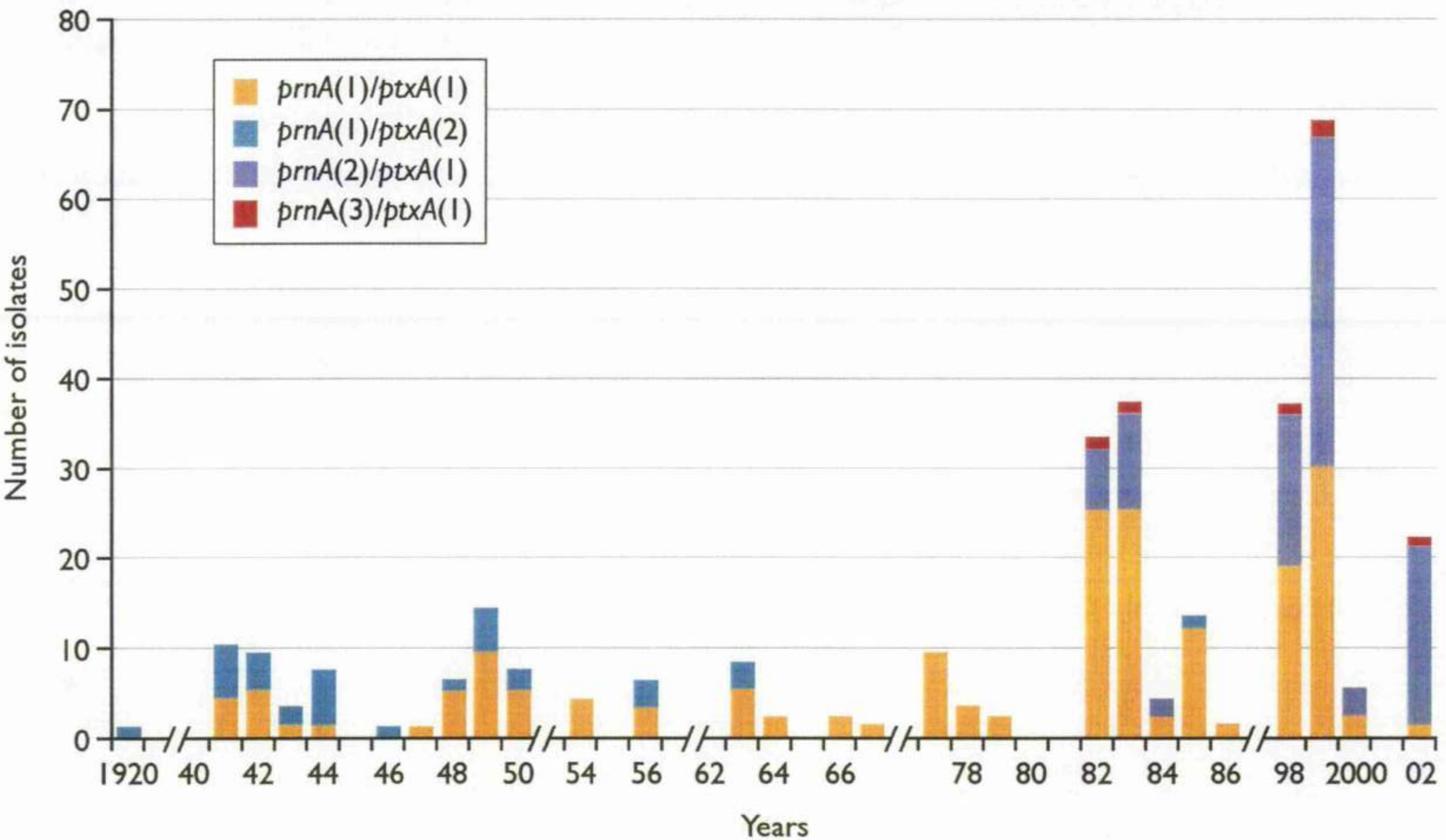
prnA(1)/*ptxA*(2); P<0.001 between *prnA*(1)/*ptxA*(1) and *prnA*(2)/*ptxA*(1); and P=0.005 between *prnA*(1)/*ptxA*(2) and *prnA*(2)/*ptxA*(1).

Restriction fragment length polymorphism types containing more than six isolates were included when investigating associations between RFLP types, pertactin, and pertussis toxin types. Therefore, only isolates from types RFLP-1, RFLP-2, RFLP-3 and RFLP-4 were investigated, and the *prnA*(3) isolate was also omitted (n=114/149). Types RFLP-2 and RFLP-4 showed a strong association with *prn*A(1) (Fisher's test, P=0.002), and RFLP-1 and RFLP-3 were exclusively associated with *ptx*A(1) (P<0.001). Pertussis toxin type, *ptxA*(2) is significantly associated with RFLP-2 (P<0.001), and also with *prnA*(1) (P=0.003).

3.3.5 Epidemiological associations and trends of pertussis cases from the UK between 1998 and 1999

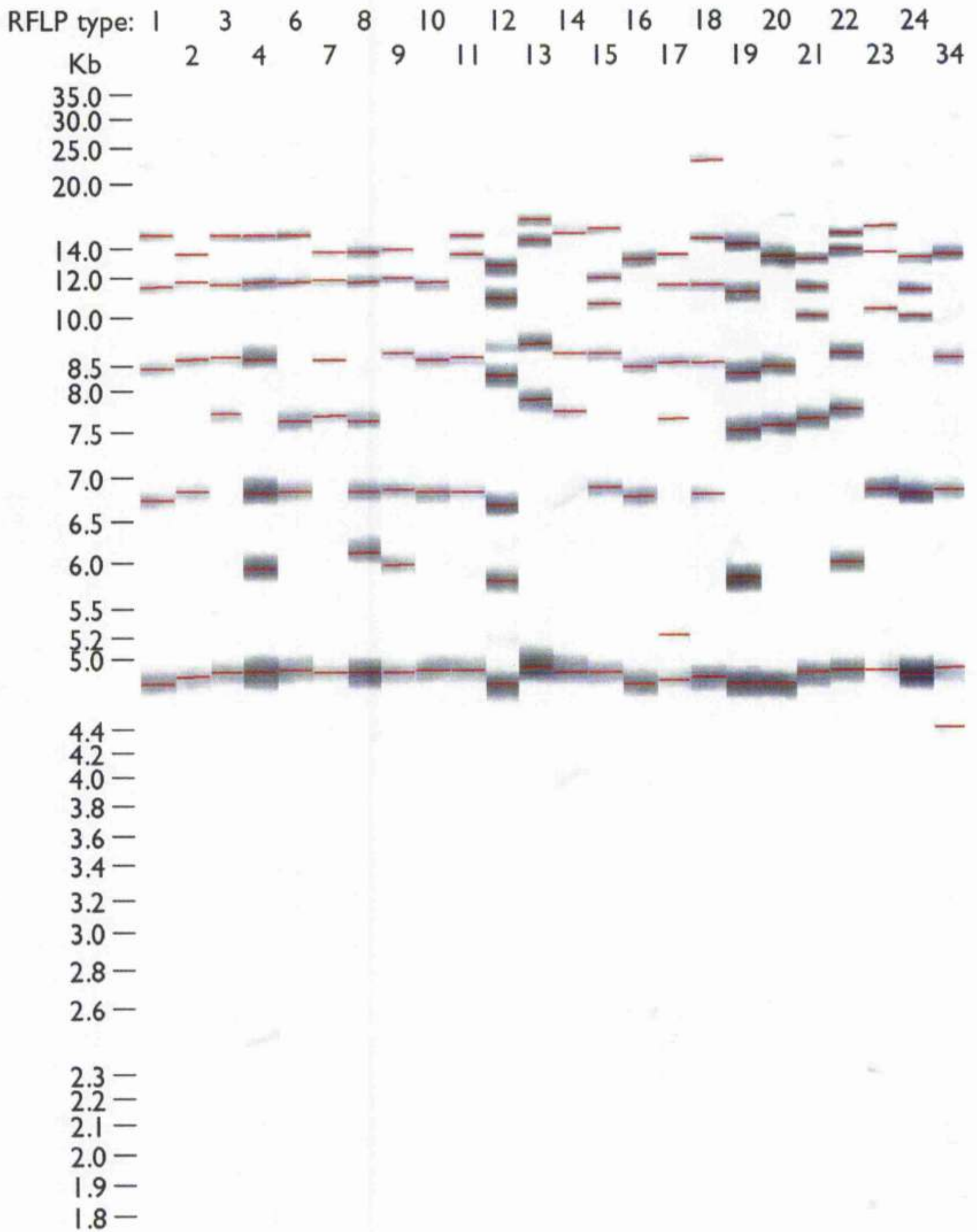
Associations within the epidemiological data were investigated, as any significant findings could skew and alter the interpretation of any further analyses. Table 3.4 (page 120) details the major aspects of the 90 cases and the proportion of cases in each category. The sub-set consisted of 43.3% males, and ages varied from three weeks to 45 years, and when sorted into five cohorts; 25.6% were under two months, 32.2% were between two and five months, 4.4% were between six and 11 months; 12.2% were between one to four years; and 25.6% were five years and over. Pertussis cases were seen in all nine geographical regions, varying between 1.1-28.9% (Table 3.4, page 120). Details of vaccination status include the number of doses the individual received (Table 3.4, page 120), and 52.2% received no vaccination against pertussis. The majority of cases showed clinically typical pertussis symptoms (75.6%) and 25.5% cases encountered other complications such as apnoea, conjunctivitis and pneumonia. Although data were available for cases with convulsions, none of the 90 cases in this sub-set presented with this complication. Sixty percent were hospitalised and three individuals of the 90 died due to pertussis in this eight-month period.

Figure 3.4. Temporal trends of combined *prnA/ptxA* types of UK *Bordetella pertussis* isolates



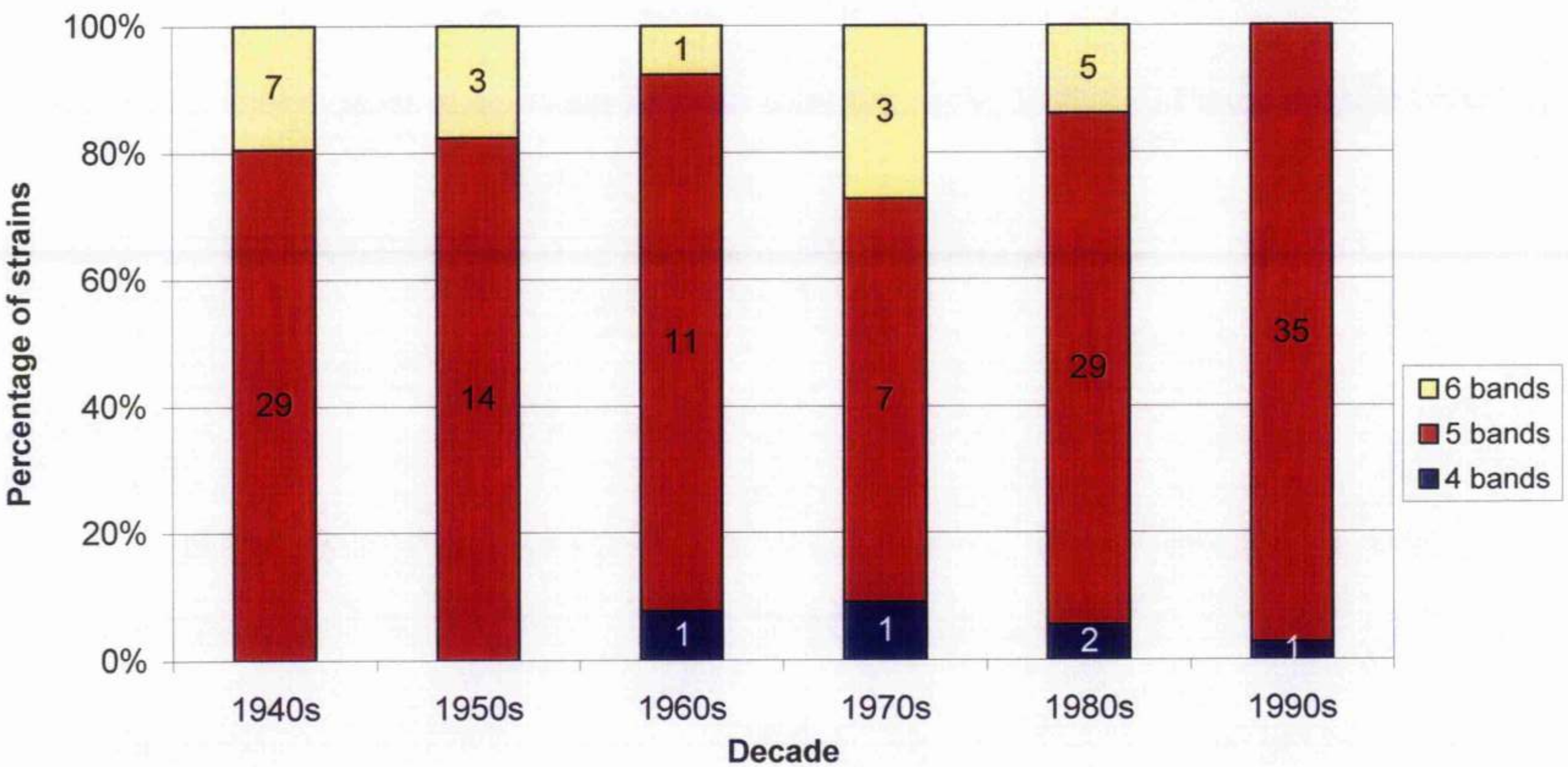
Diagonal lines on the *x* axis indicate periods for which no isolates were available. The Wellcome WCV strains are *prnA*(1)/*ptxA*(1), (n=2) or *prnA*(1)/*ptxA*(2), (n=1) and since 1996, a WCV from Aventis Pasteur has been used, containing strains of type *prnA*(1)/*ptxA*(2), (n=1) or *prnA*(1)/*ptxA*(3), (n=1) (Njamkepo *et al.*, 2002).

Figure 3.5. Normalised IS*1002*-RFLP types of UK *Bordetella pertussis* isolates from 1940 to 1999



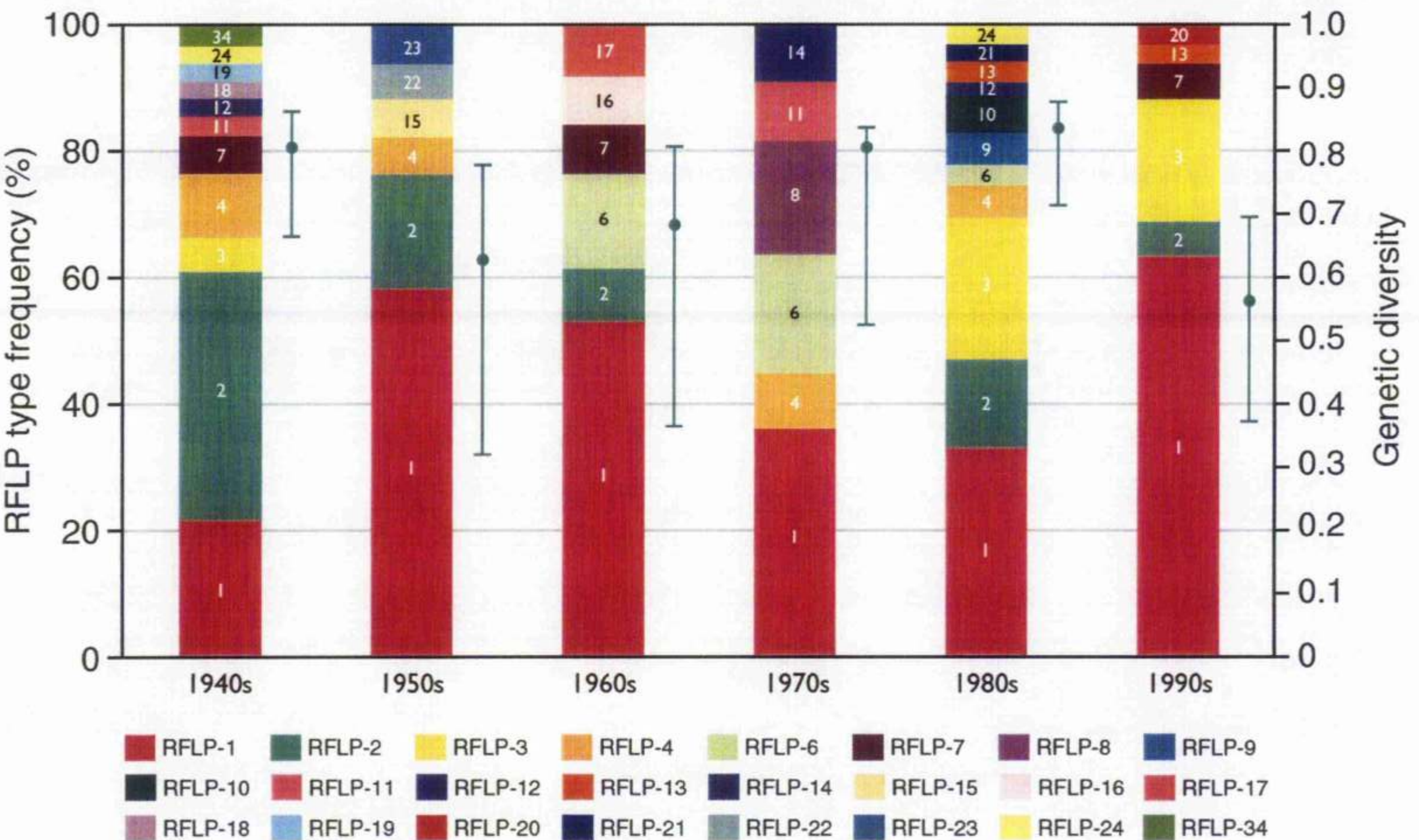
Total of IS*1002*-RFLP types=24, n=149 isolates. The thin red lines indicate bands included in the analysis.

Figure 3.6. Temporal trends of *Bordetella pertussis* according to the IS*1002*-RFLP band number in the UK



Values in bars are number of strains.

Figure 3.7. Changes in the population structure of *Bordetella pertussis* in the UK by IS*1002*-RFLP type



Genetic diversity values are shown with 95% confidence intervals as green vertical lines. These values were generated from IS*1002*-RFLP data, as previously described in 2.8.1 (page 98). *Bordetella pertussis* strains in the UK Wellcome WCV are RFLP-1, RFLP-2, and RFLP-3.

Table 3.3. Diversity of *Bordetella pertussis* IS*1002*-RFLP types, according to the pertactin and pertussis toxin S1 gene types, and combined *prnA/ptxA* types

IS*1002*-RFLP types	Pertactin gene type			Pertussis toxin S1 gene type		Combined *prnA/ptxA* type			
	prnA(1)	*prnA*(2)	*prnA*(3)	*ptxA*(1)	*ptxA*(2)	*prnA*(1)/ *ptxA*(1)	*prnA*(1)/ *ptxA*(2)	*prnA*(2)/ *ptxA*(1)	*prnA*(3)/ *ptxA*(1)
1	41	23		64		41		23	
2	24	1		7	18	6	18	1	
3	11	6		17		11		6	
4	8			6	2	6	2		
6	5			5		5			
7	5			2	3	2	3		
8	2			2		2			
9	2			2		2			
10	2			2		2			
11	2			2		2			
12	2			2		2			
13	1		1	2		1			1
14	1			1		1			
15	1				1		1		
16	1			1		1			
17	1				1		1		
18	1			1		1			
19	1			1		1			
20	1			1		1			
21	1				1		1		
22	1			1		1			
23	1				1		1		
24	2			1	1	1	1		
34	1			1		1			
Number of isolates	118	30	1	121	28	90	28	30	1
Number of types	24	3	1	20	8	20	8	3	1
GD	0.854	0.387	0	0.721	0.588	0.795	0.588	0.387	0

Table 3.4. Distribution of major epidemiological factors in the 90 UK pertussis cases analysed from September 1998 to April 1999

Factors	Category	Number of cases	Percentage of cases
Gender	Male	39	43.3
	Female	51	56.7
Age group (from 3 wks to 45 yrs)	<2 months	23	25.6
	2-5 months	29	32.2
	6-11 months	4	4.4
	1-4 years	11	12.2
	5+ years	23	25.6
Geographical region	Anglia and Oxford (1)[1]	6	6.7
	North Thames (1)	1	1.1
	North West (2)	14	15.6
	Northern and Yorkshire (2)	11	12.2
	South Thames (1)	8	8.9
	South and West (3)	26	28.9
	Trent (4)	10	11.1
	Wales (5)	12	13.3
	West Midlands (4)	2	2.2
Vaccination status	No doses	47	52.2
	1 dose	14	15.6
	2 doses	3	3.3
	3 doses	19	21.1
	4 doses	1	1.1
	Not known[2]	6	6.7
Clinically typical symptoms	Yes	68	75.6
	No	15	16.6
	Not known	7	7.8
Complications	Yes	23	25.6
	-Conjunctivitis	1	1.1
	-Pneumonia	4	4.4
	-Apnoea	11	12.2
	-Pneumonia and apnoea	5	5.6
	-Not known	2	2.2
	No	58	64.4
	Not known	9	10.0
Hospitalisation	Yes	54	60.0
	No	26	28.9
	Not known	10	11.1
Mortality	Yes	3	3.3
	Not known	87	96.7

[1]Numbers in parantheses indicate the group in which the region appeared when performing multivariable logistic regression analysis.

[2]One of the six cases was vaccinated, but the number of doses was not entered. Therefore, it has been adjusted to an unknown case.

Associations between all epidemiological variables were initially assessed by cross tabulating them and calculating Pearson's χ^2 probability to test for independence. These results are shown in Table 3.5 (page 122), as P values. Most of the significant values (P<0.05) were expected (18/66 analyses), where three were associations due to explanatory variables. Explanatory variables are those that exist before the event, i.e., a pertussis infection occurs, and include age group, sex, region, vaccinated and number of vaccine doses. The outcome variables include typical clinical symptoms, complications (conjunctivitis, pnuemonia and apnoea), hospital admission and death. The geographical regions were further grouped into five larger areas for multivariable logistic regression analysis, due to the small number of cases seen (Table 3.4, page 120).

The vaccination status of the age cohorts was significant, as was the association between the number of vaccine doses received and the age cohorts (P<0.001). This is due to the shift in proportion of those unvaccinated from 100% (23/23) of under two months old, to 40% (8/20) of over five year olds, and is directly related to the UK vaccination schedule. The other association involving explanatory variables only, is that between sex and geographical region (P=0.005). When tabulated, it is apparent that the data are sparse and with no clear pattern (Table 3.6, page 123). Moreover, when this is re-examined with the geographical regions grouped into the five larger sets (Table 3.4, page 120), the association is not significant (P=0.17).

Due to the small number of cases that died (n=3), multivariable logistic regression analyses could not be meaningfully undertaken. Furthermore, all four significant P values were associated with another outcome variable (clinical symptoms, complications, pneumonia and apnoea, see Table 3.5, page 122). All three deaths were related to a complication (P=0.005), developed apnoea (P<0.001) and were admitted to hospital (P=0.221). Two of the three cases who died also exhibited no typically clinical symptoms (P=0.030) and developed pneumonia (P=0.02).

Table 3.5. Probability values of epidemiological associations of the 90 pertussis cases from 1998 to 1999, generated from Pearson's χ^2 analysis

	Age group	Sex	Region	Vaccinated	Number of vaccine doses	Clinical symptoms	Complications	Conjunctivitis	Pneumonia	Apnoea	Hospital admission
Died	0.633	0.687	0.853	0.119	0.677	**0.030**	**0.005**	0.840	**0.002**	**<0.001**	0.221
Hospital admission	**<0.001**	**0.026**	0.158	0.111	**0.007**	0.968	**0.032**	0.481	0.467	**0.015**	
Apnoea	**0.001**	0.236	0.056	**0.007**	0.115	0.531	–	0.618	**0.004**		
Pneumonia	0.087	0.477	**0.019**	0.562	0.956	**0.001**	–	0.722			
Conjunctivitis	0.576	0.261	0.667	0.389	0.946	0.621	–				
Complications	**0.005**	0.544	0.072	0.061	0.386	0.275					
Clinical symptoms	0.777	0.390	0.094	0.443	0.842						
Number of vaccine doses	**<0.001**	0.595	0.740	–							
Vaccinated	**<0.001**	0.967	0.914								
Region	0.807	**0.005**									
Sex	0.333										

Figures in bold denote significant probability values and are described in text.

Dashes indicate those variables which were directly related (e.g., vaccinated and number of vaccine doses), and, hence, no calculation was performed.

Table 3.6. Sex distribution of pertussis cases in the nine geographical regions

	Region									
Sex	Anglia and Oxford	North Thames	North West	Northern and Yorkshire	South and West	South Thames	Trent	Wales	West Midlands	Total
Female	5 (83)	0 (0)	4 (29)	9 (82)	20 (77)	2 (25)	4 (40)	5 (42)	2 (100)	51
Male	1 (17)	1 (100)	10 (71)	2 (18)	6 (23)	6 (75)	6 (60)	7 (58)	0 (0)	39
Total	6	1	14	11	26	8	10	12	2	90

Table 3.7. Distribution of typical pertussis symptoms through age cohorts

	Age Cohort					
Typical pertussis symptoms	<2 months	2-5 months	6-11 months	1-4 years	5+ years	Total
Yes	18 (82)	20 (77)	4 (100)	9 (90)	17 (81)	68 (82)
No	4 (18)	6 (23)	0 (0)	1 (10)	4 (19)	15 (18)
Total	22	26	4	10	21	83

Table 3.8. Association of pneumonia with pertussis infection through age cohorts

	Age Cohort					
Pneumonia	<2 months	2-5 months	6-11 months	1-4 years	5+ years	Total
Yes	6 (26)	1 (4)	0 (0)	0 (0)	2 (10)	9 (11)
No	17 (74)	23 (96)	3 (100)	10 (100)	19 (90)	72 (89)
Total	23	24	3	10	21	79

Figures in parentheses are percentages.

There was only one significant association involving those cases presenting with typical clinical symptoms of pertussis. This was with pneumonia (P=0.001), where 5/9 (56%) pneumonia cases did not present with typical clinical symptoms, compared to 10/72 (14%) with no pneumonia or clinical symptoms. Further investigations to determine whether clinical symptoms and pneumonia were dependent on age were performed (Table 3.7 and 3.8, page 123). Pneumonia cases and clinical symptoms were approximately inversely proportional in each age group, and therefore the association was found not to be dependent on age. The only borderline association involving typical clinical symptoms was with geographical regions (P=0.094), and a multivariable logistic regression analysis was therefore not necessary.

There was a significant association between those encountering a complication and being admitted to hospital (P=0.032), with 18/23 (78%) encountering complications being admitted compared to only 3/23 (13%) encountering complications not being admitted. Encountering a complication was also significant with age (P=0.005), where most of those complications were represented in the under two month age group (13/23, 57%). The significant explanatory variable, age, and the borderline variables region and vaccination status, as well as the variable sex, were included in a multivariable logistic regression analysis for encountering complications (Table 3.9, page 127). The sex and vaccination status variables showed no evidence of an association, but age and region did (P=0.052 and 0.008, respectively). When sex and vaccination status were dropped from a further multivariable analysis, age and region were both significant (P=0.003 and 0.014, respectively). Those under two months old are 5.88 more likely to encounter a complication following or during a pertussis infection (P=0.003). The regional data are harder to interpret, as there seems to be no pattern (north to south, or east to west) in the odds ratio values. Region 5, which contains Wales, is highly associated with no pertussis cases encountering a complication (10/10 in Wales, 100%).

The complication conjunctivitis, was seen in only one case in this sub-set of 90, and, consequently, there were no significant or borderline associations with any other factors (Table 3.5, page 122).

Pnuemonia was significantly associated with region (P=0.019), clinical symptoms (P= 0.001, discussed earlier), and with another complication, apnoea (P=0.004). As for the more general outcome of encountering a complication, the regional association with pneumonia is hard to interpret, and the low P value could be due to the sparseness of

cases. In this sub-set, 5/16 (31%) of apnoea cases also had pneumonia, compared to 11/16 (69%) apnoea cases not having pneumonia. Multivariable logisitic regression analysis was not performed using pneumonia as an outcome variable, as there were insufficient significant associations with the explanatory variables.

Analysis of apnoea cases revealed many significant associations, including being admitted to hospital, where 14/16 (88%) of apnoea cases were admitted (P=0.015). Significant associations of apnoea are also seen with age (P=0.001), vaccination status (P=0.007), and pneumonia (P=0.004, see above). The significant explanatory variables, age group and vaccination status, together with the borderline variables, sex, regions and number of vaccine doses, were used in multivariable logisitic regression analyses with apnoea. An analysis using apnoea, sex, age, region and vaccination status did not generate significant P values (Table 3.10, page 127). The under two month old group and the unvaccinated were more likely to encounter apnoea, and both gave similar P values and odds ratios (P=0.128 for age group and P=0.174 for vaccination status). When the variable vaccination status was replaced into the analysis with the related number of vaccine doses, the P values changed for each of the variables. The number of vaccine doses generated a higher P value of 0.285. After the variables with high P values (sex and vaccination status/number of vaccine doses) were dropped from the analysis, the age group P value dropped significantly to P=0.001. This also affected the odds ratio, so, where the under two month old group has a baseline of 1.00, there is only 0.12 (0.03 to 0.45 95% CI) odds that the over two months will encounter apnoea. The region data produced the lowest P value (0.073) in relation to apnoea, but, as for the more general outcome, complication, there was no pattern in the odds ratios produced for each region.

There were five significant associations with the outcome hospital admissions, although two of these were linked to the other outcome variables, complications and apnoea (P=0.032 and P=0.015, respectively) (Table 3.5, page 122). A linear trend analysis shows that the likelihood of hospital admission decreases as the uptake of vaccine, in terms of number of doses, increases (χ^2 for trend=8.54, degrees of freedom (DF)=1, P=0.004, Table 3.11, page 128), although this has not been adjusted for age. All of the explanatory variables were used in a multivariable logistic regression analysis to calculate odds ratios influencing cases being admitted hospital (Table 3.12, page 128). There is an unusual association between gender and hospital admission, and, though the P value is not significant in the regression analysis, females were nearly twice as likely to be admitted than males, with an odds ratio of 1.85 (95% CI: 0.74 to 23.14). The under two month old group

are significantly (P<0.001) more likely to be admitted to hospital than any other age group, as calculated by the very high odds ratio, and that 21/23 (91.3%) were admitted. As for the past analyses, the region data gave varied odds ratio values for cases admitted to hospital, but there is no pattern evident. Whilst the vaccination status was not significantly associated with hospital admission (P=0.453), those vaccinated are roughly half as likely to be admitted (odds ratio of 0.51, 95% CI: 0.09 to 3.00) than those unvaccinated (Table 3.12, page 128). However, when the more descriptive variable number of vaccine doses was switched for vaccination status, the P value increased to 0.802.

3.3.6 Laboratory typing results of the 90 UK cases from 1998 to 1999

Typing information generated in the laboratory include serotype, pertactin and pertussis toxin gene sequence type, and IS*1002*-RFLP type. Of the 90 cases, 49 (54.4%) were serotype 1,2; 47 (52.2%) were *prnA*(1); 41 (45.6%) were *prnA*(2); and 2 (2.2%) were *prnA*(3) (Table 3.13, page 129). Isolates from all cases were *ptxA*(1). Figure 3.8 (page 130) shows 18 different IS*1002*-RFLP typing patterns from the 90 cases from 1998 to 1999. There were between four and six IS*1002* bands present per RFLP type. The distribution of RFLP types, shown in Table 3.13 (page 129), shows that the majority of isolates were RFLP-1 (n=49, 54%), with the second most frequent type, RFLP-3, appearing in 14 isolates (15.7%). The remaining 16 RFLP types had low frequencies, of between 1 (1.1%) and 4 (4.5%) isolates.

A combined typing system was devised by the author using the two serotypes 1,2 and 1,3, the three pertactin types, *prnA*(1-3), and the 18 different RFLP types. As all 90 cases were of *ptxA*(1), this factor was not used to produce the combined types. The 24 combined types generated, labelled A to X, are listed in Table 3.14 (page 131). The three predominant combined types A (serotype 1,2, *prnA*(1), RFLP-1), C (serotype 1,2, *prnA*(1), RFLP-3), and P (serotype 1,3, *prnA*(2), RFLP-1), represented 16/90 (18%), 13/90 (14%), and 33/90 (37%) of the isolates, and the other 21 combined types contained fewer isolates, between 1 (1%) and 3 (3%).

Table 3.9. Factors influencing likelihood of encountering a complication with a pertussis infection

Factor	Level	Odds Ratio (95% CI)	P value
Sex	Male	1.00 (baseline)	0.580
	Female	1.41 (0.41 to 4.89)	
Age group[1]	<2 months	4.42 (0.94 to 20.83)	0.052
	>2 months	1.00 (baseline)	
Region[2]	Region 1	1.00 (baseline)	0.008
	Region 2	0.27 (0.05 to 1.59)	
	Region 3	0.77 (0.16 to 3.67)	
	Region 4	1.86 (0.29 to 11.81)	
	Region 5	0	
Vaccination status	Not vaccinated	1.00 (baseline)	0.690
	Vaccinated	0.73 (0.16 to 3.34)	

Table 3.10. Factors influencing likelihood of encountering apnoea with a pertussis infection

Factor	Level	Odds Ratio (95% CI)	P value
Sex	Male	1.00 (baseline)	0.920
	Female	1.08 (0.25 to 4.61)	
Age group[1]	<2 months	1.00 (baseline)	0.128
	>2 months	0.29 (0.06 to 1.49)	
Region[2]	Region 1	1.00 (baseline)	0.073
	Region 2	0.73 (0.08 to 6.90)	
	Region 3	2.54 (0.38 to 17.16)	
	Region 4	3.12 (0.34 to 28.38)	
	Region 5	0	
Vaccination status	Not vaccinated	1.00 (baseline)	0.174
	Vaccinated	0.26 (0.04 to 1.87)	

[1]The age groups have been sorted into two larger sub-sets, as a previous analysis using the five separate age groups show similar P values, and, the above analysis has not shown any bias when comparing the log likelihood of the multivariable logistic regression analysis.

[2]Details of the regional groups are shown in Table 3.4 (page 120).

Table 3.11. Association of hospital admissions with number of pertussis vaccine doses

Hospital admission	Number of vaccine doses					
	None	1	2	3	4	Total
No	11 (25)	1 (7)	1 (33)	10 (63)	1 (100)	24 (31)
Yes	33 (75)	13 (93)	2 (67)	6 (37)	0 (0)	54 (69)
Total	44	14	3	16	1	78

Univariable analysis of score test for trend of odds: $\chi^2(DF=1)=8.54$, $P=0.004$.

Figures in parentheses are percentages.

Table 3.12. Factors influencing likelihood of admission to hospital with a pertussis infection

Factor	Level	Odds Ratio (95% CI)	P value
Sex	Male	1.00 (baseline)	0.088
	Female	1.85 (0.74 to 23.14)	
Age group	<2 months	∞[1]	<0.001
	2-5 months	1.00 (baseline)	
	6-11 months	0.02 (0.01 to 0.71)	
	1-4 years	0.02 (0.01 to 0.24)	
	5+ years	0.03 (0.01 to 0.22)	
Region[2]	Region 1	1.00 (baseline)	0.241
	Region 2	2.72 (0.24 to 30.39)	
	Region 3	5.14 (0.38 to 69.67)	
	Region 4	∞[3]	
	Region 5	9.39 (0.66 to 133.61)	
Vaccination status	Not vaccinated	1.00 (baseline)	0.453
	Vaccinated	0.51 (0.09 to 3.00)	

[1]An infinity odds ratio was generated, as all under two months old (21/21) were admitted to hospital.

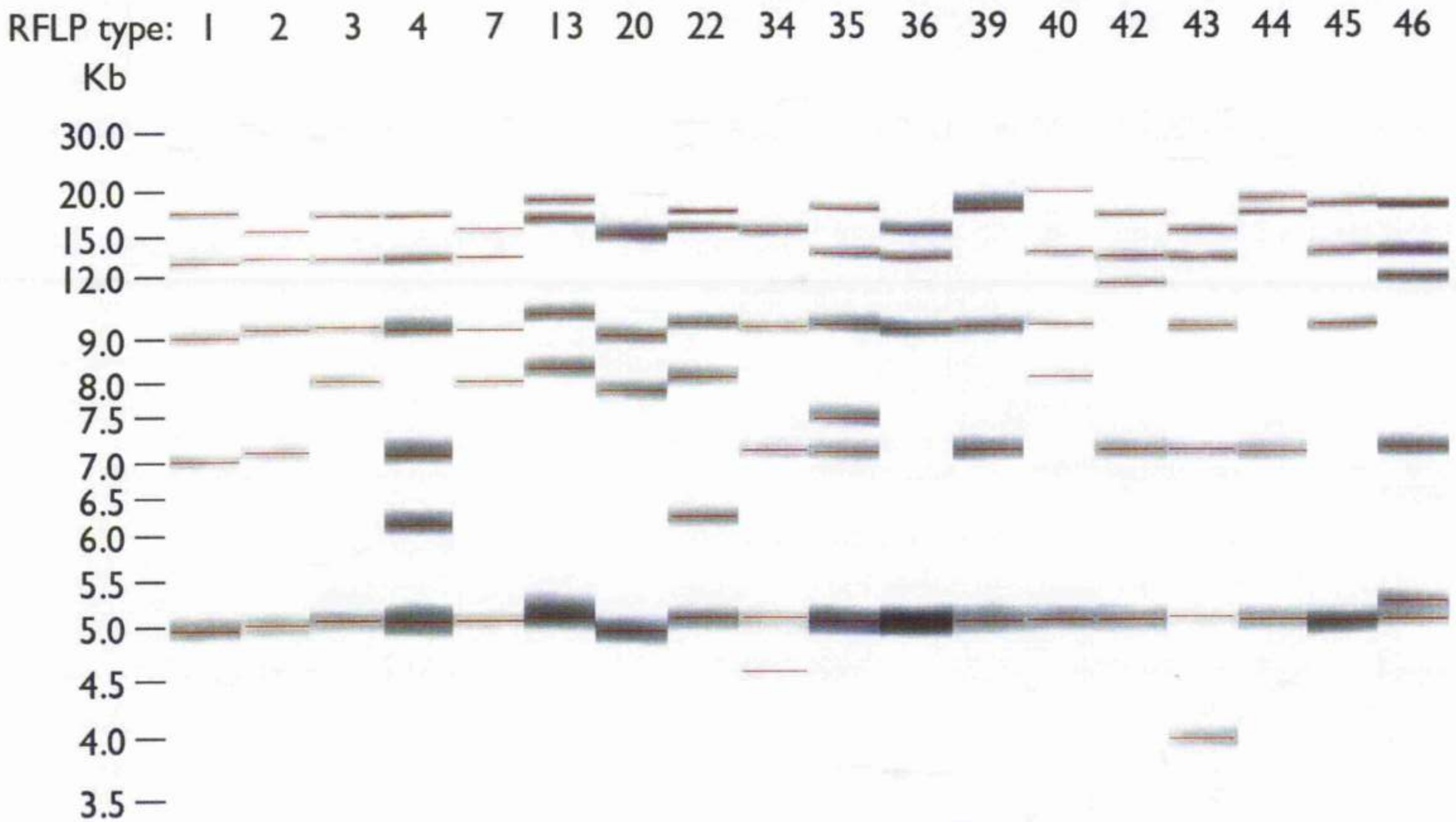
[2]Details of the regional groups are shown in Table 3.4 (page 120).

[3]An infinity odds ratio was generated, as 8/10 cases from Region 4 were admitted to hospital.

Table 3.13. The distribution of typing information of the 90 UK pertussis cases analysed from September 1998 to April 1999

Typing method	Type	Number of cases	Percentage of cases
Serotype	1,2	49	54.4
	1,3	41	45.6
Pertactin gene type	1	47	52.2
	2	41	45.6
	3	2	2.2
Pertussis toxin S1 gene type	1	90	100.0
IS*1002*-RFLP type	1	49	54.5
	2	3	3.3
	3	14	15.7
	4	3	3.3
	7	4	4.5
	13	1	1.1
	20	1	1.1
	22	2	2.2
	34	1	1.1
	35	1	1.1
	36	1	1.1
	39	3	3.3
	40	1	1.1
	42	2	2.2
	43	1	1.1
	44	1	1.1
	45	1	1.1
	46	1	1.1

Figure 3.8. IS*1002*-RFLP patterns of *Bordetella pertussis* isolates from 1998 to 1999 in the UK



Number of RFLP patterns=18. The thin red lines indicate bands included in the analysis.

Table 3.14. Combined types of *Bordetella pertussis* from 1998 to 1999 in the UK, and their corresponding serotype, pertactin gene type, and IS*1002*-RFLP type

Combined type	Serotype	Pertactin gene type	IS*1002*-RFLP type	Number of isolates (%)
A	1,2	1	1	16 (18)
B	1,2	1	2	2 (2)
C	1,2	1	3	13 (14)
D	1,2	2	3	1 (1)
E	1,2	1	4	2 (2)
F	1,2	1	7	3 (3)
G	1,2	1	13	1 (1)
H	1,2	1	20	1 (1)
I	1,2	1	22	2 (2)
J	1,2	2	35	1 (1)
K	1,2	1	39	2 (2)
L	1,2	1	42	2 (2)
M	1,2	1	43	1 (1)
N	1,2	1	44	1 (1)
O	1,2	1	46	1 (1)
P	1,3	2	1	33 (37)
Q	1,3	2	2	1 (1)
R	1,3	2	4	1 (1)
S	1,3	2	7	1 (1)
T	1,3	2	34	1 (1)
U	1,3	2	36	1 (1)
V	1,3	3	39	1 (1)
W	1,3	3	40	1 (1)
X	1,3	2	45	1 (1)
	Number of types			**Total isolates**
24	2	3	18	**90**

Analyses of associations between laboratory typing data and epidemiological data were performed using Pearson's χ^2 probability to test for independence, where the laboratory data were treated as an explanatory variable. Initial analyses were performed looking at the association between all explanatory variables: typing data, sex, age, geographical distribution, time period and vaccination status. The associations between the outcome variables (typical clinical symptoms, complications, conjunctivitis, pnuemonia and apnoea, hospital admission and death) and the typing data were then examined. Any significant associations found were then subjected to multivariable logistic regression analyses, which aims to determine the independent effect. Trends by age and time were examined in the laboratory typing data using the score test for the trend in odds.

When analysing the pertactin data, cases exhibiting *prnA*(3) isolates were omitted from calculations, as there were only two present. Similary, when using the IS*1002*-RFLP data and the combined types (serotype, pertactin type and IS*1002*-RFLP type combined), types containing less than four cases were grouped to ensure confidence when performing statistical tests. Therefore, for the RFLP data, cases were analysed using three groups, RFLP-1, RFLP-3 and the other RFLP types grouped as one type, "other RFLPs". The combined types were analysed using four groups, the combined types A, C, and P, and the remaining combined types clustered as "Z".

3.3.6.1 Associations of *Bordetella pertussis* serotypes in the 90 pertussis cases

Analyses with the other explanatory variables resulted in only two borderline significant associations with serotype: age (P=0.091) and vaccination status (P=0.128) (Table 3.15.a, page 133). When analysed as two age groups, the under 2 month old group exhibited 16 serotype 1,2 cases and seven serotype 1,3 cases. This was compared to the rest of the cases analysed (>2 months), where they were equally proportioned with 33 cases of serotype 1,2 and 34 isolates of serotype 1,3 (P=0.091). This shows some evidence that individuals under two months old were more likely to be infected with *B. pertussis* expressing serotype 1,2. Eighteen out of forty-eight (47%) of serotype 1,2 isolates and 20/37 (53%) of serotype 1,3 isolates were found in vaccinated cases compared to 30/48 (64%) of serotype 1,2 and 17/37 (36%) serotype 1,3 in unvaccinated cases (P=0.128). Age and vaccination status are however linked, where the under two month old group are not vaccinated.

Table 3.15. Probability values between serotypes of *Bordetella pertussis* isolates and associated factors

a. Single variable analysis

Factor	P value
Sex	0.598
Age group[1]	0.091
Region	0.775
Temporal trend[2]	0.851
Vaccination status	0.128
Vaccination doses	0.354
Typical pertussis symptoms	0.940
Complications	0.802
Conjunctivitis	0.356
Pneumonia	0.937
Apnoea	0.863
Hospital admission	0.017

b. Multivariable analysis with hospital admissions as the outcome

Factor	Level	Odds Ratio (95% CI)	P value
Sex	Male	1.00 (baseline)	0.083
	Female	4.67 (0.73 to 33.33)	
Age group	<2 months	∞[3]	<0.001
	2-5 months	1.00 (baseline)	
	6-11 months	0.02 (0.01 to 1.08)	
	1-4 years	0.01 (0.01 to 0.27)	
	5+ years	0.02 (0.01 to 0.24)	
Region[4]	Region 1	1.00 (baseline)	0.197
	Region 2	2.15 (0.18 to 26.05)	
	Region 3	6.14 (0.40 to 93.37)	
	Region 4	∞[5]	
	Region 5	9.10 (0.58 to 141.84)	
Vaccination status	Not vaccinated	1.00 (baseline)	0.425
	Vaccinated	0.48 (0.08 to 2.97)	
Serotype	Serotype 1,2	3.70 (0.69 to 20.00)	0.117
	Serotype 1,3	1.00 (baseline)	

[1]The age groups have been sorted into two larger sub-sets, as a previous analysis using the five separate age groups show similar P values, and, the above analysis has not shown any bias when comparing the log likelihood of the multivariable logistic regression analysis.

[2]Isolates are between September 1998 and April 1999, and are sorted into calendar months.

[3]An infinity odds ratio was generated, as all under two months old (21/21) were admitted to hospital.

[4]Details of the regional groups are shown in Table 3.4 (page 120).

[5]An infinity odds ratio was generated, as 8/10 cases from Region 4 were admitted to hospital.

Hospital admission was the only outcome variable that was significantly associated with the serotype of *B. pertussis* (P=0.017) (Table 3.15.a, page 133). Cases infected with isolates expressing serotype 1,2 were more likely to be admitted (34/54 serotype 1,2 isolates, compared to 20/54 serotype 1,3 isolates). Multivariable analysis was performed, with hospital admission as the outcome variable (Table 3.15.b, page 133). This analysis further revealed that females were over four times as likely to be admitted to hospital (P=0.083), and those infected with serotype 1,2 isolates were nearly four times as likely to be admitted to hospital than those with serotype 1,3 isolates (P=0.117).

3.3.6.2 Associations of *Bordetella pertussis* pertactin gene types in the 90 pertussis cases

The single variable analysis with the pertactin types against the other explanatory variables generated only one borderline significant association with age (P=0.109) (Table 3.16.a, page 136). When analysing the pertactin types and the proportion of isolates recovered from individuals under two months old, 15/47 (32%) of *prnA*(1) isolates were associated, compared with 7/41 (17%) of *prnA*(2) isolates and 1/2 (50%) of *prnA*(3) isolates.

As for the serotype data, the pertactin data showed a significant association with hospital admission (P=0.025). There are more *prnA*(1) and *prnA*(3) isolates associated with cases admitted to hospital than *prnA*(2) isolates (32/41 *prnA*(1) isolates (78%) and 2/2 *prnA*(3) isolates (100%), compared to 20/37 *prnA*(2) isolates (54%)). When included in a multivariable logistic regression analysis (Table 3.16.b, page 136), females are more than four times as likely to be admitted than males (P=0.083), the under two month old group are significantly more likely to be admitted (P<0.001), and cases are nearly four times as likey to be infected with *prnA*(1) isolates than *prnA*(2) isolates (P=0.118). The two *prnA*(3) isolates were excluded from the multivariable analysis, but they were associated with cases that were under three months old, exhibited typical symptoms of pertussis, developed the complication apnoea and were admitted to hospital.

3.3.6.3 Associations of *Bordetella pertussis* IS*1002*-Restriction Fragment Length Polymorphism types in the 90 pertussis cases

After the IS*1002*-RFLP data were sorted into three groups (RFLP-1, RFLP-3 and "other RFLPs"), single variable analysis indicated that only one significant association with the geographical regions (P=0.025) and a borderline association with the complication conjunctivitis (P=0.071) were present (Table 3.17, page 137). The significant association

may be due to the presence of RFLP-3 isolates in only the North West, Northern and Yorkshire, South Thames and South and West, as RFLP-1 and "other RFLPs" seem to be spread randomly throughout the regions (Table 3.18, page 137). There was only one conjunctivitis case seen in the 90 cases studied, which could make any association significant, as there are no null values. Nonetheless, the *B. pertussis* strain associated with this case was type RFLP-3. As there were no significant associations with any outcome variables, multivariable logistic regression analysis was not performed using RFLP types.

3.3.6.4 Associations of *Bordetella pertussis* combined types in the 90 pertussis cases

The combined data was analysed against other explanatory and outcome variables using the following four groups: combined types A, C, P and "Z" (the other 21 combined types). Combined type A showed a borderline association with the under two months old group (P=0.100), where there were 8/16 (50%) type A isolates associated with this age group, compared to 3/13 (23%) type C isolates, 6/33 (18%) type P isolates and 6/28 (21%) type "Z" isolates (Table 3.19, page 138). Combined type A is comprised of serotype 1,2, *prnA*(1), RFLP-1 isolates. Single variable analysis also generated a borderline significant association between the combined types and geographical regions (P=0.121). When tabulated (Table 3.20, page 138), combined type C was seen in the four regions mentioned above (3.3.6.3, page 134), and this could be due to type C consisting of RFLP-3 isolates, which may be the cause of this borderline probability. The outcome variable, complications, showed a borderline association with the combined types (P=0.125), where 7/13 (54%) of type A isolates were associated with complications, compared to 4/12 (33%) type C isolates, 7/30 (23%) type P isolates and 5/26 (19%) type "Z" isolates. As for the RFLP types, the complication conjunctivitis was associated with the combined types (P=0.121), but because there is only one case, multivariable logistic regression analyses could not be performed (Table 3.19, page 138).

Table 3.16. Probability values between pertactin gene types of *Bordetella pertussis* isolates and associated factors

a. Single variable analysis, including all three pertactin gene types present

Factor	P value
Sex	0.576
Age group[1]	0.109
Region	0.732
Temporal trend[2]	0.825
Vaccination status	0.266
Vaccination doses	0.542
Typical pertussis symptoms	0.581
Complications	0.670
Conjunctivitis	0.345
Pneumonia	0.577
Apnoea	0.742
Hospital admission	0.025

b. Multivariable analysis with hospital admissions as the outcome, excluding *prnA*(3)

Factor	Level	Odds Ratio (95% CI)	P value
Sex	Male	1.00 (baseline)	0.086
	Female	4.61 (0.71 to 30.30)	
Age group	<2 months	∞[3]	<0.001
	2-5 months	1.00 (baseline)	
	6-11 months	0.02 (0.01 to 1.08)	
	1-4 years	0.02 (0.01 to 0.28)	
	5+ years	0.02 (0.01 to 0.24)	
Region[4]	Region 1	1.00 (baseline)	0.200
	Region 2	2.16 (0.18 to 26.08)	
	Region 3	6.00 (0.39 to 91.66)	
	Region 4	∞[5]	
	Region 5	9.03 (0.58 to 140.46)	
Vaccination status	Not vaccinated	1.00 (baseline)	0.412
	Vaccinated	0.47 (0.08 to 2.90)	
Pertactin gene type	*prnA*(1)	3.70 (0.69 to 20.00)	0.118
	prnA(2)	1.00 (baseline)	

[1]The age groups have been sorted into two larger sub-sets, as a previous analysis using the five separate age groups show similar P values, and, the above analysis has not shown any bias when comparing the log likelihood of the multivariable logistic regression analysis.

[2]Isolates are between September 1998 and April 1999, and are sorted into calendar months.

[3]An infinity odds ratio was generated, as all under two months old (21/21) were admitted to hospital.

[4]Details of the regional groups are shown in Table 3.4 (page 120).

[5]An infinity odds ratio was generated, as 8/10 cases from Region 4 were admitted to hospital.

Table 3.17. Probability values between IS*1002*-RFLP types of *Bordetella pertussis* isolates and associated factors

Factor	P value
Sex	0.826
Age group[1]	0.605
Region	0.025
Temporal trend[2]	0.920
Vaccination status	0.713
Vaccination doses	0.565
Typical pertussis symptoms	0.385
Complications	0.530
Conjunctivitis	0.071
Pneumonia	0.780
Apnoea	0.839
Hospital admission	0.355

[1]The age groups have been sorted into two larger sub-sets, as a previous analysis using the five separate age groups show similar P values, and, the above analysis has not shown any bias when comparing the log likelihood of the multivariable logistic regression analysis.

[2]Isolates are between September 1998 and April 1999, and are sorted into calendar months.

Table 3.18. Geographical distribution of IS*1002*-RFLP types

Region	RFLP-1	RFLP-3	"other" RFLP	Total
Anglia and Oxford	3 (50)	0 (0)	3 (50)	6
North Thames	1 (100)	0 (0)	0 (0)	1
North West	4 (29)	6 (43)	4 (29)	14
Northern and Yorkshire	2 (18)	3 (27)	6 (55)	11
South Thames	3 (38)	1 (12)	4 (50)	8
South and West	16 (62)	4 (15)	6 (23)	26
Trent	9 (90)	0 (0)	1 (10)	10
Wales	9 (75)	0 (0)	3 (25)	12
West Midlands	2 (100)	0 (0)	0 (0)	2
Total	49 (54)	14 (16)	27 (30)	90

Figures in parantheses indicate percentages according to region distribution.

Table 3.19. Probability values between combined types of *Bordetella pertussis* isolates and associated factors

Factor	P value
Sex	0.834
Age group[1]	0.100
Region	0.121
Temporal trend[2]	0.867
Vaccination status	0.289
Vaccination doses	0.266
Typical pertussis symptoms	0.483
Complications	0.125
Conjunctivitis	0.121
Pneumonia	0.367
Apnoea	0.638
Hospital admission	0.221

[1]The age groups have been sorted into two larger sub-sets, as a previous analysis using the five separate age groups show similar P values, and, the above analysis has not shown any bias when comparing the log likelihood of the multivariable logistic regression analysis.

[2]Isolates are between September 1998 and April 1999, and are sorted into calendar months.

Table 3.20. Geographical distribution of combined types

Region	A	C	P	"Z"	Total
Anglia and Oxford	1 (17)	0 (0)	2 (33)	3 (50)	6
North Thames	0 (0)	0 (0)	1 (100)	0 (0)	1
North West	0 (0)	5 (36)	4 (28)	5 (36)	14
Northern and Yorkshire	1 (9)	3 (27)	1 (9)	6 (55)	11
South Thames	1 (13)	1 (13)	2 (24)	4 (50)	8
South and West	4 (15)	4 (15)	12 (47)	6 (23)	26
Trent	3 (30)	0 (0)	6 (60)	1 (10)	10
Wales	5 (42)	0 (0)	4 (33)	3 (25)	12
West Midlands	1 (50)	0 (0)	1 (50)	0 (0)	2
Total	16 (18)	13 (14)	33 (37)	28 (31)	90

Figures in parantheses indicate percentages according to region distribution.

3.3.6.5 Age cohorts trend analysis

The distribution of serotype and pertactin gene type in each age group was investigated and showed a similar trend (Tables 3.21 and 3.22, page 140). There were 15/23 (65%) of babies under two months old infected with *B. pertussis* whose isolates were serotype 1,2 and/or *prnA*(1), and this proportion gradually decreased to 1/4 (25%) in the six to 11 month old babies. The proportion of serotype 1,2 and *prnA*(1) isolates then rose to 11/23 (48%) of the individuals in the five year old and above age group. There was, incidently, a very significant association between serotype 1,2 with *prnA*(1) (47/90 isolates, 52%), and serotype 1,3 with *prnA*(2) (39/90, 43%) ($P<0.001$). The significance of a linear trend was borderline through the age groups for serotype (χ^2 for trend=3.15, P=0.076, Table 3.21, page 140), and for pertactin, when the *prnA*(3) cases were removed, (χ^2 for trend=2.64, P=0.104, Table 3.22, page 140).

3.3.6.6 Temporal trend analysis

The isolation dates were demoted into the month of isolation, and resulted in eight consecutive months, between September 1998 and April 1999. The distributions of serotypes and pertactin types showed a shift between February 1999 and March 1999 (Tables 3.23 and 3.24, respectively, page 141). The serotype distribution shifted from 4/8 (50%) to 11/18 (61%) serotype 1,2 isolates from September to February, to 5/13 (38%) in March. The calculated odds ratios showed an increased proportion of serotype 1,3 in March with the only value higher than 1 (1.60), but the temporal trend of this shift was not significant (χ^2 for trend=0.46, P=0.497, Table 3.23, page 140). In addition, only 5% of isolates collected after 1999, and not included in this analysis were serotype 1,2 (1 isolate of serotype 1,2 and 21 isolates of serotype 1,3).

The distribution of pertactin gene types has also changed temporally, from 4/8 (50%) to 10/18 (62%) *prnA*(1) isolates in September to February, to 5/13 (38%) *prnA*(1) in March, raising the frequency of *prnA*(2) to 8/13 (62%) (Table 3.24, page 141). To perform the linear trend analysis, *prnA*(3) isolates had to be omitted, and *prnA*(1) and *prnA*(2) formatted as binary numbers. The shift between *prnA*(1) and *prnA*(2) however, was also not significant (χ^2 for trend=0.35, DF=1, P=0.554). The 23 isolates collected between 2001 and 2002 consisted of 21 isolates possessing *prnA*(2), and one of each *prnA*(1) and *prnA*(3), indicating a later shift towards *prnA*(2) variants.

Table 3.21 Distribution of *Bordetella pertussis* serotype through age cohorts

	Age Cohort					
Serotype	**<2 months**	**2-5 months**	**6-11 months**	**1-4 years**	**5+ years**	**Total**
1,2	16 (70)	17 (59)	1 (25)	4 (36)	11 (48)	49 (54)
1,3	7 (30)	12 (41)	3 (75)	7 (64)	12 (52)	41 (46)
Total	23	29	4	11	23	90

Score test for trend of odds: χ^2(DF=1)=3.15, P=0.076.

Table 3.22. Distribution of *Bordetella pertussis* pertactin gene type through age cohorts

	Age Cohort					
Pertactin type	**<2 months**	**2-5 months**	**6-11 months**	**1-4 years**	**5+ years**	**Total**
***prnA*(1)**	15 (65)	16 (55)	1 (25)	4 (36)	11 (48)	47 (52)
***prnA*(2)**	7 (31)	12 (41)	3 (75)	7 (64)	12 (52)	41 (46)
***prnA*(3)**	1 (4)	1 (4)	0 (0)	0 (0)	0 (0)	2 (2)
Total	23	29	4	11	23	90

Score test for trend of odds: χ^2(DF=1)=2.64, P=0.104. This trend analysis was performed without *prnA*(3) isolates (n=2)

Figures in parentheses are percentages.

Table 3.23 Distribution of serotypes in UK *Bordetella pertussis* isolates from September 1998 to April 1999

Serotype	Month								
	September	October	November	December	January	February	March	April	Total
serotype 1,2	4 (50)	8 (57)	6 (60)	8 (62)	7 (54)	11 (61)	5 (38)	0 (0)	49 (54)
serotype 1,3	4 (50)	6 (43)	4 (40)	5 (38)	6 (46)	7 (39)	8 (62)	1 (100)	41 (46)
Total isolates	8	14	10	13	13	18	13	1	90
Odds ratio (95% CI)	1.00 (0.25 to 4.00)	0.75 (0.26 to 2.16)	0.67 (0.19 to 2.36)	0.63 (0.20 to 1.91)	0.86 (0.29 to 2.55)	0.64 (0.25 to 1.64)	1.60 (0.52 to 4.89)	-	-

Univariate analysis score test for trend of odds: $\chi^2(1)=0.46$, $P=0.497$.

Table 3.24 Distribution of pertactin types in UK *Bordetella pertussis* isolates from September 1998 to April 1999

Pertactin type	Month								
	September	October	November	December	January	February	March	April	Total
prnA(1)	4 (50)	7 (50)	6 (60)	8 (62)	7 (54)	10 (56)	5 (38)	0 (0)	47 (52)
prnA(2)	4 (50)	7 (50)	3 (30)	5 (38)	6 (46)	7 (39)	8 (62)	1 (100)	41 (46)
prnA(3)	0 (0)	0 (0)	1 (10)	0 (0)	0 (0)	1 (5)	0 (0)	0 (0)	2 (2)
Total isolates	8	14	10	13	13	18	13	1	90
Odds ratio (95% CI)	1.00 (0.25 to 4.00)	1.00 (0.35 to 2.85)	0.50 (0.13 to 2.00)	0.63 (0.20 to 1.91)	0.86 (0.29 to 2.55)	0.70 (0.27 to 1.84)	1.60 (0.52 to 4.89)	-	-

Univariate analysis score test for trend of odds: $\chi^2(1)=0.35$, $P=0.554$.

Figures in parentheses are percentages.

These temporal shifts in serotype and pertactin distributions coincide, and were probably due to the association between serotype and pertactin type, described in 3.3.6.5 (page 139).

3.4 Discussion

Following the recent resurgence of pertussis in countries such as the USA, Canada, Australia and The Netherlands (De Serres *et al.*, 1995; Andrews *et al.*, 1997; de Melker *et al.*, 1997; Guris *et al.*, 1999b), the genotypic diversity and population dynamics of *Bordetella pertussis* in the United Kingdom was investigated. To detect any changes in the *B. pertussis* population, strains dating from 1920 to 2002 were collected from a number of sources. Techniques which have been published previously were initially chosen for this study so that direct comparisons could be attempted with studies performed in other countries. These methods included serotyping, sequence typing of the genes coding for pertactin and pertussis toxin subunit 1 and IS*1002*-RFLP analysis (Preston, 1963; van der Zee *et al.*, 1996b; Mooi *et al.*, 1998). In the UK, only serotyping and PFGE have been used previously to study the population dynamics of *B. pertussis*. Serotyping has a limited use due to its low discriminatory power and the ability of the organisms to switch at a low frequency between serotype 1,2, 1,3 and 1,2,3 (Preston, 1988; Roberts and Parton, 2001). Pulsed-field gel electrophoresis of *B. pertussis* has had limited use in the UK and is mostly a tool for epidemiological investigations (Khattak *et al.*, 1992; Syedabubakar *et al.*, 1995). Also available were epidemiological data for a sub-set of isolates to elucidate any associations with the typing results generated.

3.4.1 Variation in the pertactin gene

The six UK clinical isolates and three Wellcome WCV strains studied for any polymorphisms outside of the polymorphic regions 1 and 2, showed no variation in over 90% of the pertactin gene (3000 base pairs) for the types that Mooi *et al.* (1998) had described. Since work on this thesis began, ten *prnA* types have been described, and the sequences can be accessed from public sequence databases, e.g., GenBank. An alignment is illustrated in Figure 1.2 (page 16). Most differences occur in region 1 of the pertactin gene, varying either in the composition or the number of GGxxP-encoded repeats, or a combination of both. However, two pertactin types, *prnA*(1) and *prnA*(7), are identical at region 1, and differ only by a point mutation approximately 150 nucleotides upstream of region 2. Therefore, 68/214 (32%) of isolates identified as *prnA*(1) by sequencing region 1, were also sequenced upstream of region 2, and were all found to be designated as *prnA*(1)

isolates. To date, *prnA*(7) has been isolated only from a strain used in some vaccine preparations and in earlier strains from The Netherlands (van Loo and Mooi, 2002).

There were three pertactin types found in the 318 UK *B. pertussis* isolates sequenced between 1920 and 2002. Between 1920 to 1979, all isolates were *prnA*(1), which is the same type as in the strains used in both the Wellcome and the Aventis Pasteur WCVs. From 1982 however, new variants *prnA*(2) and *prnA*(3) emerged and their proportions gradually increased within the circulating population so that in 2002, 20 of the 22 isolates sequenced were *prnA*(2). Although the data-set was large (n=318), there was some sampling bias due to the difficulty in obtaining earlier isolates. Consequently, this study utilised 51 isolates from the 1940s when pertussis incidence was very high (between 53,000 to 173,000 cases per year), compared to 106 isolates from the 1990s when incidence was low (between 1000 to 15,000 cases per year). Therefore the sampling fraction was a lot lower for earlier strains than later ones, and this could have possibly led to a skewed distribution of isolates tested. The distribution of pertactin types when compared both within and between the different time periods, however, remained statistically valid.

Since this work began in the UK, studies from Finland, Italy, the USA, Japan, France, Poland, Belgium, Canada and Russia, on pertactin and pertussis toxin gene variation, have been published (Mooi *et al.*, 1999; Mastrantonio *et al.*, 1999; Cassiday *et al.*, 2000; Guiso *et al.*, 2001a; Weber *et al.*, 2001; Gzyl *et al.*, 2002; de Schutter *et al.*, 2003; Peppler *et al.*, 2003; Kourova *et al.*, 2003). Some of the data from the following countries were limited in scope; (i) Japan, as only 12 clinical isolates were investigated from 1975 to 1996 (Guiso *et al.*, 2001a); (ii) Belgium, as this was a household contact study involving only ten families (de Schutter *et al.*, 2003); (iii) Canada, because the isolates analysed were from 1985 to 1994 (Peppler *et al.*, 2003); and (iv) Russia, as the number of strains available was limited (n=61) and not all of those had been investigated by pertactin and pertussis toxin S1 gene typing (Kourova *et al.*, 2003). Results from the other countries mentioned above, including The Netherlands, Finland and Italy, concur with the original hypothesis of Frits Mooi *et al.*;

> "… vaccination has selected for [*B. pertussis*] strains which are antigenically distinct from vaccine strains…".
>
> (Mooi *et al.*, 1998)

The pertactin data presented here for 318 UK isolates also appear to be concordant with this theory, but like Finland, Italy and Japan, the UK has not seen a resurgence in pertussis cases. Mooi and colleagues further suggested in their paper describing isolates in Finland that;

"… pertussis vaccines protect less well against strains with particular *prn*[*A*] alleles, such as *prn*[*A*](3), which predominate in some countries, such as The Netherlands, but are less common in Finland… ".

(Mooi *et al.*, 1999)

This statement modifies their original theory of vaccine-driven evolution by stipulating that there is a possible threshold of certain pertactin types that could be changed by vaccinating against pertussis. Furthermore, if this threshold of 'fitter' pertactin types is breached, then pertussis resurgence and ensuing epidemics could occur. This hypothesis could also be applicable to Australia, where a recent pertussis resurgence has been reported, along with a shift in pertactin types, with *prnA*(3) levels between 1989 and 1998 at 42% of the circulating *B. pertussis* population (Poynten *et al.*, 2004). However, results from the USA and Poland, which have also seen a resurgence of pertussis, have not encountered any *prnA*(3) isolates in their studies (Cassiday *et al.*, 2000; Gzyl *et al.*, 2002). In contrast, in Canada, where there has been a resurgence of pertussis, 166/3334 (5%) of isolates between 1985 and 1994 were *prnA*(3) (Peppler *et al.*, 2003). In the UK, however, *prnA*(3) comprised 4/133 (3%) of those isolates from 1998 to 2002, and there has been no pertussis resurgence.

Pertactin results from countries such as the UK and Finland, with a low incidence of pertussis, and from the high incidence countries such as USA and Poland, are concordant with the general theory of vaccine-driven evolution, but are not consistent with the suggestion that vaccines are less protective against certain *prnA* alleles. Further evidence from intranasally-challenged mouse models and a convalescent antibody study does suggest that there may be particular pertactin types that are less immunogenic than others (King *et al.*, 2001; He *et al.*, 2003). King and colleagues (2001) found that, after immunising mice with the Dutch WCV, which contains only *prnA*(1) strains, the *prnA*(2) and *prnA*(3) isolates were significantly more effective at colonisation than *prnA*(1) isolates. Humans previously infected with *prnA*(2) isolates had significantly lower levels of convalescent antibodies to *prnA*(1) than subjects infected with *prnA*(3) isolates, and those immunised with a booster ACV containing *prnA*(1) (He *et al.*, 2003). The studies above therefore indicate that non-vaccine types would be more likely to colonise and cause infection than the vaccine types. This appears to be the situation in The Netherlands, where an increase in pertussis notifications has been considered to be due to the increase in non-vaccine types (Mooi *et al.*, 1998). In the UK, 77/133 (58%) of *prnA*(2) isolates were circulating between 1998 and 2002, but there has been no increase in pertussis notifications. However, this low pertussis incidence in the UK can not be easily explained by a good vaccine efficacy and coverage, as

countries like France, who have good coverage and efficacy, have seen a resurgence in pertussis (Therre and Baron, 2000).

3.4.2 Variation in the pertussis toxin S1 gene

There have been, at the time of writing, six variants of the pertussis toxin S1 subunit gene sequence of *B. pertussis* deposited in GenBank (Figure 1.4, page 48). They differ from each other by single nucleotide polymorphisms (SNPs) at multiple sites, which can either affect the translated amino acid (non-silent) or not (silent). In the UK, two of these six *ptxA* types were found by sequencing 318 isolates from 1920-2002. Both *ptxA*(1) and *ptxA*(2) were present in strains used for the UK Wellcome WCV, and these were also circulating in strains up to 1985. In The Netherlands, where *ptxA*(2) and *ptxA*(3) strains are included in their WCV, *ptxA*(1) was considered a new variant, and was first seen there in the 1980s (Mooi *et al.*, 1998). This was also the situation in Finland (Mooi *et al.*, 1999), but, later studies in the USA, Poland and Australia revealed their first *ptxA*(1) isolates in 1970, 1960 and 1967, respectively (Cassiday *et al.*, 2000; Gzyl *et al.*, 2002, Poynten *et al.*, 2004). The first *ptxA*(1) isolates in the UK were found in 1940, clearly demonstrating that this variant was not a new emerging one, and may have had some geographical restriction in earlier years. Furthermore, *ptxA*(1) was present in the UK Wellcome WCV, and is also found in the Australian WCV (Poynten *et al.*, 2004).

Recent isolates from 1998 to 2002 in the UK showed a 100% frequency of the *ptxA*(1) type, with the disappearance of *ptxA*(2) isolates. This shift to predominantly *ptxA*(1) has been seen recently for all countries investigated. For many of these countries, strains with *ptxA*(1) are not present in their vaccine preparations and, therefore, vaccine-driven evolution seems a logical theory to explain pertussis resurgence. However, Australia has seen an increase in pertussis cases despite circulating strains being of the same *ptxA* type as the vaccine (Poynten *et al.*, 2004). The authors suggested that the change in pertactin types to 24/57 (42%) of *prnA*(3) isolates from 1989 to 1998, and a waning vaccine-induced immunity in older children and adults, may be factors in the resurgence of pertussis in Australia (Poynten *et al.*, 2004). In the UK, a large number of isolates (n=318) has been investigated for the pertussis toxin S1 genotype. The type, *ptxA*(1), present in the Wellcome vaccine, matches the circulating isolates up to 1996, and this argues against the theory of vaccine-driven evolution. However, since 1996, the UK has predominantly been using the WCV manufactured by Aventis Pasteur, which contains two strains that are *ptxA*(2) and *ptxA*(3), thus now showing a mis-match between the vaccine strains and the

circulating, recent *B. pertussis* population. The UK, nevertheless, has not seen a resurgence in pertussis cases, as reflected in the low number of clinical pertussis cases notified in 2003 (n=409).

3.4.3 Combined pertactin and pertussis toxin S1 types in the UK

The combination of pertactin and pertussis toxin S1 typing data for each individual isolate reveals a more representative picture of *B. pertussis* population dynamics and major shifts of important genotypes through time. Three *prnA* and two *ptxA* types were found in 318 isolates. This should, in theory, generate six possible combinations of *prnA/ptxA* types, but only four were seen in the UK. Similar combined results from the USA revealed five *prnA/ptxA* types out of a possible six (Cassiday *et al.*, 2000). These authors also labelled certain *prnA/ptxA* alleles as either "old", "new", or "transitional" due to the temporal shift of types (Cassiday *et al.*, 2000), indicating an association between particular *prnA* and *ptxA* types. However, when the USA data are presented in three time periods spanning 1935-1999, "transitional" and "old" *prnA/ptxA* types are seen throughout the years. If this labelling was applied to the UK data, then the "transitional" type *prnA*(1)/*ptxA*(1) (bearing an "old" *prnA* and a "new" *ptxA*) was seen from 1941 to 2002. Furthermore, in Canada this type was seen in 572/3334 (17%) of the isolates from 1985 to 1994 (Peppler *et al.*, 2003). Therefore, this "transitional" *prnA/ptxA* combination should be viewed as a successful genotype of *B. pertussis*, and up until 2000, *prnA*(1)/*ptxA*(1) has been present in the majority of isolates tested in the UK.

The UK Wellcome WCV contained the combined types *prnA*(1)/*ptxA*(1) from two strains and *prnA*(1)/*ptxA*(2) from one other. Although the proportion of type *prnA*(1)/*ptxA*(2) has decreased in recent circulating isolates, the persistance of type *prnA*(1)/*ptxA*(1) may have some influence on the efficacy of the UK Wellcome WCV. When studying the UK data by decades, all four of the *prnA/ptxA* combinations were identified in the 1980s. This evidently coincides with the emergence of new *prnA* variants *prnA*(2) and *prnA*(3) at this time, generating two of the four *prnA/ptxA* combinations. However, in the ensuing period of the 1990s, the disappearance of *ptxA*(2) from circulating strains drops the variablility to three *prnA/ptxA* types. Although there was a change of WCV manufacturer since 1996 to the Aventis Pasteur, which contains *prnA*(1)/*ptxA*(2) and *prnA*(1)/*ptxA*(3) types, the last *ptxA*(2) circulating strain was seen in 1985 and therefore before the Aventis Pasteur WCV was introduced to the UK.

3.4.4 Genotypic diversity of UK *Bordetella pertussis* between 1940 and 1999

The pertactin and pertussis toxin S1 sequencing data, as for serotyping data, did not generate enough discrimination in order to deduce whether there had been any changes in the genotypic diversity of *B. pertussis* isolates in the UK, although the data revealed some clear temporal shifts in genotypes. Therefore, IS*1002*-RFLP analysis was also used and, as this method was used previously in The Netherlands, any genotypic diversity values calculated from the UK RFLP data could be compared directly with these studies (van der Zee *et al.*, 1996b; van Loo *et al.*, 1999).

Of the 291 *B. pertussis* isolates from 1940 to 1999 studied for pertactin and pertussis toxin S1 subunit gene polymorphism, 149 were chosen at random for IS*1002*-RFLP analysis and 24 different RFLP types were generated. The difference of 24 RFLP types in 149 isolates, compared to just three *prnA* types in 318 isolates, reveals that the variability of the *B. pertussis* genome is greater when studying whole-genome profiles compared to single gene variations.

Assuming that each hybridised band contains one copy of the IS*1002* element, the RFLP types exhibited between four and six bands and therefore four to six IS*1002* elements were present in each *B. pertussis* strain (van der Zee *et al.*, 1996b). However, this assumption may not necessarily be true, as the probe is only 293 bp and the element in *B. parapertussis* is 1040 bp (van der Zee *et al.*, 1996b). The band sizes generated in IS*1002*-RFLP vary from 2096 bp to 23.3 Kb, where two or more copies of the IS*1002* element can be theoretically present. However, the intensity of the bands is consistent throughout each profile, suggesting that there is only one copy of IS*1002* in each hybridised band. The *in silico* digest of the *B. pertussis* Tohama-I genome (see 5.3.6, page 202), together with the annotated Tohama-I genome, indicates that there are six copies of the IS*1002* element, which agrees with experimental observations. In the UK, there was a statistically significant reduction in the number of IS*1002*-containing bands through time. An isolate from 1920, NCTC 10901, which was not included in this analysis, contained nine IS*1002*-containing bands. This could indicate that this particular IS element neither has any benefit or detriment to the organism, and could therefore afford to lose the elements from its' genome over time. In contrast, the IS*481* element, present in the Tohama-I genome at 238 positions, is considered to be the catalyst for multiple recombination events which resulted in the rearrangement and reduction of the chromosome when compared to the *B. bronchiseptica* genome (Parkhill *et al.*, 2003). *Bordetella pertussis* carries at least four different IS elements

(Table 1.4, page 41, and as demonstrated by the annotated Tohama-I genome), and the influence of each one on the genome could be different.

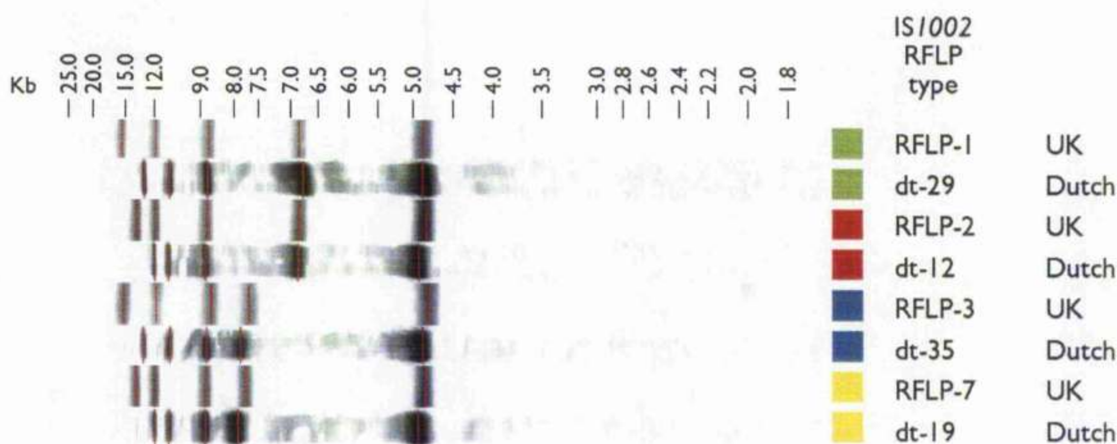
When the isolates were sorted by their RFLP type and divided into decades, there were more RFLP types present in the 1940s and 1980s than in the other decades examined. This is similar to the Dutch data, where 12 RFLP types were found in 1949 to 1954 and 11 types between 1982 and 1990 (van Loo *et al.*, 1999). However, the subsequent period of 1991 to 1996 contained 22 types, whereas in the UK only six RFLP types were seen in the 1990s. This difference may have been due to the number of isolates examined, as only 36 UK isolates were tested (0.17 types per isolate), compared to 120 isolates from The Netherlands (0.18 types per isolate), therefore revealing more RFLP types (van Loo *et al.*, 1999). This can be compared to data reported in 3.3.5 (page 114), where 18 RFLP types were seen in 90 UK isolates from September 1998 to April 1999 (0.2 types per isolate).

Unfortunately the designation of IS*1002*-RFLP profiles that the laboratory in The Netherlands use could not be applied to my data. This was because of several reasons; (i) technical differences due to an amended protocol, as the Dutch separated the restricted fragments for 16 hours (20 hours in the UK) and used chemiluminescence for detecting the hybridised bands (digoxygenin in the UK), the size of the membrane used for data analysis, and the gel analysis software used (with any user-defined parameters); (ii) the reference strains obtained from The Netherlands enabled comparison of the Dutch DNA types with the results generated in this study, but was limited by a manual, subjective comparison; and (iii) to use the Dutch designation would require a reference strain for every RFLP profile observed (n>80 types). Therefore, all the profiles observed in the present study were given an arbitrary number, and all the isolates designated appropriately. However, the major RFLP types elucidated in this study could be matched to the DNA types designated by (van Loo *et al.*, 1999) and used for comparison (Table 3.25 and Figure 3.9, page 149).

Table 3.25 Corresponding *Bordetella pertussis* IS*1002*-RFLP profiles between the UK and The Netherlands

UK designation	Dutch designation
RFLP-1	dt-29
RFLP-2	dt-12
RFLP-3	dt-35
RFLP-4	dt-28
RFLP-6	dt-51
RFLP-7	dt-19

Figure 3.9 *Bordetella pertussis* IS*1002*-RFLP profiles from the UK and The Netherlands



Distances between bands from corresponding profiles vary due to differences in electrophoresis running conditions, the area analysed, and normalisation. The thin red lines indicate bands included in the analysis, and the pairs of coloured squares group the corresponding UK and Dutch RFLP types.

The Dutch IS*1002*-RFLP types dt-28 and dt-51, which correspond to the UK types RFLP-4 and RFLP-6, could not be represented. Please refer to van Loo *et al.* (1999).

The most abundant type in the UK, RFLP-1 (dt-29), was also the prevalent type in The Netherlands, Italy, Australia and in Missouri, USA (van Loo *et al.*, 1999; Mastrantonio *et al.*, 1999, Poynten *et al.*, 2004, and 4.3.3, page 174). The second most abundant type RFLP-2 (dt-12) was also seen in the above countries, and is identical to RFLP type D1, described by van der Zee *et al.* (1996b). In their study, strains collected from Germany, China, the USA and Canada also revealed this RFLP type D1 (RFLP-2), indicating that this RFLP type has a global distribution. When analysed by decades, RFLP-2 was seen in the UK from the 1940s to the 1990s, and in The Netherlands from 1949 to 1996, also revealing the persistence over time of this succesful RFLP type.

The RFLP types of the strains used in the UK Wellcome WCV are RFLP-1, RFLP-2 and RFLP-3. Unfortunately, the strains from the Aventis Pasteur WCV were not available and therefore the RFLP types of the two strains could not be determined. However, the RFLP types in the UK Wellcome WCV correspond to the Dutch types dt-29, dt-12 and dt-35, and have been found in the circulating population of both countries. These three types were seen in 83/120 (69%) and 32/36 (89%) of the recent isolates tested in The Netherlands and the UK, respectively (van Loo *et al.*, 1999). They were also observed in 56/76 (74%) of isolates from Italy collected between 1993 and 1995 (Mastrantonio *et al.*, 1999), indicating that RFLP-1, RFLP-2 and RFLP-3 are successful lineages of *B. pertussis*. Although the UK Wellcome WCV RFLP types match with a large proportion of the UK circulating population, the Dutch WCV types, dt-19 (UK RFLP-7) and dt-43 (not seen in the UK) have not been seen in great numbers in the Dutch pertussis population (van Loo *et al.*, 1999). Partly due to this mis-match, Mooi and colleagues searched for antigenic variation within the *B. pertussis* population, by sequencing the pertactin and pertussis toxin S1 genes (Mooi *et al.*, 1998).

By using the UK pertussis vaccination coverage figures and notification rates from the 1940s onward, a concurrent association can be made between these, the RFLP data, and the genotypic diversity values. In the 1940s, before vaccination was introduced and notifications were high, there were many RFLP types (n=11/24), and a high GD of 0.81. After vaccination was introduced in the 1950s, the number of RFLP types and the GD dropped to six types and a GD of 0.64, respectively, and this change was calculated to be of borderline significance, probably due to the decreasing number of pertussis cases, and the elimination or control of less-fit strains. Interestingly, the proportion of RFLP-1 isolates increased significantly between the 1940s and the 1950s, indicating that RFLP-1 was a successful lineage and was still circulating after vaccination was introduced. Mass

vaccination was implemented in 1957, and through the 1960s and 1970s cases of pertussis dropped markedly and vaccine coverage reached approximately 80% (Health Protection Agency, 2003c). In these two decades, although the number of RFLP types was still low, the GD value rose up to 0.81 in the 1970s. This rise in GD may have been due to the low number of isolates sampled from the 1960s and 1970s compared to the other decades. In the 1980s, vaccination coverage had dropped to 30% due to reports concerning the safety of the pertussis vaccine, and two large pertussis epidemics occurred. These events were reflected in the increased number of RFLP types and the rise in genotypic diversity to 0.84. This could possibly have been due to the emergence of new variants or the re-emergence of less-fit *B. pertussis* strains, as the herd immunity in the UK was noticeably reduced. However, the vaccine coverage recovered to 94% in the late 1990s, resulting in the increase of herd immunity. Low numbers of pertussis cases was thus seen, and the reduced *B. pertussis* population experienced a significant drop in genotypic diversity to 0.57 ($P<0.01$). Combined with the fall in RFLP types and a significant increase in proportion of RFLP-1 isolates, a second reduction of less-fit strains took place, with RFLP-1 expanding as a more fit, predominant lineage of *B. pertussis.*

Due to the fall in vaccine coverage in the UK in the late 1970s, comparisons may be difficult to other countries with steady coverage, such as The Netherlands. However, after the introduction of vaccination in the early 1950s, The Netherlands also saw a decrease in GD, and suggested that it may have been due to;

> "… a decrease in the bacterial population size and/or by clonal expansion …".
>
> (van Loo *et al.*, 1999)

Thereafter, genotypic diversity gradually rose and, between 1976 and 1981, the GD was 0.69 in The Netherlands (van Loo *et al.*, 1999), whereas a higher GD value of 0.81 was seen in the UK in the 1970s. Data from the two countries gave similar high GD values once more in the 1980s (1982 to 1990 in The Netherlands). This similarity however was not reflected in the pertussis notification rates, as the UK experienced epidemic levels, yet The Netherlands did not. The drop in genotypic diversity in the 1990s was not as pronounced in The Netherlands than for the UK, and this was perhaps due to the resurgence of pertussis in The Netherlands, where some clonal expansion was still occurring, compared to the elimination of less-fit strains in the UK. So, although the UK encountered a drop in vaccination coverage and is currently experiencing low numbers of pertussis cases, both the major RFLP types and the temporal changes in genotypic diversity are similar to those in The Netherlands, and yet The Netherlands has seen a recent resurgence in pertussis.

The genotypic diversity, according to RFLP type distribution within each pertactin type, was calculated to allow comparisons with the Dutch data (van Loo *et al.*, 1999). A marked difference was found in the *prnA*(2) and *prnA*(3) values, as GD values of 0.39 and 0.00 were seen in the UK, and GD values of 0.78 and 0.57 in The Netherlands, respectively (van Loo *et al.*, 1999). This difference could partly have been due to the frequencies of pertactin types (UK, 118/149 isolates, 79% *prnA*(1), c.f. 40/134, 30% *prnA*(1) in The Netherlands) and the sample period used (UK isolates from 1940 to 1999, c.f. 1981 to 1996 in The Netherlands) (van Loo *et al.*, 1999). Van Loo *et al.* (1999) also reported that the higher GD of *prnA*(1) isolates (GD=0.91) compared to the GD values of *prnA*(2) and *prnA*(3) isolates suggested that *prnA*(1) represents the oldest lineage, and that *prnA*(3) isolates represents a younger lineage than *prnA*(2). If this hypothesis is correct, then pertactin types may not necessarily share the same RFLP type, but this does occur, and will be discussed later. Also, if *prnA*(1) is the oldest lineage, and thus a more stable genotype, then a lower GD value would be expected, with less RFLP types. Conversely, an emerging variant of pertactin would result in an expansion of strains and a higher GD value. Yet the GD values for pertactin types from both the UK and The Netherlands are similar in that *prnA*(1) exhibits the highest GD (van Loo *et al.*, 1999).

Genotypic diversity values were also calculated for the different pertussis toxin S1 and combined *prnA/ptxA* types. Higher GD values in *ptxA*(1) isolates indicate successful genotypes of *B. pertussis*, but these are not statistically significant. Van Loo *et al.* (1999) did not present the genotypic diversity in *ptxA* types and this could have been due to the small number of isolates sequenced for *ptxA* from a previous study (n=45) (Mooi *et al.*, 1998), and that the GD values calculated may not have shown a significant difference between *ptxA*(1) and *ptxA*(2) or *ptxA*(3), and hence would not have supported their theory of vaccine-driven evolution.

As mentioned earlier, associations between RFLP types and pertactin and pertussis toxin S1 types were investigated. Unfortunately, no data were available from The Netherlands concerning any statistical analysis on particular RFLP types with *prnA* or *ptxA* types (van Loo *et al.*, 1999). Strong associations between *prnA*(1) and RFLP-2 or RFLP-4 were observed in the UK isolates (114/149), but Dutch data showed a higher frequency of *prnA*(2) isolates that were RFLP-2 (van Loo *et al.*, 1999). Other significant associations in the UK included types RFLP-1 and RFLP-3 exclusively associated with *ptxA*(1), and RFLP-2 associated with *ptxA*(2). All of these associations indicate the overall clonality of the *B. pertussis* population.

3.4.5 UK *Bordetella pertussis* sub-set from 1998 to 1999

There is limited value in studying the population dynamics of *B. pertussis* by pertactin and pertussis toxin S1 sequence typing, and IS*1002*-RFLP analysis in isolation. There is no clinical relevance unless the isolates have associated information regarding patient status and epidemiological data. This was achieved using epidemiological information from the enhanced pertussis surveillance database from CDSC, which was established in 1994. Therefore, only data from isolates that were collected from the Pertussis Reference Laboratory in Manchester between September 1998 and April 1999 could be merged and evaluated.

Van Buynder *et al.* (1999) reported on *B. pertussis* surveillance in England and Wales between 1995 and 1997, using data received from the Office of National Statistics (ONS), the Department of Health hospital episodes statistics (HES), and the enhanced pertussis surveillance database. Comparing the 90 isolates analysed in this thesis with the laboratory reports interrogated by van Buynder and colleagues (1999), the percentages for each age group were almost identical. For example, the under two month olds encompassed 23 of the 90 isolates in this study (25.6%, 95% CI: 17-36%), compared to 48 of the 200 isolates (24%, 95% CI: 18-31%) reported in the cited paper (van Buynder *et al.*, 1999). The other four age groups differed by no more than three to four percent. Other characteristics that were similar between the two data-sets included sex (39/90 males, 43.3% cf. 49.8%, for the isolates reported here, and the isolates from van Buynder and co-authors (1999), respectively), vaccination status (no vaccination in 47 cases, 52.2%, CI: 41-63% cf. 424/709 cases, 60%, CI: 56-63%), complications encountered (22 cases, 24.4%, CI: 16-35% cf. 177/709, 25.0%, CI: 22-28%), hospitalised cases (54, 60.0%, CI: 49-70% cf. 465/698, 66.6%, CI: 63-70%) and mortalities (three deaths in eight months cf. 12 deaths in three years=2.7 deaths in an eight month period) (van Buynder *et al.*, 1999). The serotype distribution differed slightly between 1995 to 1997, in that 352/780 isolates (45.1%, CI: 42-49%) were serotype 1,2, 411/780 (52.7%, CI: 49-56%) were serotype 1,3, and 17/780 (2.2%, CI: 1-3%) were serotype 1,2,3 (van Buynder *et al.*, 1999). The data collected between 1998 and 1999 for this study did not contain serotype 1,2,3, and isolates have shifted to 49/90 (54.4%, CI: 44-65%) of serotype 1,2 and 41/90 (45.6%, CI: 35-56%) serotype 1,3. This is reflected in other past studies, where shifts in serotype expression in *B. pertussis* populations have been reported (Preston, 1988), and recent isolates from 2002 now appear to be 96% serotype 1,3 (71/74 isolates serotyped) (TG Harrison, pers. comm.). Overall,

this comparatively small data-set of 90 *B. pertussis* isolates is epidemiologically no different from the larger data-set analysed from 1995 to 1997 (van Buynder *et al.*, 1999).

The other typing data generated included the pertactin and pertussis toxin S1 gene sequence types, and IS*1002*-RFLP types. Eighteen different RFLP types were seen in the 90 isolates compared to six RFLP types from 36 isolates used for the genotypic diversity study (3.3.4, page 112), which were chosen from isolates between 1998 and 1999. This difference in RFLP types was probably due to the respective number of isolates tested and can be corrected when seen as ratios (0.2 types per isolate for the 90 isolate data-set and 0.17 types per isolate for the 36 isolate data-set). However, this certainly exemplifies how the sample size can affect results, analysis and subsequent conclusions. The two most dominant RFLP types occur at similar frequencies of both the 90 isolates and the 36 from the genotypic diversity study, where RFLP-1 occurs at 49/90 (54%) and 23/36 (64%), and RFLP-3 in 14/90 (16%) and 7/36 (19%), respectively.

To create a more informed view of the strains currently circulating within the *B. pertussis* population, the two different serotypes, three pertactin types and 18 RFLP types were integrated to generate 24 different combined types seen between September 1998 and April 1999. Similar to the combined *prnA/ptxA* genotypes discussed in 3.4.3 (page 146), four out of a theoretical six *prnA/ptxA* types were seen, compared to 24 out of a theoretical 108 different combined types. This suggests that *B. pertussis* is clonal in terms of these independent typing data, exemplified by the widespread and stable dominant RFLP types, RFLP-1 and RFLP-3. Therefore a mechanism of genetic restriction must be occurring to ensure the stability and fitness of the *B. pertussis* population.

This clonality is more pronounced when the 90 isolates were sorted into the 24 combined types, as only three types contained most of the strains (62/90 or 69% of the isolates). These combined types, A, C and P were generated from the following types; serotypes 1,2 and 1,3, *prnA*(1) and *prnA*(2), and RFLP-1 and RFLP-3 (see Table 3.14, page 131). It can be therefore assumed that these different types when combined, have produced a few successful lineages of *B. pertussis* which have expanded and caused infection and disease in the UK population as seen in this brief period between September 1998 to April 1999. This indicator of a clonal species is also supported by the many, less frequent combined types that have not expanded into the UK *B. pertussis* population, and are present only in small numbers. However, whether these combined types A, C, and P are persistent and stable

cannot be deducted from these data, and further studies on *B. pertussis* populations may resolve this.

3.4.6 Associations within the pertussis epidemiological data of the 90 UK cases from 1998 to 1999

A primary analysis of the epidemiological data was performed to identify any dependent variables and to eliminate any bias when further analysing associations between epidemiological and laboratory data. Associations that van Buynder *et al.* (1999) reported for data obtained between 1995 to 1997 were also seen in the data from 1998 to 1999 presented here. These included a higher likelihood of admission to hospital in those under two months old, and in the unvaccinated, in terms of number of vaccine doses received. The data here also showed an association between pertussis cases encountering complications, especially apnoea, and admission to hospital. These associations can be rooted back to the under two month old age group, where this 'risk' group is evidently one requiring hospital treatment for the complications of pertussis, and yet this group is not eligible for the receipt of pertussis vaccine under the current UK schedule of 2, 3, and 4 months. This over-burdened cohort has recently been addressed (von Konig *et al.*, 1995); (Crowcroft and Britto, 2002) and, following a study showing pertussis transmission from contacts to babies in five paediatric intensive care units (PICUs), a pertussis vaccine booster for 4 to 5 year olds was introduced in the UK, in November 2001 (Crowcroft *et al.*, 2003). Future analysis concerning burden of pertussis of this vulnerable group needs to be addressed, to assess whether the booster is improving the situation, or if further measures, such as an additional pertussis booster is required for adolescents or women at child-bearing age.

Geographical regions in the UK also gave two significant associations; one with the complication pneumonia, and the other with gender. Studying the breakdown of data however, revealed that the cases seem to be random and, due to this random nature, the proportions between regions can vary markedly. A significant probability was therefore generated, and if a larger data-set was analysed, these associations would possibly become insignificant. Borderline associations with the regions were seen with the following outcomes: typical clinical symptoms, encountering any complication, apnoea and hospital admission. These data, even after a multivariable analysis, were hard to interpret, probably due to the low number and random distribution of cases.

Those presenting with typical clinical symptoms of pertussis were less likely than those not presenting with typical clinical symptoms to have pneumonia. This association was independent of age group, and there were only nine pneumonia cases seen in the total of 90. An association between deaths and absence of typically clinical symptoms was also seen; of the three cases that died, two were under two months and the third case was in the two to five month age group. Neonates generally do not present with typical symptoms of pertussis (von Konig *et al.*, 2002). The England and Wales enhanced pertussis surveillance database derives data from questionnaires, and more emphasis may be directed to the complication pneumonia or death, than to whether a case has typical clinical symptoms.

The most unusual association was between gender and hospital admission, as females were more likely to be admitted than males. However, this may have been affected by various facets; i) a skewed proportion of females to males (51:39 respectively); ii) a higher proportion of females in the younger age group (15:8, female:male) and; iii) reports of morbidity and mortality rates being higher in females than males are documented (Edwards *et al.*, 1999). So although van Buynder *et al.* (1999) reported no gender differences with hospital admissions, an association was seen in this small data-set of 90 isolates from 1998 to 1999, although this was not significant in the multivariable logistic regression analysis (P=0.088, Table 3.12, page 128). Overall, the epidemiological data of these 90 cases seemed to reflect earlier findings of pertussis surveillance in England and Wales (van Buynder *et al.*, 1999).

3.4.7 Association between epidemiological data and typing results of the 90 UK *Bordetella pertussis* isolates from 1998 to 1999

In the single variable analyses, three statistically significant associations were found between the epidemiological data and typing results, where two of these associated the severity measure of hospital admission with either serotype or pertactin gene type. Serotype 1,2 isolates were nearly four times more likely than serotype 1,3 to be associated with an admission to hospital. Similar results were seen in a previous study in the UK, with pertussis cases between 1995 and 1997, where serotype 1,2 was twice as likely as serotype 1,3 to be associated with hospitalised cases (van Buynder *et al.*, 1999). The other statistical association was that *prnA*(1) isolates were nearly four times more likely to be seen in hospital admissions than *prnA*(2). These two observations with serotype 1,2 and *prnA*(1) may not be independent, as there was a strong association between serotype and pertactin type, where serotype 1,2 isolates were seen with *prnA*(1) isolates, and serotype 1,3 with

prnA(2). As serotype 1,2,3 was not seen, and *prnA*(3) was observed in only two isolates, associations for these types could not be determined. The mechanism behind the association between serotype and pertactin has not been deduced; genome analysis and expression studies concerning the regulation of the related genes may reveal this.

The third significant association involved the geographical regions with RFLP types. Although RFLP-3 isolates were only found in four of the nine regions, the large number of parameters generate many degrees of freedom and high χ^2 values, resulting in small P values. The χ^2 value is also dependent on the number of isolates analysed, so this sub-set of 90 was not sufficient to calculate associations for the aspects mentioned above, and more isolates are required for a more thorough analysis.

Generally, statistical significance is measured at the 95% confidence level (P=0.05) to reduce the number of false positives, or in this case, false associations. But with this strict confidence limit some associations may be overlooked, particularly in a small data-set. Therefore, any associations with a P value of 0.20 or under were noted as borderline significance, and were treated as such. If more data had been available, these borderline associations might have been ascertained at a more confident level. Examples of these borderline associations include; vaccination status with serotype, geographical regions with combined types, those encountering any complication with combined type A isolates, and the conjunctivitis case with combined type C. Other borderline associations link the under two month age group with serotype 1,2, *prnA*(1) and combined type A. An earlier study in the UK showed a similar association, focusing on the serotype 1,3 proportion between the three to 12 month age group (38%) and the 12 to 24 month age group (74%) (Syedabubakar *et al.*, 1995). The more recent study by van Buynder *et al.* (1999) did not reveal any significant associations between serotype and age distribution.

Trend analysis was also performed on the time-related aspects, age cohorts and the interval of isolation collection (temporal trend). The age groups showed a downward trend in the distribution of serotype 1,2 and *prnA*(1) isolates from under two months to the six to 11 month age group. These isolate types then recovered slightly in the one to four, and five years and above age groups. This fall and rise in *prnA*(1) isolates through age cohorts has also been seen in The Netherlands (Mooi *et al.*, 1998), although the actual proportions were different to those seen in the UK (for the under two months in the UK, 15/23 (65%) were *prnA*(1), and the under three months group in The Netherlands, 22% were *prnA*(1)) Mooi *et al.* (1998). Mooi *et al.* (1998) also showed that the age groups can be converted into the

vaccination status of individuals as they receive the pertussis vaccine at 2, 3 and 4 months of age, with an additional booster at 11 months. The UK pertussis vaccine is also given at 2, 3 and 4 months of age (the booster for 4-5 year olds was introduced after the collection of isolates in November 2001), and the vaccine coverage was 94% in 1999 in those up to 24 months old. Hence, the following can be deduced; individuals under two months are not immune, those between two and five months are partially immunised, those between six and 11 months, and from one to four years are fully immunised and, from five years onward, individuals have a waning immunity to pertussis infection. Therefore, in the UK most infections in unvaccinated babies are caused by serotype 1,2 and *prnA*(1) isolates and, as individuals get older, their pertussis vaccination course would be complete, and hence their immune response against *B. pertussis* should be at their highest. At this point, the serotype 1,2/*prnA*(1) infections decrease in the one to four year age group (Table 3.21 and 3.22, page 140). As the individuals grow older, their immunity to pertussis wanes and serotype 1,2/*prnA*(1) isolates cause an increased number in infections. In The Netherlands, this was observed and interpreted as;

> "… the notion that vaccine-induced immunity against *prnA*(1) strains is stronger than that against *prnA*(2) and *prnA*(3) strains …".
>
> (Mooi *et al.*, 1998)

The Aventis Pasteur WCV used in the UK contains two *prnA*(1) strains and, thus, if agreeing with Mooi *et al.* (1998) observations, any individual who received the vaccine may have a strong vaccine-induced immunity against *prnA*(1) and may be more likely to acquire a non-vaccine *prnA* type of pertussis infection. Therefore this trend of pertactin type is not affected by the vaccine constituents *per se*, but by the immune status of the individual. This fall and rise through the age groups can also be observed in the notification data of pertussis (Notifications of Infectious Diseases, 2003). However, both trends of serotype and pertactin type across age groups are only of borderline significance, and thus should only be seen as an observation and not as evidence of a definite association.

The other trend analysis regarding the collection period with either serotype or pertactin type, was also not significant (score test for trend of odds at P=0.497 and 0.554, respectively). The proportion of both serotype 1,2 and *prnA*(1) isolates dropped from 4/8 (50%) in September 1998, to 5/13 (38%) in March 1999. As the last isolate was collected on the 2nd April 1999, this one isolate could not validate whether this shift was true. However, isolates from 2002, which were used for the pertactin and pertussis toxin data (3.3.1 and 3.3.2, pages 105 and 106, respectively), give a more pronounced reduction in serotype 1,2 and *prnA*(1), to just 1/22 isolates (5%), and epidemiological data from the

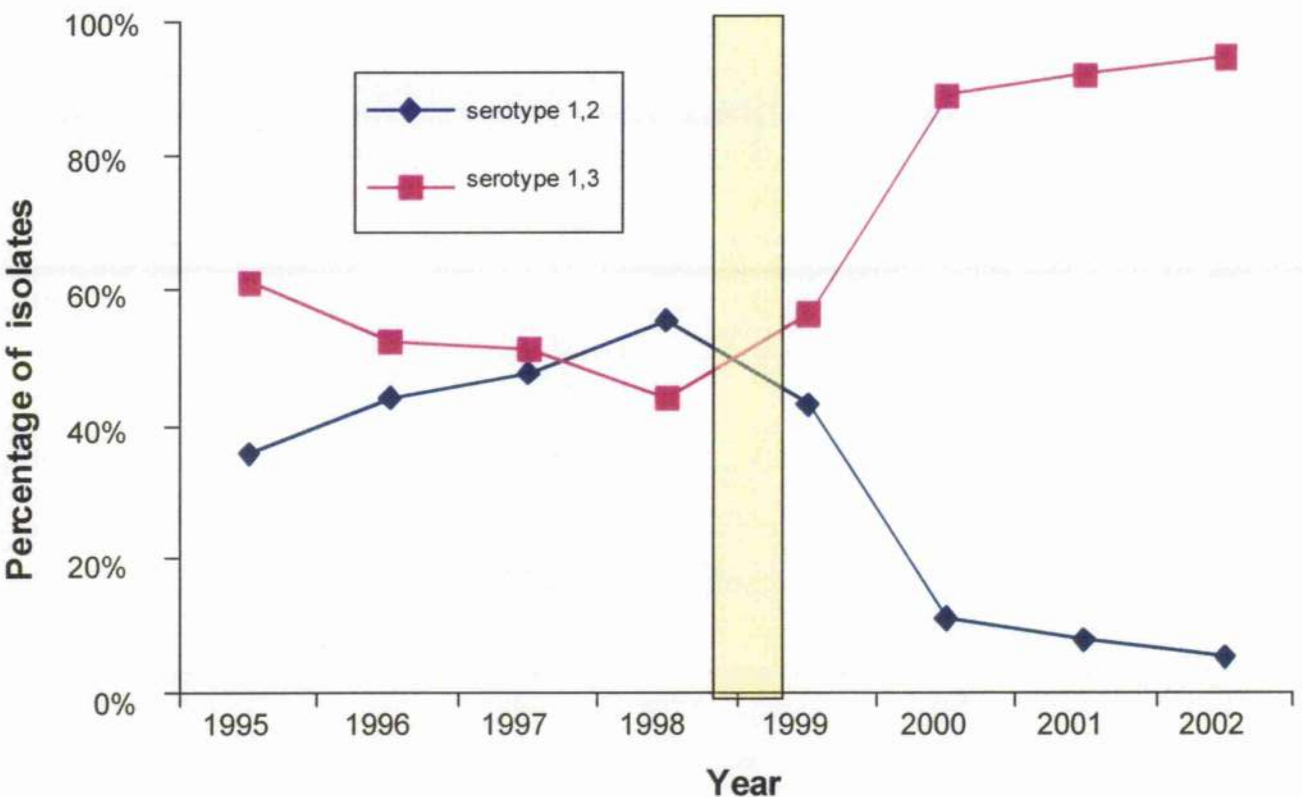
enhanced laboratory surveillance shows the dramatic decline of serotype 1,2 isolates in 1999 to under 10% (Figure 3.10, page 161) (Immunisation Division, CDSC, HPA, 2004b). Reasons for this dramatic shift in serotype and pertactin type are unknown, although it may be due to a change in vaccination policy, such as the change of WCV in 1996 to Aventis Pasteur, the temporary introduction of acellular pertussis vaccine in Autumn 1999, and the booster in pre-school children in November 2001, which all contain *prnA*(1) strains (serotypes not known). Another factor that could have caused this shift in proportions is a decrease in vaccine quality and/or efficacy, but this is highly unlikely because the UK has not seen an overall increase or resurgence of pertussis, which would be indicative of inadequate vaccine manufacture. A further point to address concerns the shift in serotype distribution, which has been observed throughout the decades in the UK, via slipped-strand mispairing, and this probably reflects a natural shift in the *B. pertussis* population.

Although the two *prnA*(3) isolates were not used in the multivariable logistic regression analysis, they were from cases under three months old, who had typical clinical symptoms, developed apnoea and were admitted to hospital. These factors, although not statistically significant, might point to a more virulent strain of *B. pertussis*. Once more, additional isolates are required, especially those of type *prnA*(3), to elucidate whether these associations could be significant. If so, it may be supportive of a previously described hypothesis (Mooi *et al.*, 1999; Gzyl *et al.*, 2002). These authors suggested that an increase of *prnA*(3) isolates may affect the epidemiology of pertussis, and this may be one reason why The Netherlands and Finland are seeing a resurgence, and Poland is not; approximately 65/133 (49%) and 5/43 (12%) of *prnA*(3) isolates, compared to 0/19 (0%), respectively (Gzyl *et al.*, 2002).

This unique set of isolates, incorporating epidemiological data and laboratory typing results, albeit small, was still able to identify two significant associations and other borderline observations. Previous studies have either focused on epidemiology and surveillance using clinical data (de Melker *et al.*, 2000a; Guris *et al.*, 1999b; Skowronski *et al.*, 2002), or genetic diversity and outbreak investigations using laboratory typing (de Moissac *et al.*, 1994; Syedabubakar *et al.*, 1995; van Loo *et al.*, 1999). This is unsatisfactory for the scientific community, as more valuable data can be generated using an integrated approach, especially in countries where pertussis resurgence occurs, and this could reveal a better picture of *B. pertussis* dynamics. Knowledge gained from such studies could have important vaccine implications, whether in formulating the composition or the schedule. With such qualitative and detailed data, overinterpretation can be a problem, and associations that are

not significant may get too much emphasis. More significant associations may have been identified if there were more isolates in this data-set, as borderline associations need to be verified. The inclusion of more isolates may still not be useful when analysing RFLP and the combined types however, as these methods give higher discrimination, and an increase of types can impair interpretation when used with epidemiological data. However, IS*1002*-RFLP analysis and combined types do reveal a more dynamic population of *B. pertussis* than the individual serotype, pertactin and pertussis toxin gene types.

Figure 3.10. *Bordetella pertussis* serotype distribution in England and Wales, from 1995 to 2002



Composite data from the Enhanced Pertussis Surveillance, CDSC, Manchester University and RSIL (Immunisation Division, Communicable Disease and Surveillance Centre, Health Protection Agency, 2004b). The yellow shaded area encompasses the period when isolates were collected for association analyses with the case data.

4 Characterisation of *Bordetella pertussis* isolates from Missouri, USA

The study of *B. pertussis* isolates from the UK (Chapter 3) investigated any epidemiological trends and long-term shifts in the bacterial population and the influence, if any, of the introduction and widespread use of the pertussis vaccine. The UK has not undergone an apparent resurgence of pertussis, unlike many other countries with high vaccination coverage. This chapter focuses on a smaller set of isolates from Missouri, USA, to question whether any short-term shifts could be identified, using a combined typing approach.

Since the late 1980s, the reported incidence of pertussis has increased in the state of Missouri, USA, from 0.46 cases per 100,000 in 1984 to a peak of 2.77 per 100,000 in 1989. Cases remained high with a recent incidence of 1.91 per 100,000 in 2001 (Missouri Department of Health and Senior Services, 2002). In 1999 there were 1.38 cases per 100,000 in Missouri (75 cases), peaking in June to July and September to October. The rise in June and July coincided with an increase of isolates from Kansas City sent to the Missouri State Public Health Laboratory (Figure 4.1, page 164). In Missouri, culture-positive cases are used in conjunction with the CDC clinical criteria to report a pertussis case (Guris *et al.*, 1999a), but epidemiological typing of *B. pertussis* is not currently performed.

4.1 Strategy

The clinical isolates obtained from Missouri, described in Appendix 3, Part D, were investigated using the following methods; serotyping, sequencing of the pertactin and pertussis toxin S1 genes, IS*1002*-RFLP analysis and PFGE with *Xba*I. No strains used in the USA vaccine preparations were available for examination. However, *prnA* and *ptxA* data from a previous USA study describing specific vaccines and strains will be referred to in the discussion (Cassiday *et al.*, 2000). The index of discrimination and associations between typing methods were assessed. Data gathered here were then compared to the UK data from Chapter 3 and other reported countries, and geographical differences and vaccine regimens are discussed.

Figure 4.1. The state of Missouri, USA, with main cities located, including Kansas City

Taken from http://www.mwsc.edu/education/question.html. Last accessed 22nd March 2004.

Forty-six *B. pertussis* isolates were sent to RSIL for analysis, one of which was subsequently identified as *B. parapertussis* and therefore excluded from this study. One other isolate was excluded because it was recovered from Kansas City five weeks after the last isolate from Kansas City; therefore it could not be determined whether it was an outbreak isolate or an unrelated isolate. The remaining 44 isolates were grouped into three subsets; (i) ten isolates from Kansas City, which were geographically and temporally clustered within seven weeks between June and July 1999 (referred to as "Kansas City" isolates throughout this chapter); (ii) 19 recent isolates that were epidemiologically-unrelated and were collected between 1996 and 2001 and; (iii) 15 earlier isolates also epidemiologically-unrelated, collected between 1984 and 1989. Isolates from 1990-1995 were not available.

Sequences comprising over 90% of the *prnA* coding region were investigated in six of the isolates, and were deposited in GenBank via the BankIt submission entry (http://www.ncbi.nlm.nih.gov/BankIt/index.html). This was to determine the level of sequence variation outside the two known polymorphic regions 1 and 2 of the pertactin gene (Mooi *et al.*, 1998), and required amplification and subsequent sequencing of the three overlapping PCR fragments described in 2.5.4.1 (page 68). The six isolates were chosen for their isolation date and *prnA/ptxA* combination. Region 1 and region 2 (Figure 1.2, page 16), were sequenced for the remaining 38 strains, following amplification of a 1428 bp fragment with primers AF and BR.

4.2 Methods

The following techniques used in this chapter are described in Chapter 2; serotyping, pertactin and pertussis toxin S1 subunit gene typing (PCR amplification using primer pairs AF and BR, and SF and SR, respectively), IS*1002*-RFLP analysis and PFGE using *Xba*I. The serotyping scheme used for this thesis differs from the scheme used in the USA, where three agglutinogens are tested, compared to six (Robinson *et al.*, 1989). The discriminatory indices were calculated using the formula expressed in section 2.8.3 (page 99).

4.3 Results

4.3.1 Serotyping results

Serotype 1,3 was predominant, occurring in 39/44 (89%) of the isolates, and two isolates were serotype 1,2,3 and one was serotype 1,2 (Table 4.1, page 167). Two isolates gave discrepant results when performing the slide-agglutination test. Serotyping was attempted

on these two isolates a further three times using fresh culture each time and resulted in the expression of agglutinogen 1 only in MO-301, and isolate MO-404 auto-agglutinated. All ten of the "Kansas City" isolates were serotype 1,3.

4.3.2 Polymorphism in the pertactin and pertussis toxin S1 genes

The nucleotide sequences of four of the six strains for which over 90% of the *prnA* coding region was determined (Table 4.2, page 167), were identical to the previously designated pertactin types *prnA*(1), *prnA*(2) or *prnA*(6), (GenBank accession numbers AJ011091, AJ011092 or AJ132095, respectively) (Mooi *et al.*, 1998). The nucleotide sequences from two isolates with the novel pertactin type *prnA*(9) found in this study were identical (100% for 2817 nucleotides). This novel type *prnA*(9) contains seven GGxxP repeats $(GGAVP)_2$, $(GGFGP)_5$ in region 1, and five PQP repeats in region 2, and differs from *prnA*(2) by the presence of an extra GGFGP repeat (Figure 4.2, page 168). This novel pertactin type had not been previously reported and sequences from both strains were deposited in GenBank under the following accession numbers; AF218785 (strain MO-908) and AF456356 (strain MO-121).

Region 1 and region 2 were sequenced for all 44 isolates. No variation was seen in region 2 and all isolates contained the nucleotide sequence encoding five PQP repeats. The point mutation at approximately 150 nucleotides upstream of region 2 that differentiates *prnA*(7) from *prnA*(1), previously described by Mooi *et al.* (1998), was not found. Three previously described *prnA* variants, *prnA*(1), *prnA*(2), and *prnA*(6), were recovered from the isolates tested. The novel variant, designated *prnA*(9), was seen in two of the 44 isolates.

Thirty-four of the 44 isolates were sorted into temporal groups labelled "Early" (1984-1989) and "Recent" (1996-2001). All four *prnA* types were found in 1984-1989, nine of the 15 isolates were *prnA*(1), three were *prnA*(2), one *prnA*(6) and two were of the novel type *prnA*(9) (Table 4.1, page 167). The frequency of *prnA*(2) isolates increased from 3/15 (20%) isolates in 1984-1989 to 18/19 (95%) isolates in 1996-2001, with the remaining isolate from 1996-2001 being *prnA*(1). This shift from predominantly *prnA*(1) to *prnA*(2) is significant even in such small numbers (Fisher's exact test P=0.001). The ten isolates from the suscept outbreak cluster "Kansas City" were all *prnA*(2).

Table 4.1. Distribution of serotype, pertactin and pertussis toxin S1 subunit variants from 44 *Bordetella pertussis* isolates from Missouri, USA, by isolation period

Typing method and type	Year period		
	"Early", 1984–1989, n=15	"Recent", 1996–2001, n=19	"Kansas City", 1999, n=10
Serotype			
1 only		1	
1,2	1		
1,3	12	17	10
1,2,3	1	1	
auto-agglutination	1		
Pertactin gene type			
prnA(1)	9	1	
prnA(2)	3	18	10
prnA(6)	1		
prnA(9)	2		
Pertussis toxin S1 subunit gene type			
ptxA(1)	12	18	10
ptxA(2)	3	1	
Pertactin/Pertussis toxin S1 subunit alleles			
prnA(1)/*ptxA*(1)	7		
prnA(1)/*ptxA*(2)	2	1	
prnA(2)/*ptxA*(1)	3	18	10
prnA(6)/*ptxA*(2)	1		
prnA(9)/*ptxA*(1)	2		

Values are the number of strains. The "Kansas City" data are in grey, as the isolates were not included in the temporal analysis.

Table 4.2. Details of the six *Bordetella pertussis* isolates from Missouri, USA, for which over 90% of the coding region of the pertactin gene was determined

Isolate	Year isolated	Pertactin gene type	GenBank Accession number of pertactin gene sequence	Pertussis toxin S1 gene type
MO-20	1984	*prnA*(6)	AF456357	*ptxA*(2)
MO-814	1984	*prnA*(1)	AF456355	*ptxA*(1)
MO-908	1988	*prnA*(9)	AF218785	*ptxA*(1)
MO-121	1989	*prnA*(9)	AF456356	*ptxA*(1)
MO-1440	1999	*prnA*(2)	AF456358	*ptxA*(1)
MO-1503	1999	*prnA*(1)	AF456359	*ptxA*(2)

Figure 4.2. Nucleic acid alignment of the *Bordetella pertussis* pertactin gene types, *prnA*(2) and *prnA*(9)

```
       CONSENSUS: CGCGGTGCGGGCATGGCGGGCGTCTGCTCTCCACCTGGCATCCAATGAAC   6
                                                              M  N
AJ011092 prnA(2): -----------------GGGCGTCTGCTCTCCACCTGGCATCCAATGAAC
AF456356 prnA(9): CGCGGTGCGGGCATGGCGGGCGTCTGCTCTCCACCTGGCATCCAATGAAC
AF218785 prnA(9): --------------------------------------------------

       CONSENSUS: ATGTCTCTGTCACGCATTGTCAAGGCGGCGCCCCTGCGCCGCACCACGCT   56
                   M  S  L  S  R  I  V  K  A  A  P  L  R  R  T  T  L
AJ011092 prnA(2): ATGTCTCTGTCACGCATTGTCAAGGCGGCGCCCCTGCGCCGCACCACGCT
AF456356 prnA(9): ATGTCTCTGTCACGCATTGTCAAGGCGGCGCCCCTGCGCCGCACCACGCT
AF218785 prnA(9): ---------------------AAGGCGGCGCCCCTGCGCCGCACCACGCT

       CONSENSUS: GGCCATGGCGCTGGGCGCGCTGGGCGCCGCCCCGGCGGCGCATGCCGACT   106
                   A  M  A  L  G  A  L  G  A  A  P  A  A  H  A  D
AJ011092 prnA(2): GGCCATGGCGCTGGGCGCGCTGGGCGCCGCCCCGGCGGCGCATGCCGACT
AF456356 prnA(9): GGCCATGGCGCTGGGCGCGCTGGGCGCCGCCCCGGCGGCGCATGCCGACT
AF218785 prnA(9): GGCCATGGCGCTGGGCGCGCTGGGCGCCGCCCCGGCGGCGCATGCCGACT

       CONSENSUS: GGAACAACCAGTCCATCGTCAAGACCGGTGAGCGCCAGCATGGCATCCAT   156
                   W  N  N  Q  S  I  V  K  T  G  E  R  Q  H  G  I  H
AJ011092 prnA(2): GGAACAACCAGTCCATCGTCAAGACCGGTGAGCGCCAGCATGGCATCCAT
AF456356 prnA(9): GGAACAACCAGTCCATCGTCAAGACCGGTGAGCGCCAGCATGGCATCCAT
AF218785 prnA(9): GGAACAACCAGTCCATCGTCAAGACCGGTGAGCGCCAGCATGGCATCCAT

       CONSENSUS: ATCCAGGGCTCCGACCCGGGCGGCGTACGGACCGCCAGCGGAACCACCAT   206
                   I  Q  G  S  D  P  G  G  V  R  T  A  S  G  T  T  I
AJ011092 prnA(2): ATCCAGGGCTCCGACCCGGGCGGCGTACGGACCGCCAGCGGAACCACCAT
AF456356 prnA(9): ATCCAGGGCTCCGACCCGGGCGGCGTACGGACCGCCAGCGGAACCACCAT
AF218785 prnA(9): ATCCAGGGCTCCGACCCGGGCGGCGTACGGACCGCCAGCGGAACCACCAT

       CONSENSUS: CAAGGTAAGCGGCCGTCAGGCCCAGGGCATCCTGCTAGAAAATCCCGCGG   256
                   K  V  S  G  R  Q  A  Q  G  I  L  L  E  N  P  A
AJ011092 prnA(2): CAAGGTAAGCGGCCGTCAGGCCCAGGGCATCCTGCTAGAAAATCCCGCGG
AF456356 prnA(9): CAAGGTAAGCGGCCGTCAGGCCCAGGGCATCCTGCTAGAAAATCCCGCGG
AF218785 prnA(9): CAAGGTAAGCGGCCGTCAGGCCCAGGGCATCCTGCTAGAAAATCCCGCGG

       CONSENSUS: CCGAGCTGCAGTTCCGGAACGGCAGTGTCACGTCGTCGGGACAGTTGTCC   306
                  A  E  L  Q  F  R  N  G  S  V  T  S  S  G  Q  L  S
AJ011092 prnA(2): CCGAGCTGCAGTTCCGGAACGGCAGTGTCACGTCGTCGGGACAGTTGTCC
AF456356 prnA(9): CCGAGCTGCAGTTCCGGAACGGCAGTGTCACGTCGTCGGGACAGTTGTCC
AF218785 prnA(9): CCGAGCTGCAGTTCCGGAACGGCAGTGTCACGTCGTCGGGACAGTTGTCC

       CONSENSUS: GACGATGGCATCCGGCGCTTTCTGGGCACCGTCACCGTCAAGGCCGGCAA   356
                   D  D  G  I  R  R  F  L  G  T  V  T  V  K  A  G  K
AJ011092 prnA(2): GACGATGGCATCCGGCGCTTTCTGGGCACCGTCACCGTCAAGGCCGGCAA
AF456356 prnA(9): GACGATGGCATCCGGCGCTTTCTGGGCACCGTCACCGTCAAGGCCGGCAA
AF218785 prnA(9): GACGATGGCATCCGGCGCTTTCTGGGCACCGTCACCGTCAAGGCCGGCAA

       CONSENSUS: GCTGGTCGCCGATCACGCCACGCTGGCCAACGTTGGCGACACCTGGGACG   406
                   L  V  A  D  H  A  T  L  A  N  V  G  D  T  W  D
AJ011092 prnA(2): GCTGGTCGCCGATCACGCCACGCTGGCCAACGTTGGCGACACCTGGGACG
AF456356 prnA(9): GCTGGTCGCCGATCACGCCACGCTGGCCAACGTTGGCGACACCTGGGACG
AF218785 prnA(9): GCTGGTCGCCGATCACGCCACGCTGGCCAACGTTGGCGACACCTGGGACG

       CONSENSUS: ACGACGGCATCGCGCTCTATGTGGCCGGCGAACAGGCCCAGGCCAGCATC   456
                  D  D  G  I  A  L  Y  V  A  G  E  Q  A  Q  A  S  I
AJ011092 prnA(2): ACGACGGCATCGCGCTCTATGTGGCCGGCGAACAGGCCCAGGCCAGCATC
AF456356 prnA(9): ACGACGGCATCGCGCTCTATGTGGCCGGCGAACAGGCCCAGGCCAGCATC
AF218785 prnA(9): ACGACGGCATCGCGCTCTATGTGGCCGGCGAACAGGCCCAGGCCAGCATC
```

Figure 4.2. Nucleic acid alignment of the *Bordetella pertussis* pertactin gene types, *prnA*(2) and *prnA*(9), continued

```
       CONSENSUS: GCCGACAGCACCCTGCAGGGCGCTGGCGGCGTGCAGATCGAGCGCGGCGC    506
                   A  D  S  T  L  Q  G  A  G  G  V  Q  I  E  R  G  A
AJ011092 prnA(2): GCCGACAGCACCCTGCAGGGCGCTGGCGGCGTGCAGATCGAGCGCGGCGC
AF456356 prnA(9): GCCGACAGCACCCTGCAGGGCGCTGGCGGCGTGCAGATCGAGCGCGGCGC
AF218785 prnA(9): GCCGACAGCACCCTGCAGGGCGCTGGCGGCGTGCAGATCGAGCGCGGCGC

       CONSENSUS: CAATGTCACGGTCCAACGCAGCGCCATCGTCGACGGGGGCTTGCATATCG    556
                   N  V  T  V  Q  R  S  A  I  V  D  G  G  L  H  I
AJ011092 prnA(2): CAATGTCACGGTCCAACGCAGCGCCATCGTCGACGGGGGCTTGCATATCG
AF456356 prnA(9): CAATGTCACGGTCCAACGCAGCGCCATCGTCGACGGGGGCTTGCATATCG
AF218785 prnA(9): CAATGTCACGGTCCAACGCAGCGCCATCGTCGACGGGGGCTTGCATATCG

       CONSENSUS: GCGCCCTGCAGTCATTGCAGCCGGAAGACCTTCCGCCCAGCCGGGTGGTG    606
                  G  A  L  Q  S  L  Q  P  E  D  L  P  P  S  R  V  V
AJ011092 prnA(2): GCGCCCTGCAGTCATTGCAGCCGGAAGACCTTCCGCCCAGCCGGGTGGTG
AF456356 prnA(9): GCGCCCTGCAGTCATTGCAGCCGGAAGACCTTCCGCCCAGCCGGGTGGTG
AF218785 prnA(9): GCGCCCTGCAGTCATTGCAGCCGGAAGACCTTCCGCCCAGCCGGGTGGTG

       CONSENSUS: CTGCGCGACACCAACGTGACCGCCGTGCCCGCCAGCGGCGCGCCCGCGGC    656
                   L  R  D  T  N  V  T  A  V  P  A  S  G  A  P  A  A
AJ011092 prnA(2): CTGCGCGACACCAACGTGACCGCCGTGCCCGCCAGCGGCGCGCCCGCGGC
AF456356 prnA(9): CTGCGCGACACCAACGTGACCGCCGTGCCCGCCAGCGGCGCGCCCGCGGC
AF218785 prnA(9): CTGCGCGACACCAACGTGACCGCCGTGCCCGCCAGCGGCGCGCCCGCGGC

       CONSENSUS: GGTGTCTGTGTTGGGGGCCAGTGAGCTTACGCTCGACGGCGGGCACATCA    706
                    V  S  V  L  G  A  S  E  L  T  L  D  G  G  H  I
AJ011092 prnA(2): GGTGTCTGTGTTGGGGGCCAGTGAGCTTACGCTCGACGGCGGGCACATCA
AF456356 prnA(9): GGTGTCTGTGTTGGGGGCCAGTGAGCTTACGCTCGACGGCGGGCACATCA
AF218785 prnA(9): GGTGTCTGTGTTGGGGGCCAGTGAGCTTACGCTCGACGGCGGGCACATCA

       CONSENSUS: CCGGCGGGCGGGCAGCGGGGGTGGCGGCCATGCAAGGGGCGGTCGTGCAT    756
                  T  G  G  R  A  A  G  V  A  A  M  Q  G  A  V  V  H
AJ011092 prnA(2): CCGGCGGGCGGGCAGCGGGGGTGGCGGCCATGCAAGGGGCGGTCGTGCAT
AF456356 prnA(9): CCGGCGGGCGGGCAGCGGGGGTGGCGGCCATGCAAGGGGCGGTCGTGCAT
AF218785 prnA(9): CCGGCGGGCGGGCAGCGGGGGTGGCGGCCATGCAAGGGGCGGTCGTGCAT

       CONSENSUS:CTGCAGCGCGCGACGATACGGCGCGGGGACGCGCCTGCC GGCGGTGCGGT    806
                  L  Q  R  A  T  I  R  R  G  D  A  P  A   G  G  A  V
AJ011092 prnA(2):CTGCAGCGCGCGACGATACGGCGCGGGGACGCGCCTGCC GGCGGTGCGGT
AF456356 prnA(9):CTGCAGCGCGCGACGATACGGCGCGGGGACGCGCCTGCC GGCGGTGCGGT
AF218785 prnA(9):CTGCAGCGCGCGACGATACGGCGCGGGGACGCGCCTGCC GGCGGTGCGGT

       CONSENSUS:TCCC GGCGGTGCGGTTCCC GGCGGCTTCGGTCCC GGCGGCTTCGGTCCC G    856
                   P  G  G  A  V  P  G  G  F  G  P  G  G  F  G  P
AJ011092 prnA(2):TCCC GGCGGTGCGGTTCCC GGCGGCTTCGGTCCC GGCGGCTTCGGTCCC G
AF456356 prnA(9):TCCC GGCGGTGCGGTTCCC GGCGGCTTCGGTCCC GGCGGCTTCGGTCCC G
AF218785 prnA(9):TCCC GGCGGTGCGGTTCCC GGCGGCTTCGGTCCC GGCGGCTTCGGTCCC G

       CONSENSUS:GCGGCTTCGGTCCC GGCGGCTTCGGTCCC GGCGGCTTCGGTCCC GTCCTC    906
                 G  G  F  G  P  G  G  F  G  P  G  G  F  G  P  V  L
AJ011092 prnA(2):GCGGCTTCGGTCCC GGCGGCTTCGGTCCC --------------- GTCCTC
AF456356 prnA(9):GCGGCTTCGGTCCC GGCGGCTTCGGTCCC GGCGGCTTCGGTCCC GTCCTC
AF218785 prnA(9):GCGGCTTCGGTCCC GGCGGCTTCGGTCCC GGCGGCTTCGGTCCC GTCCTC

       CONSENSUS: GACGGCTGGTATGGCGTGGACGTATCGGGCTCCAGCGTGGAGCTCGCCCA    956
                   D  G  W  Y  G  V  D  V  S  G  S  S  V  E  L  A  Q
AJ011092 prnA(2): GACGGCTGGTATGGCGTGGACGTATCGGGCTCCAGCGTGGAGCTCGCCCA
AF456356 prnA(9): GACGGCTGGTATGGCGTGGACGTATCGGGCTCCAGCGTGGAGCTCGCCCA
AF218785 prnA(9): GACGGCTGGTATGGCGTGGACGTATCGGGCTCCAGCGTGGAGCTCGCCCA
```

Figure 4.2. Nucleic acid alignment of the *Bordetella pertussis* pertactin gene types, *prnA*(2) and *prnA*(9), continued

```
         CONSENSUS: GTCGATCGTCGAGGCGCCGGAGCTGGGCGCCGCAATCCGGGTGGCCGCG    1006
                     S  I  V  E  A  P  E  L  G  A  A  I  R  V  G  R
AJ011092 prnA(2):   GTCGATCGTCGAGGCGCCGGAGCTGGGCGCCGCAATCCGGGTGGCCGCG
AF456356 prnA(9):   GTCGATCGTCGAGGCGCCGGAGCTGGGCGCCGCAATCCGGGTGGCCGCG
AF218785 prnA(9):   GTCGATCGTCGAGGCGCCGGAGCTGGGCGCCGCAATCCGGGTGGCCGCG

         CONSENSUS: GCGCCAGGGTGACGGTGTCGGGCGGCAGCTTGTCCGCACCGCACGGCAAT   1056
                    G  A  R  V  T  V  S  G  G  S  L  S  A  P  H  G  N
AJ011092 prnA(2):   GCGCCAGGGTGACGGTGTCGGGCGGCAGCTTGTCCGCACCGCACGGCAAT
AF456356 prnA(9):   GCGCCAGGGTGACGGTGTCGGGCGGCAGCTTGTCCGCACCGCACGGCAAT
AF218785 prnA(9):   GCGCCAGGGTGACGGTGTCGGGCGGCAGCTTGTCCGCACCGCACGGCAAT

         CONSENSUS: GTCATCGAGACCGGCGGCGCGCGTCGCTTTGCGCCTCAAGCCGCGCCCCT   1106
                    V  I  E  T  G  G  A  R  R  F  A  P  Q  A  A  P  L
AJ011092 prnA(2):   GTCATCGAGACCGGCGGCGCGCGTCGCTTTGCGCCTCAAGCCGCGCCCCT
AF456356 prnA(9):   GTCATCGAGACCGGCGGCGCGCGTCGCTTTGCGCCTCAAGCCGCGCCCCT
AF218785 prnA(9):   GTCATCGAGACCGGCGGCGCGCGTCGCTTTGCGCCTCAAGCCGCGCCCCT

         CONSENSUS: GTCGATCACCTTGCAGGCCGGCGCGCATGCCCAGGGGAAAGCGCTGCTGT   1156
                     S  I  T  L  Q  A  G  A  H  A  Q  G  K  A  L  L
AJ011092 prnA(2):   GTCGATCACCTTGCAGGCCGGCGCGCATGCCCAGGGGAAAGCGCTGCTGT
AF456356 prnA(9):   GTCGATCACCTTGCAGGCCGGCGCGCATGCCCAGGGGAAAGCGCTGCTGT
AF218785 prnA(9):   GTCGATCACCTTGCAGGCCGGCGCGCATGCCCAGGGGAAAGCGCTGCTGT

         CONSENSUS: ACCGGGTCCTGCCGGAGCCCGTGAAGCTGACGCTGACCGGGGGCGCCGAT   1206
                    Y  R  V  L  P  E  P  V  K  L  T  L  T  G  G  A  D
AJ011092 prnA(2):   ACCGGGTCCTGCCGGAGCCCGTGAAGCTGACGCTGACCGGGGGCGCCGAT
AF456356 prnA(9):   ACCGGGTCCTGCCGGAGCCCGTGAAGCTGACGCTGACCGGGGGCGCCGAT
AF218785 prnA(9):   ACCGGGTCCTGCCGGAGCCCGTGAAGCTGACGCTGACCGGGGGCGCCGAT

         CONSENSUS: GCGCAGGGCGACATCGTCGCGACGGAGCTGCCCTCCATTCCCGGCACGTC   1256
                    A  Q  G  D  I  V  A  T  E  L  P  S  I  P  G  T  S
AJ011092 prnA(2):   GCGCAGGGCGACATCGTCGCGACGGAGCTGCCCTCCATTCCCGGCACGTC
AF456356 prnA(9):   GCGCAGGGCGACATCGTCGCGACGGAGCTGCCCTCCATTCCCGGCACGTC
AF218785 prnA(9):   GCGCAGGGCGACATCGTCGCGACGGAGCTGCCCTCCATTCCCGGCACGTC

         CONSENSUS: GATCGGGCCGCTCGACGTGGCGCTGGCCAGCCAGGCCCGATGGACGGGCG   1306
                     I  G  P  L  D  V  A  L  A  S  Q  A  R  W  T  G
AJ011092 prnA(2):   GATCGGGCCGCTCGACGTGGCGCTGGCCAGCCAGGCCCGATGGACGGGCG
AF456356 prnA(9):   GATCGGGCCGCTCGACGTGGCGCTGGCCAGCCAGGCCCGATGGACGGGCG
AF218785 prnA(9):   GATCGGGCCGCTCGACGTGGCGCTGGCCAGCCAGGCCCGATGGACGGGCG

         CONSENSUS: CTACCCGCGCGGTCGACTCGCTGTCCATCGACAACGCCACCTGGGTCATG   1356
                    A  T  R  A  V  D  S  L  S  I  D  N  A  T  W  V  M
AJ011092 prnA(2):   CTACCCGCGCGGTCGACTCGCTGTCCATCGACAACGCCACCTGGGTCATG
AF456356 prnA(9):   CTACCCGCGCGGTCGACTCGCTGTCCATCGACAACGCCACCTGGGTCATG
AF218785 prnA(9):   CTACCCGCGCGGTCGACTCGCTGTCCATCGACAACGCCACCTGGGTCATG

         CONSENSUS: ACGGACAACTCGAACGTCGGTGCGCTACGGCTGGCCAGCGACGGCAGCGT   1406
                     T  D  N  S  N  V  G  A  L  R  L  A  S  D  G  S  V
AJ011092 prnA(2):   ACGGACAACTCGAACGTCGGTGCGCTACGGCTGGCCAGCGACGGCAGCGT
AF456356 prnA(9):   ACGGACAACTCGAACGTCGGTGCGCTACGGCTGGCCAGCGACGGCAGCGT
AF218785 prnA(9):   ACGGACAACTCGAACGTCGGTGCGCTACGGCTGGCCAGCGACGGCAGCGT

         CONSENSUS: CGATTTCCAGCAGCCGGCCGAAGCTGGGCGGTTCAAGGTCCTGACGGTCA   1456
                     D  F  Q  Q  P  A  E  A  G  R  F  K  V  L  T  V
AJ011092 prnA(2):   CGATTTCCAGCAGCCGGCCGAAGCTGGGCGGTTCAAGGTCCTGACGGTCA
AF456356 prnA(9):   CGATTTCCAGCAGCCGGCCGAAGCTGGGCGGTTCAAGGTCCTGACGGTCA
AF218785 prnA(9):   CGATTTCCAGCAGCCGGCCGAAGCTGGGCGGTTCAAGGTCCTGACGGTCA
```

Figure 4.2. Nucleic acid alignment of the *Bordetella pertussis* pertactin gene types, *prnA*(2) and *prnA*(9), continued

```
       CONSENSUS: ATACGCTGGCGGGTTCGGGGCTGTTCCGCATGAATGTCTTCGCGGACCTG   1506
                  N  T  L  A  G  S  G  L  F  R  M  N  V  F  A  D  L
AJ011092 prnA(2): ATACGCTGGCGGGTTCGGGGCTGTTCCGCATGAATGTCTTCGCGGACCTG
AF456356 prnA(9): ATACGCTGGCGGGTTCGGGGCTGTTCCGCATGAATGTCTTCGCGGACCTG
AF218785 prnA(9): ATACGCTGGCGGGTTCGGGGCTGTTCCGCATGAATGTCTTCGCGGACCTG

       CONSENSUS: GGGCTGAGCGACAAGCTGGTCGTCATGCAGGACGCCAGCGGCCAGCACAG   1556
                  G  L  S  D  K  L  V  V  M  Q  D  A  S  G  Q  H  R
AJ011092 prnA(2): GGGCTGAGCGACAAGCTGGTCGTCATGCAGGACGCCAGCGGCCAGCACAG
AF456356 prnA(9): GGGCTGAGCGACAAGCTGGTCGTCATGCAGGACGCCAGCGGCCAGCACAG
AF218785 prnA(9): GGGCTGAGCGACAAGCTGGTCGTCATGCAGGACGCCAGCGGCCAGCACAG

       CONSENSUS: GCTGTGGGTCCGCAACAGCGGCAGCGAGCCGGCCAGCGCCAACACCCTGC   1606
                   L  W  V  R  N  S  G  S  E  P  A  S  A  N  T  L
AJ011092 prnA(2): GCTGTGGGTCCGCAACAGCGGCAGCGAGCCGGCCAGCGCCAACACCCTGC
AF456356 prnA(9): GCTGTGGGTCCGCAACAGCGGCAGCGAGCCGGCCAGCGCCAACACCCTGC
AF218785 prnA(9): GCTGTGGGTCCGCAACAGCGGCAGCGAGCCGGCCAGCGCCAACACCCTGC

       CONSENSUS: TGCTGGTGCAGACGCCACTAGGCAGCGCGGCGACCTTTACCCTTGCCAAC   1656
                  L  L  V  Q  T  P  L  G  S  A  A  T  F  T  L  A  N
AJ011092 prnA(2): TGCTGGTGCAGACGCCACTAGGCAGCGCGGCGACCTTTACCCTTGCCAAC
AF456356 prnA(9): TGCTGGTGCAGACGCCACTAGGCAGCGCGGCGACCTTTACCCTTGCCAAC
AF218785 prnA(9): TGCTGGTGCAGACGCCACTAGGCAGCGCGGCGACCTTTACCCTTGCCAAC

       CONSENSUS: AAGGACGGCAAGGTCGATATCGGTACCTATCGCTATCGATTGGCCGCCAA   1706
                   K  D  G  K  V  D  I  G  T  Y  R  Y  R  L  A  A  N
AJ011092 prnA(2): AAGGACGGCAAGGTCGATATCGGTACCTATCGCTATCGATTGGCCGCCAA
AF456356 prnA(9): AAGGACGGCAAGGTCGATATCGGTACCTATCGCTATCGATTGGCCGCCAA
AF218785 prnA(9): AAGGACGGCAAGGTCGATATCGGTACCTATCGCTATCGATTGGCCGCCAA

       CONSENSUS: CGGCAATGGGCAGTGGAGCCTGGTGGGCGCGAAGGCGCCGCCGGCGCCCA   1756
                    G  N  G  Q  W  S  L  V  G  A  K  A  P  P  A  P
AJ011092 prnA(2): CGGCAATGGGCAGTGGAGCCTGGTGGGCGCGAAGGCGCCGCCGGCGCCCA
AF456356 prnA(9): CGGCAATGGGCAGTGGAGCCTGGTGGGCGCGAAGGCGCCGCCGGCGCCCA
AF218785 prnA(9): CGGCAATGGGCAGTGGAGCCTGGTGGGCGCGAAGGCGCCGCCGGCGCCCA

       CONSENSUS: AGCCCGCGCCGCAGCCGGGTCCCCAGCCGCCGCAGCCGCCGCAGCCGCAG   1806
                  K  P  A  P  Q  P  G  P  Q  P  P  Q  P  P  Q  P  Q
AJ011092 prnA(2): AGCCCGCGCCGCAGCCGGGTCCCCAGCCGCCGCAGCCGCCGCAGCCGCAG
AF456356 prnA(9): AGCCCGCGCCGCAGCCGGGTCCCCAGCCGCCGCAGCCGCCGCAGCCGCAG
AF218785 prnA(9): AGCCCGCGCCGCAGCCGGGTCCCCAGCCGCCGCAGCCGCCGCAGCCGCAG

       CONSENSUS: CCGGAAGCGCCGGCGCCGCAACCGCCGGCGGGCAGGGAGTTGTCCGCCGC   1856
                   P  E  A  P  A  P  Q  P  P  A  G  R  E  L  S  A  A
AJ011092 prnA(2): CCGGAAGCGCCGGCGCCGCAACCGCCGGCGGGCAGGGAGTTGTCCGCCGC
AF456356 prnA(9): CCGGAAGCGCCGGCGCCGCAACCGCCGGCGGGCAGGGAGTTGTCCGCCGC
AF218785 prnA(9): CCGGAAGCGCCGGCGCCGCAACCGCCGGCGGGCAGGGAGTTGTCCGCCGC

       CONSENSUS: CGCCAACGCGGCGGTCAACACGGGTGGGGTGGGCCTGGCCAGCACGCTCT   1906
                    A  N  A  A  V  N  T  G  G  V  G  L  A  S  T  L
AJ011092 prnA(2): CGCCAACGCGGCGGTCAACACGGGTGGGGTGGGCCTGGCCAGCACGCTCT
AF456356 prnA(9): CGCCAACGCGGCGGTCAACACGGGTGGGGTGGGCCTGGCCAGCACGCTCT
AF218785 prnA(9): CGCCAACGCGGCGGTCAACACGGGTGGGGTGGGCCTGGCCAGCACGCTCT

       CONSENSUS: GGTACGCCGAAAGCAATGCGTTGTCCAAGCGCCTGGGCGAGTTGCGCCTG   1956
                  W  Y  A  E  S  N  A  L  S  K  R  L  G  E  L  R  L
AJ011092 prnA(2): GGTACGCCGAAAGCAATGCGTTGTCCAAGCGCCTGGGCGAGTTGCGCCTG
AF456356 prnA(9): GGTACGCCGAAAGCAATGCGTTGTCCAAGCGCCTGGGCGAGTTGCGCCTG
AF218785 prnA(9): GGTACGCCGAAAGCAATGCGTTGTCCAAGCGCCTGGGCGAGTTGCGCCTG
```

Figure 4.2. Nucleic acid alignment of the *Bordetella pertussis* pertactin gene types, *prnA*(2) and *prnA*(9), continued

```
       CONSENSUS: AATCCGGACGCCGGCGGCGCCTGGGGCCGCGGCTTCGCGCAACGCCAGCA   2006
                   N  P  D  A  G  G  A  W  G  R  G  F  A  Q  R  Q  Q
AJ011092 prnA(2): AATCCGGACGCCGGCGGCGCCTGGGGCCGCGGCTTCGCGCAACGCCAGCA
AF456356 prnA(9): AATCCGGACGCCGGCGGCGCCTGGGGCCGCGGCTTCGCGCAACGCCAGCA
AF218785 prnA(9): AATCCGGACGCCGGCGGCGCCTGGGGCCGCGGCTTCGCGCAACGCCAGCA

       CONSENSUS: GCTGGACAACCGCGCCGGGCGGCGCTTCGACCAGAAGGTGGCCGGCTTCG   2056
                    L  D  N  R  A  G  R  R  F  D  Q  K  V  A  G  F
AJ011092 prnA(2): GCTGGACAACCGCGCCGGGCGGCGCTTCGACCAGAAGGTGGCCGGCTTCG
AF456356 prnA(9): GCTGGACAACCGCGCCGGGCGGCGCTTCGACCAGAAGGTGGCCGGCTTCG
AF218785 prnA(9): GCTGGACAACCGCGCCGGGCGGCGCTTCGACCAGAAGGTGGCCGGCTTCG

       CONSENSUS: AGCTGGGCGCCGACCACGCGGTGGCGGTGGCCGGCGGACGCTGGCACCTG   2106
                  E  L  G  A  D  H  A  V  A  V  A  G  G  R  W  H  L
AJ011092 prnA(2): AGCTGGGCGCCGACCACGCGGTGGCGGTGGCCGGCGGACGCTGGCACCTG
AF456356 prnA(9): AGCTGGGCGCCGACCACGCGGTGGCGGTGGCCGGCGGACGCTGGCACCTG
AF218785 prnA(9): AGCTGGGCGCCGACCACGCGGTGGCGGTGGCCGGCGGACGCTGGCACCTG

       CONSENSUS: GGCGGGCTGGCCGGCTATACGCGCGGCGACCGCGGCTTCACCGGCGACGG   2156
                   G  G  L  A  G  Y  T  R  G  D  R  G  F  T  G  D  G
AJ011092 prnA(2): GGCGGGCTGGCCGGCTATACGCGCGGCGACCGCGGCTTCACCGGCGACGG
AF456356 prnA(9): GGCGGGCTGGCCGGCTATACGCGCGGCGACCGCGGCTTCACCGGCGACGG
AF218785 prnA(9): GGCGGGCTGGCCGGCTATACGCGCGGCGACCGCGGCTTCACCGGCGACGG

       CONSENSUS: CGGCGGCCACACCGACAGCGTGCATGTCGGGGGCTATGCCACATATATCG   2206
                    G  G  H  T  D  S  V  H  V  G  G  Y  A  T  Y  I
AJ011092 prnA(2): CGGCGGCCACACCGACAGCGTGCATGTCGGGGGCTATGCCACATATATCG
AF456356 prnA(9): CGGCGGCCACACCGACAGCGTGCATGTCGGGGGCTATGCCACATATATCG
AF218785 prnA(9): CGGCGGCCACACCGACAGCGTGCATGTCGGGGGCTATGCCACATATATCG

       CONSENSUS: CCGACAGCGGTTTCTACCTGGACGCGACGCTGCGCGCCAGCCGCCTGGAG   2256
                  A  D  S  G  F  Y  L  D  A  T  L  R  A  S  R  L  E
AJ011092 prnA(2): CCGACAGCGGTTTCTACCTGGACGCGACGCTGCGCGCCAGCCGCCTGGAG
AF456356 prnA(9): CCGACAGCGGTTTCTACCTGGACGCGACGCTGCGCGCCAGCCGCCTGGAG
AF218785 prnA(9): CCGACAGCGGTTTCTACCTGGACGCGACGCTGCGCGCCAGCCGCCTGGAG

       CONSENSUS: AATGACTTCAAGGTGGCGGGCAGCGACGGGTACGCGGTCAAGGGCAAGTA   2306
                   N  D  F  K  V  A  G  S  D  G  Y  A  V  K  G  K  Y
AJ011092 prnA(2): AATGACTTCAAGGTGGCGGGCAGCGACGGGTACGCGGTCAAGGGCAAGTA
AF456356 prnA(9): AATGACTTCAAGGTGGCGGGCAGCGACGGGTACGCGGTCAAGGGCAAGTA
AF218785 prnA(9): AATGACTTCAAGGTGGCGGGCAGCGACGGGTACGCGGTCAAGGGCAAGTA

       CONSENSUS: CCGCACCCATGGGGTGGGCGCCTCGCTCGAGGCGGGCCGGCGCTTTACCC   2356
                    R  T  H  G  V  G  A  S  L  E  A  G  R  R  F  T
AJ011092 prnA(2): CCGCACCCATGGGGTGGGCGCCTCGCTCGAGGCGGGCCGGCGCTTTACCC
AF456356 prnA(9): CCGCACCCATGGGGTGGGCGCCTCGCTCGAGGCGGGCCGGCGCTTTACCC
AF218785 prnA(9): CCGCACCCATGGGGTGGGCGCCTCGCTCGAGGCGGGCCGGCGCTTTACCC

       CONSENSUS: ATGCCGACGGCTGGTTCCTCGAGCCGCAGGCCGAGCTGGCGGTATTCCGG   2406
                  H  A  D  G  W  F  L  E  P  Q  A  E  L  A  V  F  R
AJ011092 prnA(2): ATGCCGACGGCTGGTTCCTCGAGCCGCAGGCCGAGCTGGCGGTATTCCGG
AF456356 prnA(9): ATGCCGACGGCTGGTTCCTCGAGCCGCAGGCCGAGCTGGCGGTATTCCGG
AF218785 prnA(9): ATGCCGACGGCTGGTTCCTCGAGCCGCAGGCCGAGCTGGCGGTATTCCGG

       CONSENSUS: GCCGGCGGCGGTGCGTACCGCGCGGCCAACGGCCTGCGGGTGCGCGACGA   2456
                   A  G  G  G  A  Y  R  A  A  N  G  L  R  V  R  D  E
AJ011092 prnA(2): GCCGGCGGCGGTGCGTACCGCGCGGCCAACGGCCTGCGGGTGCGCGACGA
AF456356 prnA(9): GCCGGCGGCGGTGCGTACCGCGCGGCCAACGGCCTGCGGGTGCGCGACGA
AF218785 prnA(9): GCCGGCGGCGGTGCGTACCGCGCGGCCAACGGCCTGCGGGTGCGCGACGA
```

Figure 4.2. Nucleic acid alignment of the *Bordetella pertussis* pertactin gene types, *prnA*(2) and *prnA*(9), continued

```
       CONSENSUS: AGGCGGCAGCTCGGTGCTGGGTCGCCTGGGCCTGGAGGTCGGCAAGCGCA    2506
                   G  G  S  S  V  L  G  R  L  G  L  E  V  G  K  R
AJ011092 prnA(2): AGGCGGCAGCTCGGTGCTGGGTCGCCTGGGCCTGGAGGTCGGCAAGCGCA
AF456356 prnA(9): AGGCGGCAGCTCGGTGCTGGGTCGCCTGGGCCTGGAGGTCGGCAAGCGCA
AF218785 prnA(9): AGGCGGCAGCTCGGTGCTGGGTCGCCTGGGCCTGGAGGTCGGCAAGCGCA

       CONSENSUS: TCGAACTGGCAGGCGGCAGGCAGGTGCAGCCATACATCAAGGCCAGCGTG    2556
                  I  E  L  A  G  G  R  Q  V  Q  P  Y  I  K  A  S  V
AJ011092 prnA(2): TCGAACTGGCAGGCGGCAGGCAGGTGCAGCCATACATCAAGGCCAGCGTG
AF456356 prnA(9): TCGAACTGGCAGGCGGCAGGCAGGTGCAGCCATACATCAAGGCCAGCGTG
AF218785 prnA(9): TCGAACTGGCAGGCGGCAGGCAGGTGCAGCCATACATCAAGGCCAGCGTG

       CONSENSUS: CTGCAGGAGTTCGACGGCGCGGGTACGGTACACACCAACGGCATCGCGCA    2606
                   L  Q  E  F  D  G  A  G  T  V  H  T  N  G  I  A  H
AJ011092 prnA(2): CTGCAGGAGTTCGACGGCGCGGGTACGGTACACACCAACGGCATCGCGCA
AF456356 prnA(9): CTGCAGGAGTTCGACGGCGCGGGTACGGTACACACCAACGGCATCGCGCA
AF218785 prnA(9): CTGCAGGAGTTCGACGGCGCGGGT--------------------------

       CONSENSUS: CCGCACCGAACTGCGCGGCACGCGCGCCGAACTGGGCCTGGGCATGGCCG    2656
                    R  T  E  L  R  G  T  R  A  E  L  G  L  G  M  A
AJ011092 prnA(2): CCGCACCGAACTGCGCGGCACGCGCGCCGAACTGGGCCTGGGCATGGCCG
AF456356 prnA(9): CCGCACCGAACTGCGCGGCACGCGCGCCGAACTGGGCCTGGGCATGGCCG
AF218785 prnA(9): --------------------------------------------------

       CONSENSUS: CCGCGCTGGGCCGCGGCCACAGCCTGTATGCCTCGTACGAGTACTCCAAG    2706
                  A  A  L  G  R  G  H  S  L  Y  A  S  T  E  Y  S  K
AJ011092 prnA(2): CCGCGCTGGGCCGCGGCCACAGCCTGTATGCCTCGTACGAGTACTCCAAG
AF456356 prnA(9): CCGCGCTGGGCCGCGGCCACAGCCTGTATGCCTCGTACGAGTACTCCAAG
AF218785 prnA(9): --------------------------------------------------

       CONSENSUS: GGCCCGAAGCTGGCCATGCCGTGGACCTTCCACGCGGGCTACCGGTACAG    2756
                    G  P  K  L  A  M  P  W  T  F  H  A  G  Y  R  Y  S
AJ011092 prnA(2): GGCCCGAAGCTGGCCATGCCGTGGACCTTCCACGCGGGCTACCGGTACAG
AF456356 prnA(9): GGCCCGAAGCTGGCCATGCCGTGGACCTTCCACGCGGGCTACCGGTACAG
AF218785 prnA(9): --------------------------------------------------

       CONSENSUS: CTGGTAAAGCGAGGAGGGTCTATCCCCCGCGGAGGAGGTTTTCCTGGAGC    2763
                   W  *
AJ011092 prnA(2): CTGGTAAAGCGAGGAGGGTCTATCCCCCGCGGAGGAGGTTTTCCTGGAGC
AF456356 prnA(9): CTGGTAAAGCGAGGAGGGTCTATCCCCCGCGGAGGAGGTTTTC-------
AF218785 prnA(9): --------------------------------------------------

       CONSENSUS: TT
AJ011092 prnA(2): TT
AF456356 prnA(9): --
AF218785 prnA(9): --
```

The GenBank sequence, AJ011092, a *prnA*(2) strain, was used to align the sequences generated from MO-121 (AF456356) and MO-908 (AF218785). The amino acids are coded as the universal single letter designations (Appendix 1), and are shown in black underneath the nucleotide coding sequence. The black horizontal line and black asterix (*) indicate the start and stop codons of pertactin, respectively. Dashed lines indicate that the sequence is not found. The boxed area encompasses region 1 of the pertactin gene, where both *prnA*(9) sequences possess an extra repeat.

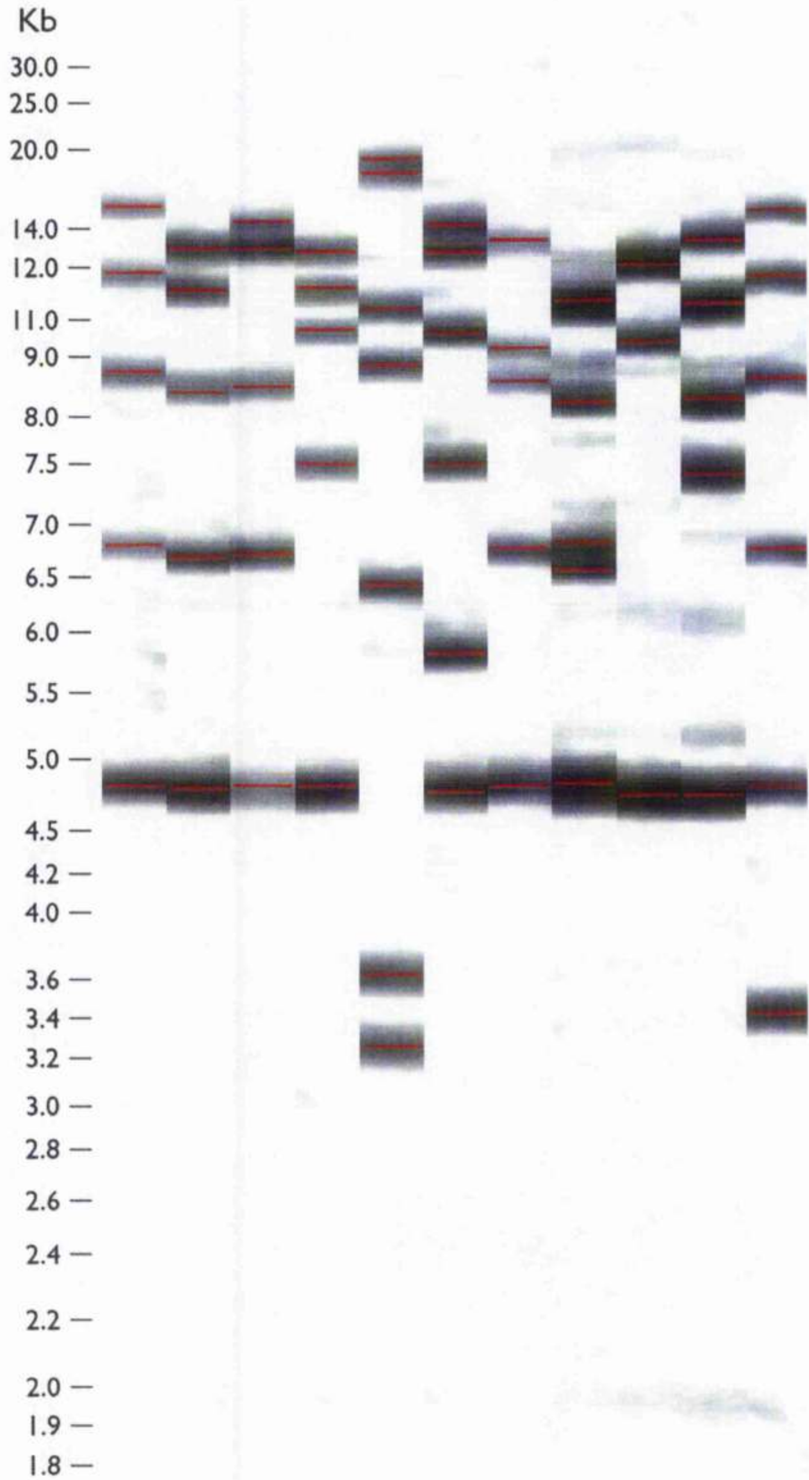
Six *ptxA* types have been described to date, *ptxA*(1), *ptxA*(2), *ptxA*(3) and *ptxA*(4) (Mooi *et al.*, 2000), and *ptxA*(5) and *ptxA*(6) (Poynten *et al.*, 2004) (Figure 1.4, page 48), but only two types, *ptxA*(1) and *ptxA*(2) were found in the Missouri *B. pertussis* isolates (Table 4.1, page 167). The *ptxA*(1) variant dominated the *B. pertussis* population, occurring at a frequency of 80% (12/15) in 1984-1989 and 95% (18/19) in 1996-2001 (Fisher's exact test, P=0.299, not significantly different). All "Kansas City" isolates were *ptxA*(1).

The 44 isolates generated four *prnA* types and two *ptxA* types and these were combined to give a definitive *prnA*/*ptxA* allele. Out of eight possible *prnA*/*ptxA* combinations, only five were found; *prnA*(1)/*ptxA*(1), *prnA*(1)/*ptxA*(2), *prnA*(2)/*ptxA*(1), *prnA*(6)/*ptxA*(2), and *prnA*(9)/*ptxA*(1), (Table 4.2, page 167). All five of these combinations were found from 1984 to 1989, and type *prnA*(1)/*ptxA*(1) was present in seven of the 15 strains, (47%). Type *prnA*(2)/*ptxA*(1) has been seen almost exclusively from 1996 to 2001, and found in 18/19 (95%) of the isolates, with the remaining isolate designated as *prnA*(1)/*ptxA*(2). This shift towards *prnA*(2)/*ptxA*(1) isolates is very significant (Fisher's exact test, P<0.001). The ten isolates from Kansas City in 1999 were all *prnA*(2)/*ptxA*(1).

4.3.3 IS*1002*-Restriction Fragment Length Polymorphism analysis

A total of 11 different IS*1002*-RFLP types were generated from the 44 isolates after normalisation using BioNumerics (Applied Maths) (Figure 4.3, page 175). The number of bands per RFLP type varied between three and seven, and most RFLP types (7/11) contained five. When grouped temporally, 15 isolates generated ten different RFLP types from 1984-1989 and 19 isolates from 1996-2001 produced four RFLP types (Table 4.3, page 176). The most frequent RFLP type, RFLP-1, was found in 29 (66%) of the total isolates tested, and was seen in six of the 15 isolates (40%) from 1984-1989, and 13 of 19 (68%) from 1996-2001 (Table 4.3, page 176). The second most frequently seen type, RFLP-2 occurred at a much lower frequency of 5/44 (11%) for the total number of isolates tested. When divided into the two subsets, RFLP-2 is present in one isolate of the 15 (7%) from 1984-1989 and four of the 19 (21%) isolates from 1996-2001. Of the 11 RFLP types seen, eight comprised one isolate each, where seven of the eight types were present in the years 1984-1989. All ten "Kansas City" isolates were of the same RFLP type, RFLP-1.

Figure 4.3. IS*1002*-RFLP types of *Bordetella pertussis* isolates from Missouri, USA



The thin red lines indicate bands included in the analysis, as normalised using BioNumerics (Applied Maths).

Table 4.3. Temporal distribution of IS*1002*-RFLP types in *Bordetella pertussis* isolates from Missouri, USA

IS*1002*-RFLP type	"Early", 1984-1989	"Recent", 1996-2001	"Kansas City", 1999
1	6	13	10
2	1	4	
11	1		
21	1		
25	1		
26	1		
27	1	1	
28	1		
31	1		
32	1		
33		1	
Total isolates	15	19	10

Table 4.4. Temporal distribution of PFGE types, produced with *Xba*I, in *Bordetella pertussis* isolates from Missouri, USA

PFGE type	"Early", 1984-1989	"Recent", 1996-2001	"Kansas City", 1999
1	5		
2		10	5
3		4	5
4		1	
6	1		
9	2		
10	1	1	
14	2		
15		2	
17	1		
21		1	
24	1		
27	1		
31	1		
Total isolates	15	19	10

Numbers of isolates are shown. The "Kansas City" data are in grey, as the isolates were not included in the temporal analysis.

4.3.4 Pulsed-Field Gel Electrophoresis analysis

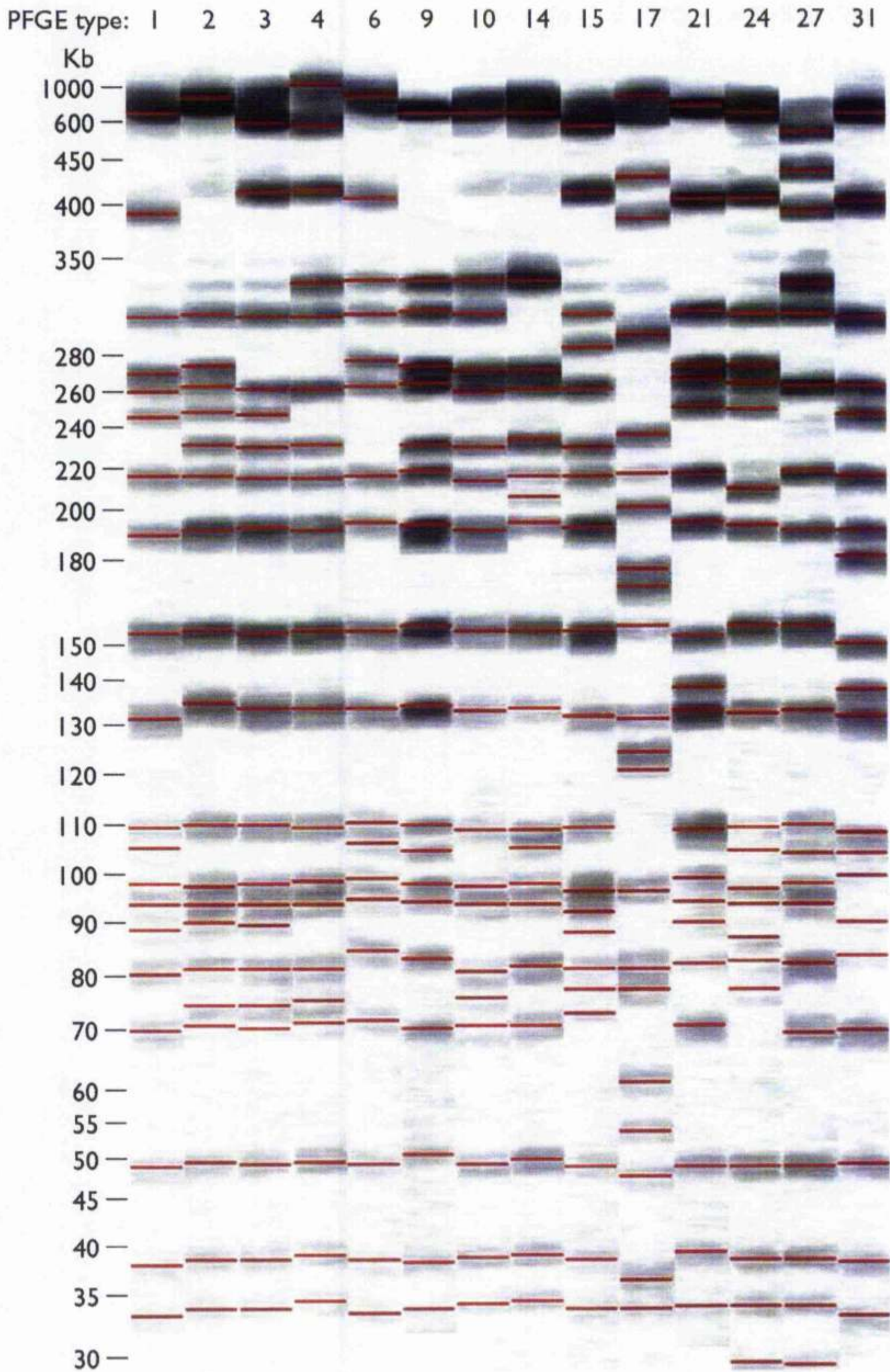
The 44 isolates from Missouri, USA, produced 14 different PFGE profiles using *Xba*I (Table 4.4 and Figure 4.4, pages 176 and 178). There were nine different PFGE profiles from the 1984-1989 isolates and six profiles from the 1996-2001 isolates, but only one profile, PFGE-10, was seen in both time periods. Profile PFGE-2 was the dominant type, occurring in 15/44 of the total isolates (34%), and was also present in 5/10 (50%) of the "Kansas City" isolates. Interestingly, serotyping, gene sequencing and IS*1002*-RFLP analysis, grouped all the "Kansas City" isolates as serotype 1,3, *prnA*(2)/*ptxA*(1) and RFLP-1, whereas PFGE divided this subset into two profiles, PFGE-2 (n=5) and PFGE-3 (n=5). These two profiles are closely related to each other and are clustered with another profile PFGE-15, to produce a clade with a similarity of 91.6% (Figure 4.5, page 179). The main difference between PFGE-2 and PFGE-3 is the presence of a 41,696 bp band in PFGE-3, and a 27,576 bp band in PFGE-2. Similar to IS*1002*-RFLP, seven of the 14 PFGE profiles were unique, and five of those were present only in the 1984-1989 isolates.

4.3.5 Discriminatory indices and associations of the various typing methods

In order to calculate the indices of discrimination (*D*) (section 2.8.3, page 99), nine of the "Kansas City" isolates were removed, as these were presumed to be epidemiologically-related. Therefore, 35 isolates were evaluated.

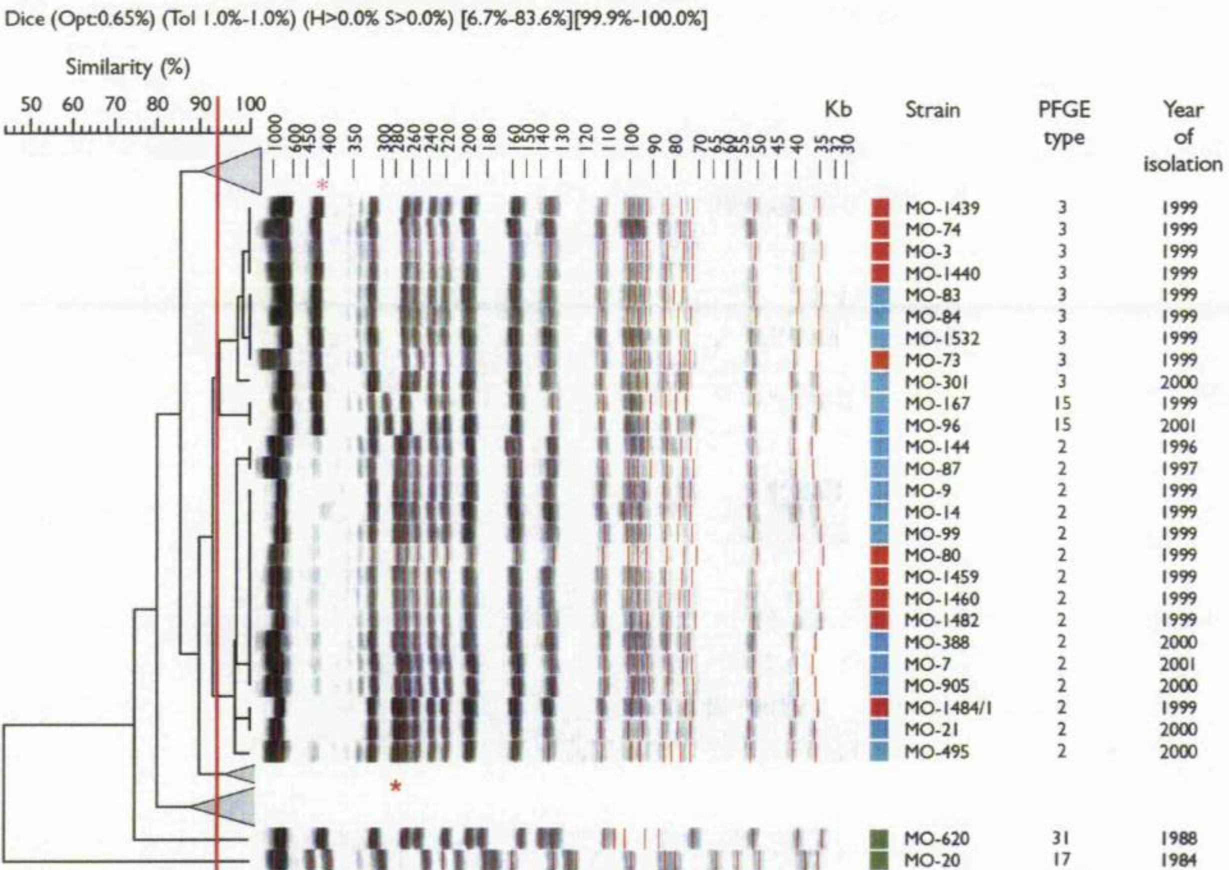
The highest index of discrimination value was calculated for PFGE (D=0.88), and the lowest was D=0.21 for the pertussis toxin S1 gene typing (Table 4.5, page 180). The other *D* values for the other typing methods varied between 0.54 and 0.68 for pertactin gene typing and serotyping, respectively. These values however, are based on only 35 isolates, within a limited time period between 1984 and 2001. Chapter 5 describes a more detailed assessment of the discrimination of these typing methods.

Figure 4.4. PFGE types, with *Xba*I, of *Bordetella pertussis* isolates from Missouri, USA



The thin red lines indicate bands included in the analysis as normalised using BioNumerics (Applied Maths).

Figure 4.5. A dendrogram of *Bordetella pertussis* isolates from Missouri, USA, analysed by PFGE with *Xba*I, focusing on the clade containing PFGE-2 and PFGE-3 isolates



Cluster analysis was performed using the Dice coefficient of similarity and the dendrogram generated by the unweighted pair group method with arithmetic averages (UPGMA). The solid red lines indicate bands included in the analysis. Green boxes denotes 1984-1989 isolates, red denotes "Kansas City 1999" isolates, and blue denotes other isolates collected between 1996-2001. Collapsed branches are unrelated to the cluster in question, which has 91.63% similarity (red vertical line). *, indicates both the 41,696 bp band in PFGE-3 and the 27,576 bp band in PFGE-2, which are the main differences between these two PFGE types.

Table 4.5. Indices of discrimination of the typing methods performed on the 43 *Bordetella pertussis* isolates from Missouri, USA

Typing method	Number of types	Index of discrimination (*D*)
Serotyping	4	0.68
Pertactin gene typing	4	0.54
Pertussis toxin S1 gene typing	2	0.21
prnA/*ptxA* combination	5	0.58
IS*1002*-RFLP analysis	11	0.66
PFGE using *Xba*I	14	0.88

The Missouri isolates were predominantly serotype 1,3, (39/44), with 28/31 *prnA*(2) isolates, 25/29 RFLP-1 isolates and 35/40 *ptxA*(1) isolates associated with this serotype. Associations between RFLP types and pertactin and pertussis toxin S1 gene variants were also investigated. All 29 RFLP-1 isolates were type *ptxA*(1), and 25 of these were *prnA*(2). The remaining four RFLP-1 isolates were either pertactin variants *prnA*(1) or *prnA*(9), and present in the more diverse period of 1984-1989. Thirteen of the 19 isolates from 1996 to 2001 were RFLP-1 and *prnA*(2)/*ptxA*(1), and in addition, all the "Kansas City" isolates (n=10) were also RFLP-1, *prnA*(2)/*ptxA*(1) and serotype 1,3. Associations between PFGE type and pertactin and pertussis toxin S1 gene variants were also examined. All PFGE profiles with more than one isolate expressed the same *prn*A type, but PFGE-1 isolates exhibited either *ptxA*(1) or *ptxA*(2). The two predominant PFGE profiles, PFGE-2 (n=15) and PFGE-3 (n=9), were *prnA*(2)/*ptxA*(1), and all the isolates within these two PFGE types were recovered between 1996 and 2001.

Associations of the isolates between IS*1002*-RFLP and PFGE types were also examined. Four PFGE types with more than one isolate had unique genotypes; PFGE-3, RFLP-1, *prnA*(2)/*ptxA*(1) (n=9); PFGE-9, RFLP-1, *prnA*(2)/*ptxA*(1) (n=2); PFGE-14, RFLP-1, *prnA*(9)/*ptxA*(1) (n=2); and PFGE-15, RFLP-2, *prnA*(2)/*ptxA*(1) (n=2). There were three PFGE types that contained isolates with more than one RFLP type, and similarly, 3 RFLP types consisting of more than one PFGE type, indicating no conformity between these two typing methods.

4.4 Discussion

A re-emergence of pertussis in the USA prompted colleagues from the Missouri State Laboratory to send isolates to RSIL for genotyping. In particular, a cluster of *B. pertussis* strains isolated within a six-week period in the summer of 1999 from Kansas City were received by Missouri State Laboratory. The question arose as to whether these isolates were; i) part of a specific outbreak; ii) represented an upsurge of a particular strain, or iii) were part of a general increase in pertussis, caused by a variety of strains due to an unidentified event in pertussis epidemiology. To assess their relatedness, and whether any differences in the *B. pertussis* population had occurred, isolates from 1984 to 1989 and 1999 to 2001 were compared with the isolates from Kansas City.

4.4.1.1 Genotypic diversity of *Bordetella pertussis* isolates from Missouri, USA

A large-scale study of 152 *B. pertussis* strains from at least 12 states in the USA determined the pertactin and pertussis toxin S1 gene type, and PFGE profile (Cassiday *et al.*, 2000). Included in their study was at least one isolate from Missouri, but it is unclear from the paper which *prnA/ptxA* type were designated (*prnA*(1)/*ptxA*(1), *prnA*(1)/*ptxA*(2) or *prnA*(2)/*ptxA*(1)). Although the study had a wider spectrum of strains ranging from 1935 to 1999 that originated from the entire country, it did not show a high diversity of strains compared to this study from a single state. Cassiday *et al.* (2000) reported three different *prnA/ptxA* alleles compared to five *prnA/ptxA* types from this study of Missouri isolates in the same time span of 1989-1999. They also labelled certain *prnA/ptxA* alleles as either "old", "new", or "transitional" (section 3.4.3, page 146). This labelling could have been applied to these isolates from Missouri, but it does not explain how other alleles, such as *prnA*(6)/*ptxA*(2), could be placed in this scheme.

During the course of this thesis, a new pertactin type, designated *prnA*(9), was discovered in two isolates, one from 1988, the other in 1989. These differed from *prnA*(2) by an extra GGFGP repeat. It cannot be ascertained whether *prnA*(9) is a "transitional" type or a "new" and recent variant, but a very recent study from Canada also describes this pertactin type, and suggests that *prnA*(9)/*ptxA*(1) is a "new" strain (Peppler *et al.*, 2003). However, *prnA*(9) is, at the time of writing, not as successful as its close counterpart *prnA*(2), which is the present dominant pertactin type in Missouri.

The 44 isolates produced 11 different RFLP types, with a predominance of two types, RFLP-1 and RFLP-2. When divided into the three subsets, the 1984-1989 isolates showed a higher diversity in RFLP types (10/11 types) compared to isolates from 1996-2001 (3/11) and the "Kansas City" isolates (1/11, all RFLP-1). When analysed using PFGE with *Xba*I, a higher diversity was also seen in the 1984-1989 isolates (10/14 PFGE profiles) compared to 6/14 in 1996-2001 and 2/14 in the "Kansas City" isolates. This higher diversity in earlier isolates compared to later ones was also reflected in the serotype, pertactin gene type and the combined *prnA/ptxA* type. The drop in diversity did not coincide with a change of vaccination schedule, a switch from whole-cell vaccine (WCV) to acellular vaccine (ACV), first licensed for infants in the USA in 1996, or a difference in reporting or improved laboratory diagnostic methods. However, it did coincide with the rise of pertussis incidence in Missouri (from 0.46 per 100,000 in 1984 to 1.91 per 100,000 in 2001) (Missouri Department of Health and Senior Services, 2002).

These data from Missouri, USA, do not support the hypothesis of vaccine-driven evolution of *B. pertussis* for a number of reasons. The pertussis vaccines used in the United States since introduction in 1947 have been varied and widely distributed. Evaluation of the USA vaccine strains was confounded when Cassiday *et al.* (2000) investigated, although they reported that the whole-cell vaccines derived from the Tohama-I strain, and all acellular vaccines currently in use, all exhibit type *prnA*(1)/*ptxA*(2). This genotype indeed differs from those of isolates circulating between 1996-2001 in Missouri, which were predominantly *prnA*(2)/*ptxA*(1), but the number of isolates was limited, and genotypes before 1984 cannot be ascertained. Also, the introduction of the ACV in 1996 in Missouri did not coincide with any significant change in pertussis incidence, although the isolates tested between 1996-2001 were almost exclusively serotype 1,3, *prnA*(2)/*ptxA*(1), RFLP-1 and either PFGE-2 or PFGE-3.

4.4.2 Geographical distribution of *Bordetella pertussis* types

Pertactin gene types *prnA*(1), *prnA*(2), *prnA*(6) and *prnA*(9) were found in the isolates from Missouri, USA. The first two types have been seen globally, but *prnA*(6) has been seen only in circulating strains in this study and *prnA*(9) was reported recently in Canada in 41 isolates from 1985 to 1994 (Peppler *et al.*, 2003). This Canadian study also reported that approximately 5% of the 3334 isolates were *prnA*(3) (Peppler *et al.*, 2003). It was previously thought that *prnA*(3) isolates were absent from North America after a large study involving 152 *B. pertussis* isolates collected between 1935 and 1999 from various states of the USA (Cassiday *et al.*, 2000). Results reported here also showed that *prnA*(3) isolates were absent in Missouri state, but *prnA*(6) and *prnA*(9) isolates were present, suggesting some geographical constraint of *B. pertussis*.

Due to the lack of inter-laboratory standardisation of the IS*1002*-RFLP method, the only comparisons that could be performed were with the other data presented in this thesis from the UK, and with the presence of the predominant RFLP type, RFLP-1 in other countries. The temporal distribution of the Missouri RFLP types differed from those in the UK. When comparing the data-sets 1984-1989 and 1996-2001 from Missouri (n=34, Table 4.3, page 176) with the 1980s and 1990s data-sets from the UK (n=72, Figure 3.7, page 118), they both show a fall in diversity in recent years and the expansion of RFLP-1 isolates. A drop in genotypic diversity of *B. pertussis* has also been seen in The Netherlands (van Loo *et al.*, 1999). The successful type RFLP-1, which is found in the UK (Chapter 3), The Netherlands (known as dt-29) (van Loo *et al.*, 1999), Italy (Mastrantonio *et al.*, 1999),

Germany, Canada (known as C2) (van der Zee *et al.*, 1996b) and Australia (Poynten *et al.*, 2004) can be seen as a prolific lineage which accounts for a large percentage of the global *B. pertussis* population.

Similarly to IS*1002*-RFLP, data from pulsed-field gel electrophoresis cannot be compared directly with data from other countries because of differing running conditions and type designations. However, PFGE results from the evaluation panel used in Chapter 5 (Table 5.2, page 193), which were predominantly UK isolates, can be compared with these results from the Missouri isolates. The PFGE types found in both the Missouri and the UK were; PFGE-1, PFGE-3, PFGE-4, PFGE-6, PFGE-9 and PFGE-17 (Figure 4.4, page 178), although the latter type was from the type strain of *B. pertussis* (NCTC 10739^{T}) present in the panel and originally from the USA. The Missouri and the UK isolates both showed a large diversity of PFGE types, but the distributions of PFGE types differed, and although PFGE-2 was present in ten of the recent Missouri isolates, this type was markedly absent in the UK isolates. These PFGE data thus reveal the different geographical distribution of *B. pertussis* isolates.

Interestingly, where there was geographical variation in the pertactin gene, IS*1002*-RFLP and PFGE type distributions, there was no apparent constraint in the serotype or pertussis toxin type. Serotype distribution fluctuates periodically, and serotypes 1,2, 1,3 and 1,2,3 have been seen in the UK and The Netherlands (Miller *et al.*, 1992; Mooi, 1994). The *ptxA* types *ptxA*(1) and *ptxA*(2) are found in most countries that have reported polymorphisms for this gene: The Netherlands (Mooi *et al.*, 1998), Finland (Mooi *et al.*, 1999), the USA (Cassiday *et al.*, 2000), Japan (Guiso *et al.*, 2001a), France (Weber *et al.*, 2001), Poland (Gzyl *et al.*, 2002), Canada (Peppler *et al.*, 2003), Russia (Kourova *et al.*, 2003) and the UK (Chapter 3). The geographical constraint of pertactin, IS*1002*-RFLP and PFGE types, compared to the comparative universal distribution of serotypes and pertussis toxin S1 types, raises the question of how the individual *B. pertussis* genomes and the bacterial population are regulated. In addition, IS*1002*-RFLP analysis and PFGE investigate whole genome changes, serotyping is dependent on the expression of agglutinogens 1, 2 and 3, and pertactin and pertussis toxin S1 types focuses on differences in the chosen genes. Further work is needed regarding the regulatory differences between conservative and variable areas of the genome.

The discriminatory indices of the typing methods from both the Missouri isolates (Table 4.5, page 180) and the UK evaluation panel (Table 5.7, page 211) were compared. Overall,

although the UK evaluation panel was chosen for its diversity, and the Missouri isolates were limited from one state in the USA and taken from 1984 to 2001, the discriminatory indices were very similar for most typing methods. The exceptions were the *ptxA* gene typing (*D*=0.37 cf. *D*=0.21, UK versus Missouri, respectively), and the RFLP typing (*D*=0.85 cf. *D*=0.66, UK versus Missouri, respectively), where the lower discriminatory indices from Missouri were probably due to the limited time period (i.e., no isolates were recovered before 1984).

To ascertain whether the same *B. pertussis* lineages were circulating in the UK and Missouri, associations of method types were compared using data from Chapters 3 and 5. In Missouri, most RFLP-1 isolates were serotype 1,3, *prnA*(2) and *ptxA*(1), and this successful lineage has also appeared in the UK as combined type P (Table 3.14, page 131). This *B. pertussis* lineage is therefore comparatively clonal, and has expanded and spread to distant geographical locations. However, this lineage (combined type P) can be divided into many different PFGE types. Other common *B. pertussis* lineages included: PFGE-3, RFLP-1, *prnA*(2)/*ptxA*(1) and serotype 1,3 isolates (n=2 in the UK and n=7 in Missouri); and PFGE-1, *prnA*(1)/*ptxA*(1) isolates (n=15 in the UK and n=4 from Missouri). Most of the other PFGE or RFLP types were not shared between the UK and Missouri, and *B. pertussis* has therefore diversified in various geographical niches.

4.4.3 *Bordetella pertussis* isolates from Kansas City, Missouri, USA

All of the ten temporally-clustered isolates from Kansas City were serotype 1,3, *prnA*(2)/*ptxA*(1) and RFLP-1, which was also seen in 13 of the 19 isolates from 1996-2001 that were epidemiologically distinct. Therefore it could not be ascertained as to whether the Kansas City isolates were part of an epidemiologically-linked outbreak. However, the Kansas City isolates were split evenly into two groups using PFGE (PFGE-2 and PFGE-3), although the PFGE types cluster at 91.6% similarity, indicating a recent and close evolution of these two lineages. Furthermore, these PFGE types differ by two bands, indicating only one genetic difference between them, and thus being closely related according to Tenover *et al.* (1995). de Moissac *et al.* (1994) studied the epidemiology using PFGE in a three-year epidemic period and found that more than one PFGE type was circulating, but that study was long-term, compared to this study of just two months. Hence, these isolates from Kansas City may not have been from the same clone, but had very similar genotypes. However, both of these PFGE types were also found in the

geographically distinct 1996-2001 isolates, so that the Kansas City isolates are still indistinguishable from contemporary isolates.

The ten Kansas City isolates from 1999 were not derived from one clone or lineage of *B. pertussis*, and the majority of strains around this time were mostly of the same type (serotype 1,3, *prnA*(2)/*ptxA*(1), RFLP-1 and PFGE-2 or PFGE-3). An unknown event might have occurred between 1990-1995 (no isolates available), which led to a sustained rise in pertussis incidence in Missouri. This event may also have seen the emergence and expansion of a few successful genotypes of *B. pertussis*. It appears that more than one strain was circulating concurrently and causing disease.

5 Evaluation of typing methods for *Bordetella pertussis*

The UK and Missouri, USA *B. pertussis* studies in this thesis showed varying discrimination for the typing methods performed (Chapters 3 and 4, respectively). In particular, the ten Kansas City (Missouri, USA) isolates from 1999 clustered within the same IS*1002*-RFLP type, but were divided into two PFGE types (4.4.3, page 185). For epidemiological purposes, this result makes it hard to distinguish whether the isolates and the associated pertussis cases are closely related or not. This chapter, therefore, attempts to evaluate the various typing methods used for *B. pertussis*, and recommends a method, according to the criteria described below.

5.1 Strategy

At the time of writing, there has not been a published and comprehensive evaluation of typing methods on a well-defined panel of *B. pertussis* isolates. Therefore *B. pertussis* isolates were chosen to calculate the typability, reproducibility, index of discrimination and epidemiological concordance of the following typing methods; serotyping, IS*1002* restriction fragment length polymorphism (RFLP) analysis and pulsed-field gel electrophoresis (PFGE) using *Xba*I. Pertactin (*prnA*) and pertussis toxin S1 (*ptxA*) gene sequencing were not included for this evaluation as this information was available and used to select the isolates. Linkage between different types of these various typing methods were also investigated, although not for statistical significance, as the number of strains were too low. The availability of the unpublished annotated *B. pertussis* genome sequence allowed the comparison of fragments derived experimentally from the Tohama-I strain with those predicted by *in silico* analysis of its genome for both the IS*1002*-RFLP and the PFGE techniques.

5.1.1 Criteria used to select the representative *Bordetella pertussis* panel

The objective was to choose 40 to 50 *B. pertussis* isolates from the entire collection described in Appendix 3, such that the maximum discrimination of each method could be estimated. This was carried out after the majority of *prnA* and *ptxA* gene typing was completed and, hence, these characteristics were used to select the greatest variation for this sub-set; i.e., representatives of all three UK *prnA* types, *prnA*(1-3) and the two *ptxA* types, *ptxA*(1) and *ptxA*(2). It was decided to select clinical *B. pertussis* isolates of UK origin only, and to pick at least one isolate from every year available to maximise temporal spread. Therefore, there were 24 isolates selected between 1920 and 1985 for the panel (Table 5.1, page 190). These included two isolates from 1983, one previously investigated by

determining 95% of the pertactin gene, and the other was the first known *prnA*(3) isolate in the UK.

Isolates submitted to Manchester University in 1998 to 1999, also had epidemiological information available, such as name (not shown), date of birth, vaccination status and severity of disease (see Appendix 3, part B). Fourteen isolates from this data-set were selected for the panel, and the reasons for inclusion are listed in Table 5.1, page 190. This included two pairs of epidemiologically-linked isolates. A further two epidemiologically-linked pairs were added to the panel which were submitted to RSIL in 2002 (n=4) (Table 5.1, page 190). Two isolates from 1998 to 1999 were serotype 1,2 and *prnA*(2), which were exceptions from the two common relationships of serotype 1,2 and *prnA*(1), or serotype 1,3 and *prnA*(2) (3.3.6.5, page 139). An isolate from 1999 that was submitted for the Paediatric Intensive Care Unit (PICU) study to RSIL, PHLS CPHL, London (Crowcroft *et al.*, 2003), was also selected.

The *B. pertussis* type strain (18323^T=NCTC 10739^T), isolated originally in the USA, was also included in the panel as a well-characterised reference strain. Other reference strains included were all the available UK vaccine strains (n=4). In order to calculate the reproducibility of the typing methods, replicates of six isolates were assigned, making a total of 54 *B. pertussis* isolates. To avoid observer bias, all of the selected strains were randomly allocated a study control number, from UK 1 to UK 54[4].

5.2 Methods

The isolates were assessed using the following methods, as described in Chapter 2; serotyping, IS*1002*-RFLP analysis and PFGE with *Xba*I. Formulae for evaluating the above methods are expressed in section 2.8.3 (page 99).

5.2.1 The *in silico* digest of the Tohama-I genome

The *B. pertussis* genome was downloaded from the Sanger Institute web-site (http://www.sanger.ac.uk/Projects/B_pertussis/) after permission from Julian Parkhill, and then imported into Kodon software.

---

[4]After the allocation of study code numbers, and during the generation of laboratory data, it was discovered that one isolate (UK 38) was contaminated with *B. parapertussis*. This isolate was thus removed from the analysis, so that the panel contained 53 isolates.

Table 5.1. *Bordetella pertussis* panel for the evaluation of typing techniques

Original culture collection	Strain number	Coded number in panel	Year deposited/ isolated	Pertactin allele (*prnA*)	Pertussis toxin S1 allele (*ptxA*)	Serotype	Reason for inclusion in panel
NCTC	NCTC 10901	UK 5	1920	1	2	NK	date of deposition/isolation
Wellcome Collection	CN137	UK 2	1941	1	2	NK	date of deposition/isolation
Wellcome Collection	CN351	UK 3	1942	1	1	NK	date of deposition/isolation and 95% pertactin gene determined
Wellcome Collection	CN741	UK 4	1943	1	2	NK	date of deposition/isolation
Wellcome Collection	CN909	UK 1	1944	1	2	NK	date of deposition/isolation
Wellcome Collection	CN1407	UK 6	1946	1	2	NK	date of deposition/isolation
Wellcome Collection	CN2055	UK 7	1947	1	1	NK	date of deposition/isolation
Wellcome Collection	CN2420	UK 8	1948	1	1	NK	date of deposition/isolation
Wellcome Collection	CN2992B	UK 42	1949	1	2	1,2,3	in WCV
Wellcome Collection	CN2998	UK 9	1949	1	1	NK	date of deposition/isolation
Wellcome Collection	CN3099	UK 43 & 48	1950	1	1	1,2	in WCV
Wellcome Collection	CN3108	UK 10	1950	1	1	NK	date of deposition/isolation
Wellcome Collection	CN3701	UK 11	1954	1	1	NK	date of deposition/isolation
Wellcome Collection	CN4154	UK 12	1956	1	1	NK	date of deposition/isolation
Wellcome Collection	CN5447	UK 13	1963	1	2	NK	date of deposition/isolation
Wellcome Collection	CN5631	UK 14	1964	1	1	1,2,3	date of deposition/isolation
NCTC	NCTC 10909	UK 15	1966	1	1	1,3	date of deposition/isolation
NCTC	NCTC 10910	UK 16	1967	1	1	1,3	date of deposition/isolation
NCTC	NCTC10739[T]	UK 33 & 41	1970	6	4	NK	type strain
Glasgow University	77/19110	UK 17	1977	1	1	NK	date of deposition/isolation
Manchester University	PRL18335	UK 18	1978	1	1	1,2,3	date of deposition/isolation and serotype 1,2,3
Manchester University	PRL8002	UK 19	1979	1	1	1,2	date of deposition/isolation
CAMR	DCH53	UK 20	1982	1	1	NK	date of deposition/isolation and 95% pertactin gene determined
CAMR	DCH132	UK 21	1983	2	1	NK	date of deposition/isolation and 95% pertactin gene determined
CAMR	DCH164	UK 22	1983	3	1	NK	date of deposition/isolation and first known UK *prnA*(3)
Wellcome Collection	CN8410	UK 23	1984	1	2	NK	date of deposition/isolation

Table 5.1. *Bordetella pertussis* panel for the evaluation of typing techniques, cont.

Original culture collection	Strain number	Coded number in panel	Year deposited/ isolated	Pertactin allele (*prnA*)	Pertussis toxin S1 allele (*ptxA*)	Serotype	Reason for inclusion in panel
Wellcome Collection	CN8540	UK 24	1985	1	1	NK	date of deposition/isolation
Manchester University	98K300	UK 25	1998	1	1	1,2	date of deposition/isolation and 95% pertactin gene determined
Manchester University	98K316	UK 26	1998	2	1	1,2	rare serotype and *prnA* type combination
Manchester University	98K320	UK 27	1998	2	1	1,3	95% pertactin gene determined
Manchester University	98K322	UK 28	1998	1	1	1,2	} father/daughter related
Manchester University	98K323	UK 29	1998	[illegible]	1	1,2	
Manchester University	98K328	UK 30	1998	3	1	1,3	*prnA*(3) isolate
Manchester University	99K010	UK 31 & 46	1999	1	1	1,2	patient died
Manchester University	99K023	UK 32	1999	2	1	1,3	date of deposition/isolation
Manchester University	99K039	UK 34	1999	2	1	1,3	4 pertussis vaccine doses and 8 years old
Manchester University	99K045	UK 35	1999	3	1	1,3	95% pertactin gene determined
Manchester University	99K046	UK 36	1999	2	1	1,2	rare serotype and *prnA* type combination
Manchester University	99K062	UK 37	1999	2	1	1,3	} brother/sister related
Manchester University	99K063		1999	2	1	1,3	
Manchester University	99K079	UK 39	1999	2	1	1,3	Case was 45 years old
RSIL	PICU 475	UK 40	1999	2	1	1,3	from separate study
RSIL	Bp706	UK 51	2002	2	1	1,3	} brother/sister related
RSIL	Bp707	UK 53	2002	2	1	1,3	
RSIL	Bp710	UK 52	2002	2	1	1,3	} father/son related
RSIL	Bp711	UK 54	2002	2	1	1,3	
NIBSC	CN5476	UK 44 & 49		1	1	1,3	in WCV
NIBSC	Tohama-I	UK 45 & 50		1	2	NK	in ACV

Appendix 3 contain additional details of these strains (n=48). Strains in orange are related cases and are not included when calculating the discrimination indices. NK=not known; NCTC=National Collection of Type Cultures; CAMR=Centre for Applied Microbiological Research; RSIL=Respiratory and Systemic Infection Laboratory; NIBSC=National Institute for Biological Standards and Control; WCV=whole-cell vaccine and ; ACV=acellular vaccine.

To compare the observed and expected results of a IS*1002*-RFLP profile generated with the Tohama-I strain, a two-step procedure was required. An initial *in silico* digest, was performed using the recognition sequence of *Sma*I (CCC↓GGG) in Kodon. The fragments were then sized and listed according to their position in the genome. The positions of the IS*1002* element within the Tohama-I genome were then located and then matched with the appropriate fragments digested by *Sma*I as containing an IS*1002* element. These "virtual" fragment sizes were then compared with the calculated band sizes from the actual IS*1002*-RFLP pattern of the Tohama-I strain, following normalisation in BioNumerics.

To compare the observed PFGE profile of the Tohama-I genome with the expected one, a virtual digest using the recognition sequence of *Xba*I (T↓CTAGA) was performed. The Kodon software generated a virtual picture of the migration of bands produced after restriction. The band sizes were recorded and compared (for standard deviation values) with the band sizes calculated after normalisation (in BioNumerics) of a PFGE experiment.

5.3 Results

5.3.1 Typability of typing methods

Table 5.2 (page 193) shows the results of the three typing methods (n=53). All four serotypes were found using the panel (1 only; 1,2; 1,3; and 1,2,3). There were 19 different RFLP types determined using the panel and over a third of isolates (20/53) were RFLP-1. The PFGE data produced more types than the RFLP data (23 PFGE types), and also revealed a dominant profile, PFGE-1, found in 16/54 isolates.

To evaluate typability, one isolate from each replicate pair (used to calculate the reproducibility values), was omitted (n=47), and the typability value (T), was calculated using the formula in section 2.8.3 (page 99), (Table 5.2, page 193). One isolate underwent two separate attempts for serotyping, as it auto-agglutinated, but typing was successful on a further third attempt. The typability value was T=0.98 before re-testing the auto-agglutinating isolate and, when successful, the typability value became T=1.00. This however, lowers the confidence in the typability value when re-testing is performed, and should thus be expressed with caution. The typability values for both RFLP and PFGE were also T=1.00.

Table 5.2. Results of the typing methods derived from the *Bordetella pertussis* typing panel

Typing method	Types observed	Number of isolates	Typability value
Serotyping	**Serotype**		**1.00***
	1	6	
	1,2	20	
	1,3	19	
	1,2,3	9	
IS*1002*-RFLP analysis	**IS*1002*-RFLP type**		**1.00**
	RFLP-1	20	
	RFLP-2	6	
	RFLP-3	6	
	RFLP-6	1	
	RFLP-7	3	
	RFLP-11	1	
	RFLP-12	1	
	RFLP-13	1	
	RFLP-21	3	
	RFLP-22	1	
	RFLP-24	1	
	RFLP-25	2	
	RFLP-29	1	
	RFLP-35	1	
	RFLP-36	1	
	RFLP-38	1	
	RFLP-39	1	
	RFLP-40	1	
	RFLP-47	1	
PFGE analysis	**PFGE type (with *Xba*I)**		**1.00**
	PFGE-1	16	
	PFGE-3	3	
	PFGE-4	4	
	PFGE-5	4	
	PFGE-6	2	
	PFGE-7	3	
	PFGE-8	3	
	PFGE-9	1	
	PFGE-11	2	
	PFGE-12	2	
	PFGE-13	2	
	PFGE-16	1	
	PFGE-17	1	
	PFGE-20	1	
	PFGE-22	1	
	PFGE-23	1	
	PFGE-25	1	
	PFGE-26	1	
	PFGE-28	1	
	PFGE-29	1	
	PFGE-32	1	
	PFGE-33	1	
	PFGE-34	1	

*The typability value was previously 0.98 for serotyping, as one isolate was untypable. After three attempts, the isolate gave a serotype result, thus achieving the typability value of 1.00 (5.3.1, page 192).

5.3.2 Reproducibility of typing methods

In total there were six replicates in the panel, although a review of the results revealed an obvious discrepancy between the replicates of CN5476 (UK 44 and UK 49). These replicates differed from each other in all three typing methods evaluated here. Given the reported stability of serotyping and the marked differences in IS*1002*-RFLP and PFGE profiles, this is unlikely to have arisen from experimental variation, therefore this pair was excluded from the analyses[5]. Therefore, the typing results from the remaining five replicates were used to calculate the reproducibility value as described in section 2.8.3 (page 99).

Serotyping, PFGE and IS*1002*-RFLP analysis generated the same R value of 1.00 (Table 5.3, Figure 5.1 and Figure 5.2, pages 195 to 197).

5.3.3 Discrimination indices of typing methods

Indices of discrimination (*D*) were calculated as described in 2.8.3 (page 99) using the unrelated isolates (n=44), which are in black in Table 5.1, page 190.

The phenotypic method serotyping generated a value of *D*=0.70 and is low, as only four possible serotypes can be generated for *B. pertussis* (Preston, 1963). The DNA fingerprinting techniques IS*1002*-RFLP analysis and PFGE gave the highest indices of discrimination (*D*=0.85 and 0.89, respectively). However, these values still do not meet the requirements of an ideal typing system (*D*>0.95) (Struelens *et al.*, 1996), but they do approach the desired value described by Hunter and Gaston (*D*>0.90) (1988).

---

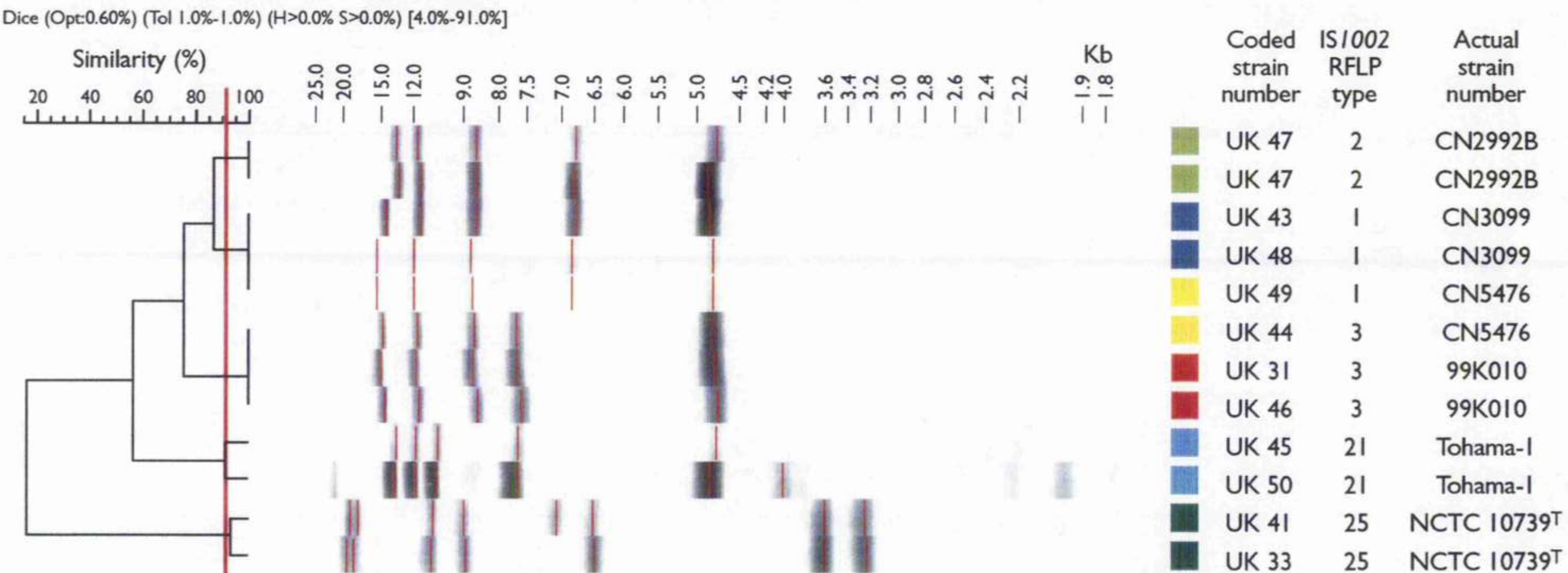
[5]Subsequent to the preparation of this thesis, the isolate CN5476, and its coded replicates UK 44 and UK 49, were re-examined by serotyping. Further isolates CN3099 and Tohama-I, and their respective coded replicates were also repeated as controls, and UK 10 and the original strain CN3108 were tested because they possessed the same genotype as that of the replicate UK 49. These serotyping results were consistent with those found in the coded trial; i.e., CN5476 and UK 44 were serotype 1,3; CN3099 and Tohama-I were serotype 1,2; and CN3108, UK 10 and UK 49 were serotype 1.

Table 5.3. Type designations of replicate isolates from the panel, with reproducibility values for each *Bordetella pertussis* typing method

Strain number	Coded replicates	Serotype	PFGE type (*Xba*I)	IS*1002*-RFLP type	Other details
CN3099	UK 43	1,2	1	1	UK WCV strain
	UK 48	1,2	1	1	*prnA*(1)/*ptxA*(1)
CN2992B	UK 42	1,2,3	5	2	UK WCV strain
	UK 47	1,2,3	5	2	*prnA*(1)/*ptxA*(2)
Tohama-I	UK 45	1,2	7	21	ACV strain
	UK 50	1,2	7	21	*prnA*(1)/*ptxA*(2)
NCTC10739[T]	UK 33	1	16	25	Type strain
	UK 41	1	16	25	*prnA*(6)/*ptxA*(4)
99K010	UK 31	1,2	8	3	Clinical isolate
	UK 46	1,2	8	3	*prnA*(1)/*ptxA*(1)
Reproducibility values:		**1.00**	**1.00**	**1.00**	

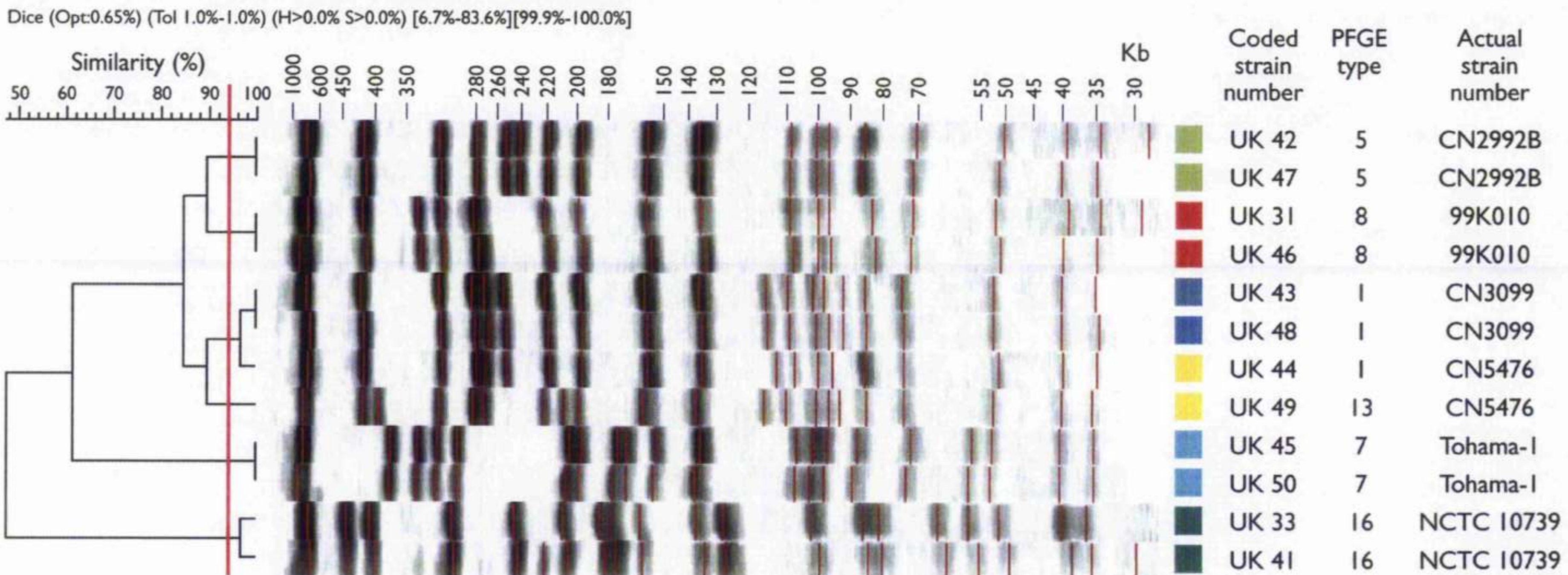
The pertactin and pertussis toxin S1 subunit gene types were not included in the evaluation and are shown for information only.

Figure 5.1. Reproducibility of *Bordetella pertussis* genotyping by IS*1002*-RFLP analysis



Cluster analysis was performed using the Dice coefficient of similarity and the dendrogram generated by the unweighted pair group method with arithmetic averages (UPGMA), where the thick red line shows the 91.8% similarity cut-off. The thin red lines indicate bands included in the analysis, and the coloured squares indicate the replicate strains. The replicates UK 44 and UK 49 differed in all three typing methods and were subsequently revealed to be from different strains. These replicates were therefore removed from the reproducibility calculations (see 5.3.2, page 194).

Figure 5.2. Reproducibility of *Bordetella pertussis* genotyping by PFGE using *Xba*I



Cluster analysis was performed using the Dice coefficient of similarity and the dendrogram generated by the unweighted pair group method with arithmetic averages (UPGMA), where the thick red line shows the 94% similarity cut-off. The thin red lines indicate bands included in the analysis, and the coloured squares indicate the replicate strains. The replicates UK 44 and UK 49 differed in all three typing methods and were subsequently revealed to be from different strains. These replicates were therefore removed from the reproducibility calculations (see 5.3.2, page 194).

5.3.4 Epidemiological concordance of typing methods

Originally, there were four pairs of isolates included in the panel that were epidemiologically-linked through family contact (Table 5.4 page 199). However, one pair of isolates from a brother and sister, 99K062 and 99K063, when tested after coding, were not identical by serotyping, PFGE typing or IS*1002*-RFLP analysis. The serotyping results for isolate 99K063 were also inconsistent. This strain was previously serotyped at Manchester University as serotype 1,3 but, when performed here as UK 38, it was serotype 1,2,3. Also, the profiles generated by PFGE with *Xba*I and IS*1002*-RFLP analysis for this isolate differed from the rest of the profiles from the epidemiological isolates, with 48.4% and 13.2% similarities respectively, when clustered by Dice and UPGMA (Figures 5.3 and 5.4, pages 200 and 201). The IS*1002*-RFLP profile of 99K063 closely resembled *B. parapertussis* profiles obtained by van der Zee *et al.* (1996a), and this isolate was thus revealed, by slide-agglutination with specific *B. parapertussis* antisera, to be mixed with *B. parapertussis.*

The only pair of isolates that gave the same types with all three typing methods was the father/daughter isolates 98K322 and 98K323, bearing serotype 1,2, PFGE-1 and RFLP-1. The second epidemiologically-related pair (brother and sister), BP706 and BP707, differed in only the RFLP type, generating RFLP-47 and RFLP-2, respectively (Figure 5.3, page 200). These two RFLP types differ only by a single band of approximately 7.09 Kb in RFLP-47 and are 90.9% similar when clustered. The other epidemiologically-related pair of isolates, BP710 and BP711, also differed by one typing method, PFGE, and produced types PFGE-4 and PFGE-3, respectively. These too are 90.9% similar, notably with the loss of an approximately 334 Kb band and gain of two lower bands in BP711 (Figure 5.4, page 201).

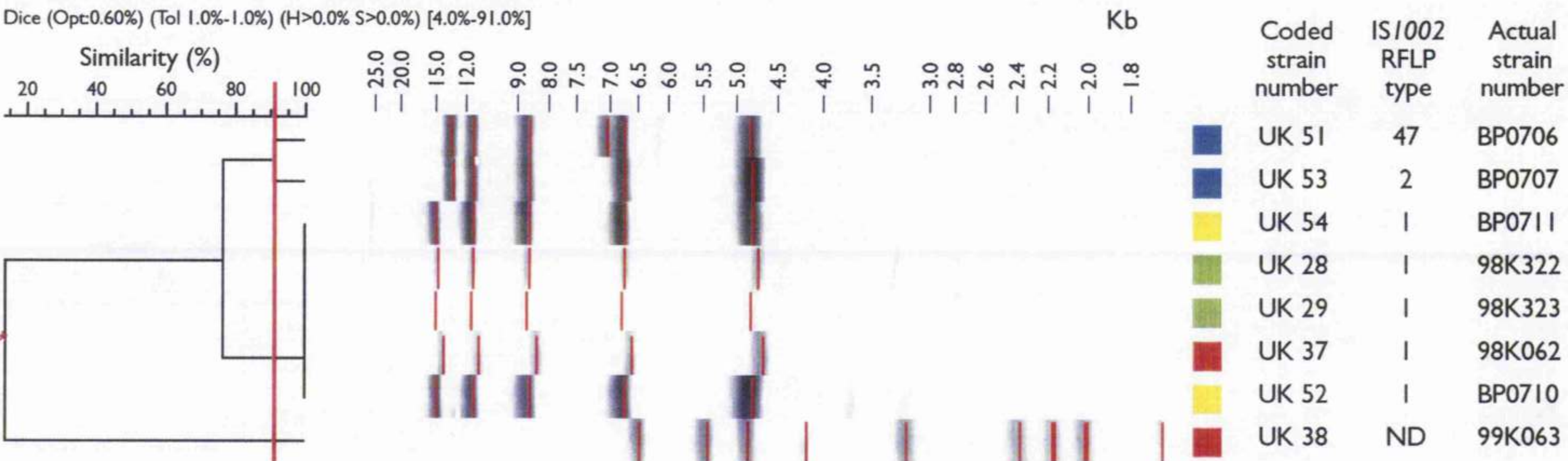
When calculating the epidemiological concordance (E), 99K062 and 99K063 were excluded, as 99K063 was contaminated and not recoverable. Serotyping showed the highest epidemiological concordance (E=1.00). Due to the mismatch of two of the three pairs of isolates by PFGE or IS*1002*-RFLP analysis, these methods gave low values, both at E=0.67.

Table 5.4. Type designations of 'epidemiologically-related' *Bordetella pertussis* isolates from the panel, with epidemiological concordance values

Strain number	Coded designation	Serotype	PFGE (*Xba*I) type	IS*1002*-RFLP type	Contact details	Pertactin and pertussis toxin S1 gene types
98K322	UK 28	1,2	1	1	Father	*prnA*(1)/*ptxA*(1)
98K323	UK 29	1,2	1	1	Daughter	*prnA*(1)/*ptxA*(1)
BP0706	UK 51	1,3	4	47	Brother	*prnA*(2)/*ptxA*(1)
BP0707	UK 53	1,3	4	2	Sister	*prnA*(2)/*ptxA*(1)
BP0710	UK 52	1,3	4	1	Father	*prnA*(2)/*ptxA*(1)
BP0711	UK 54	1,3	3	1	Son	*prnA*(2)/*ptxA*(1)
Epidemiological concordance:		**1.00**	**0.67**	**0.67**		**1.00 1.00**

The pertactin and pertussis toxin S1 subunit gene types were not included in the evaluation and are shown for information only.

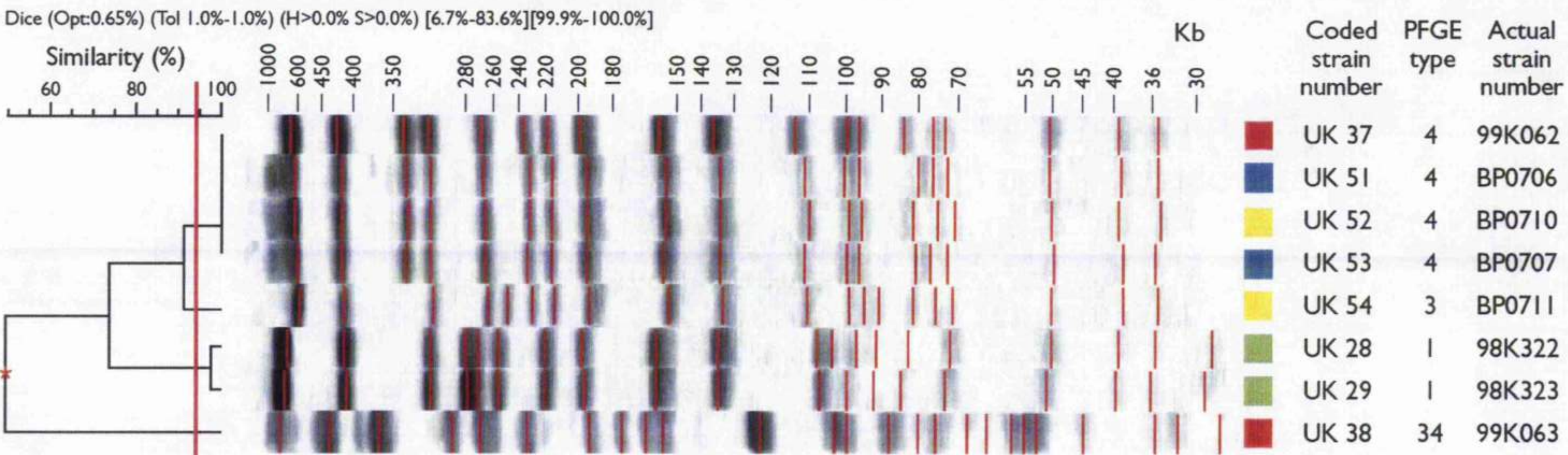
Figure 5.3. Cluster analysis of 'epidemiologically-related' *Bordetella pertussis* isolates by IS*1002*-RFLP analysis



Coded strain number	IS*1002* RFLP type	Actual strain number
UK 51	47	BP0706
UK 53	2	BP0707
UK 54	1	BP0711
UK 28	1	98K322
UK 29	1	98K323
UK 37	1	98K062
UK 52	1	BP0710
UK 38	ND	99K063

Cluster analysis was performed using the Dice coefficient of similarity and the dendrogram generated by the unweighted pair group method with arithmetic averages (UPGMA), where the thick red line shows the 91.8% similarity cut-off . The thin red lines indicate bands included in the analysis and the coloured squares indicate the 'epidemiologically-related' strains. The red asterix denotes the low similarity (13.2%) of UK 38 with the other seven isolates tested. This isolate was later found to be mixed with *B. parapertussis*. ND=not designated.

Figure 5.4. Cluster analysis of 'epidemiologically-related' *Bordetella pertussis* isolates by PFGE using *Xba*I



Cluster analysis was performed using the Dice coefficient of similarity and the dendrogram generated by the unweighted pair group method with arithmetic averages (UPGMA), where the thick red line shows the 94% similarity cut-off. The thin red lines indicate bands included in the analysis and the coloured squares indicate the 'epidemiologically-related' strains. The red asterix denotes the low similarity (48.4%) of UK 38 with the other seven isolates tested. This isolate was later found to be mixed with *B. parapertussis*.

5.3.5 Linkage between typing methods

Linkage between one particular type by one typing method with another type of a different typing method, was investigated within this panel of *B. pertussis* isolates. This was assessed after the code was broken, and one isolate from each known replicate or epidemiologically-linked pair was removed to avoid any false associations. Thus, nine related isolates were removed, leaving 45 isolates for analysis, and types with three or more isolates were investigated. The pertactin and pertussis toxin S1 subunit gene typing data were included in this analysis. All 12 *prnA*(2) isolates were *ptxA*(1), and the nine *ptxA*(2) isolates were all *prnA*(1). The predominant RFLP and PFGE types also had strong associations with pertussis toxin S1 gene types; RFLP-1 isolates were all *ptxA*(1) (n=16); and all but one PFGE-1 isolates were *ptxA*(1) (n=13/14). The same 13 PFGE-1 isolates were mostly *prnA*(1) (n=12/13). RFLP-3 isolates were *ptxA*(1) (n=5) and RFLP-7 isolates were *prnA*(1) (n=3). Other linkage was related to PFGE types, but was based on groups with only three isolates; three PFGE-4 isolates were serotype 1,3, *prnA*(2) and *ptxA*(1); and three PFGE-5 isolates were serotype 1,2,3, *prnA*(1) and *ptxA*(2).

There was no apparent continuity between the association of predominant types by one typing method and another type by another method, except for those stated above. For example, RFLP-1 isolates (n=20) were associated with eight different PFGE types and conversely PFGE-1 isolates (n=16) are associated with seven RFLP types.

5.3.6 Comparison of IS*1002*-RFLP analysis and PFGE profiles by *in silico* and observed analysis of the Tohama-I strain

The IS*1002*-RFLP band sizes estimated from the experimental analysis of the Tohama-I strain, after cleavage with *Sma*I enzyme and hybridisation of the IS*1002* element, were exported from BioNumerics. The sequenced Tohama-I genome was cleaved *in silico* using the *Sma*I recognition site in the software program Kodon, to locate the six predicted copies of the IS*1002* element. The restriction endonuclease *Sma*I is a frequent cutter in the GC-rich (68% G+C mol) *B. pertussis* genome, and the *in silico* digest generated 1564 fragments ranging from 6 to 27,563 bp from the 4 Mb genome. This does not however reveal where the six copies of the IS*1002* element are located. A second *in silico* step was thus required, searching the annotated genome for the IS*1002* elements, recording their positions and matching corresponding positions from the *Sma*I-digested fragments. Table 5.5 (page 204) shows the comparison of the actual fragment sizes with the *in silico* IS*1002*-containing *Sma*I fragments.

Only five major bands were produced experimentally by IS*1002*-RFLP analysis with the Tohama-I strain, but there was a lower, fainter band of 4070 bp (Figure 5.5, page 205). When this sixth, smaller fragment was included into the comparison analysis, the differences in fragment sizes between the observed and expected IS*1002*-RFLP profiles were small, ranging from 39 to 1968 bp (0.8-15.5% difference).

The *in silico* digest using the rare cutter recognition site of *Xba*I, to simulate the PFGE results with the Tohama-I strain, was more straightforward than the IS*1002*-RFLP analysis. The observed band sizes of the experimental PFGE analysis were exported from BioNumerics. Twenty-six fragments were generated ranging from 1236 to 789,751 bp. Kodon can also generate an image of the fragments as predicted if electrophoresed on an agarose gel. Figure 5.6 (page 207) shows the virtual gel analysis of the expected *Xba*I-digest and of the observed profile of the Tohama-I genome and strain, respectively. The fragment range was limited to show those between 48,500 and 1,018,500 bp, as this was the range of the fragments visualised from the lambda ladder used in the PFGE experiments. The two profiles were very similar and this was supported by the band size comparison (Table 5.6, page 206). Although there were only 19 fragments from the observed Tohama-I PFGE profile, the two extra fragments generated from the *in silico* digest (134,788 and 107,399 bp) were relatively close to two other *in silico* fragments (135,100 and 102,981 bp, respectively) and have probably co-migrated (Figure 5.6, page 207). The difference in band size ranges from 21 to 13,641 bp (0.1-14.1% difference).

5.4 Discussion

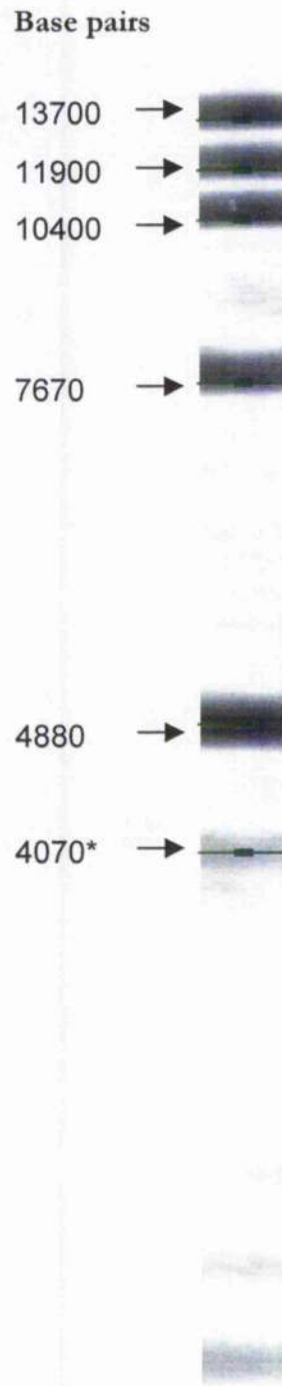
Various typing methods have been recommended for *B. pertussis* in either; comparing the three typing methods PFGE, RAPD and ERIC-PCR (Moissenet *et al.*, 1996); studying population genetics using IS*1002*-RFLP analysis (Musser *et al.*, 1986; van der Zee *et al.*, 1997); or investigating outbreak situations with PFGE (de Moissac *et al.*, 1994; Beall *et al.*, 1995; Mooi *et al.*, 2000). These recommendations are, however, subjective and the method is chosen usually to support the initial aims and objectives of the authors. Therefore, in this project, various typing methods were evaluated with a carefully selected panel of coded isolates, which included epidemiologically-linked and replicate isolates.

Table 5.5. A comparison of observed and predicted band sizes of fragments containing the IS*1002* elements of the *Bordetella pertussis* Tohama-I strain

Observed/estimated RFLP band sizes (bp)	Expected/*in silico* band sizes (bp)	Difference in bp between observed and expected (% difference)
13700	11732	1968 (15.48)
11900	10671	1229 (10.89)
10400	9867	533 (5.26)
7670	6819	851 (11.75)
4880	4919	39 (0.80)
4070*	4169	99 (2.40)

*This band was omitted for analysis of the coded typing panel. It is highlighted for the purposes of the comparison with the *in silico* results only.

Figure 5.5. The observed IS*1002*-RFLP profile of the *Bordetella pertussis* Tohama-I strain



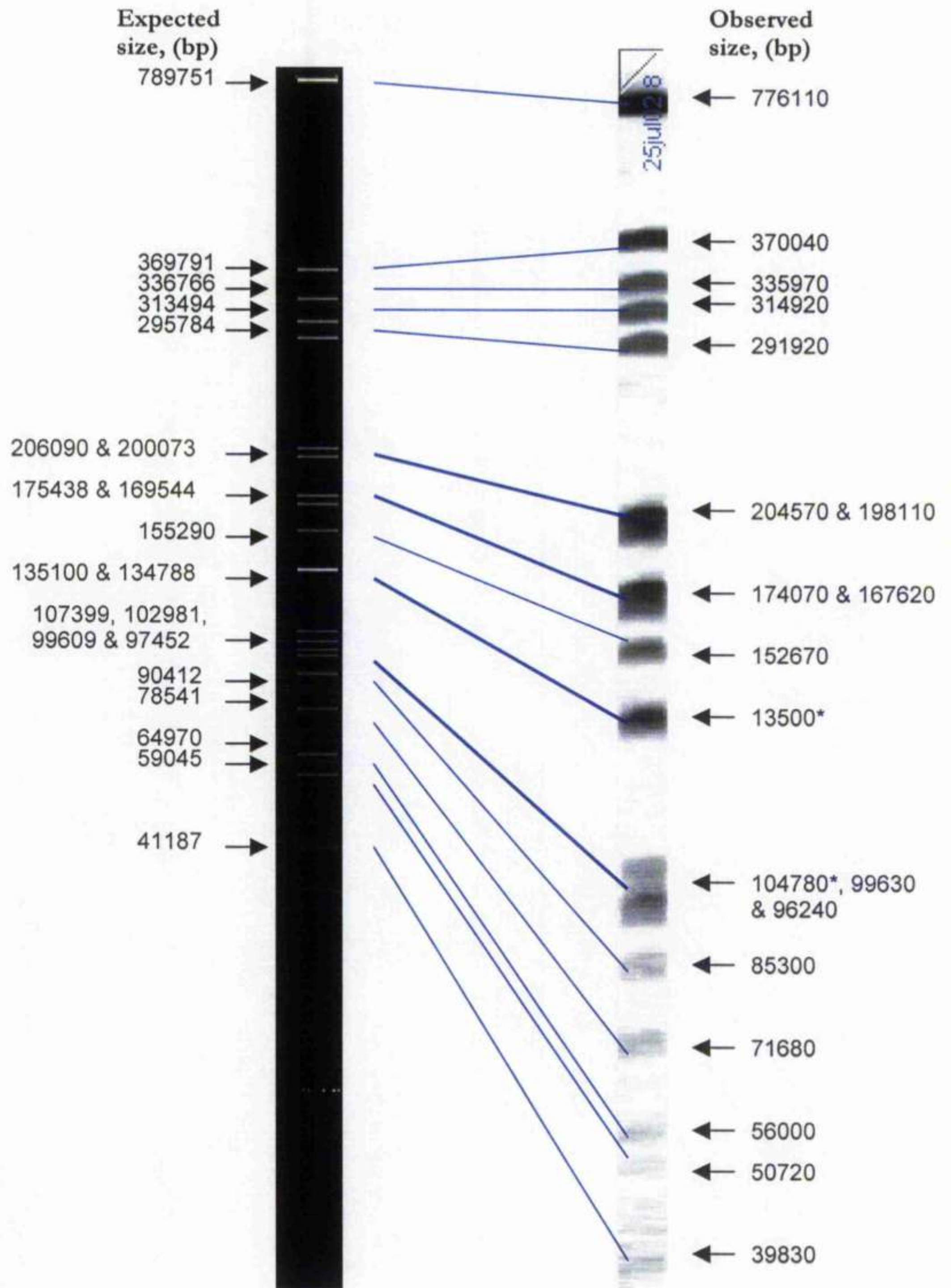
*This band was omitted for analysis of the coded typing panel. It is highlighted for the purposes of the comparison with the *in silico* results only.

Table 5.6. A comparison of observed and expected band sizes of the PFGE fragments of *Bordetella pertussis* Tohama-I with *Xba*I

Observed/estimated band sizes (bp)	Expected/*in silico* band sizes (bp)	Difference in bp between observed and expected (% difference)
776110	789751	13641 (1.73)
370040	369791	249 (0.07)
335970	336766	796 (0.24)
314920	313494	1426 (0.45)
291920	295784	3864 (1.31)
204570	206090	1520 (0.74)
198110	200073	1963 (0.98)
174070	175438	1368 (0.78)
167620	169544	1924 (1.13)
152670	155290	2620 (1.69)
135000*	135100	100 (0.07)
	134788	212 (0.16)
104780*	107399	2619 (2.44)
	102981	1799 (1.75)
99630	99609	21 (0.02)
96240	97452	1212 (1.24)
85300	90412	5112 (5.65)
71680	78541	6861 (8.74)
56000	64970	8970 (13.81)
50720	59045	8325 (14.10)
39830	41187	1357 (3.29)

*Fragments visualised as one band on the observed PFGE profile, probably due to co-migration of the corresponding expected bands.

Figure 5.6. A comparison of observed and expected band sizes of the PFGE fragments of *Bordetella pertussis* Tohama-I with *Xba*I



*The observed fragments correspond with two expected bands on the adjacent profile.

5.4.1 The suitability of the *Bordetella pertussis* typing panel

The panel of isolates comprised only 53 coded strains, and was used to calculate typability, reproducibility, index of discrimination and epidemiological concordance values to allow accurate comparisons of the different typing methods. However, there were a number of shortfalls of the panel. The availability of well-characterised, epidemiologically-linked isolates was limited, therefore only pairs of isolates originating from related family members were included in the panel. Unfortunately, one pair was eliminated from the analysis as one of the isolates was mixed with *B. parapertussis*. Hence, a future panel for further studies should contain more related isolates to calculate the epidemiological concordance value with more confidence, which would enable the panel to adhere to recommended guidelines to include five or more outbreak-related sets (Struelens *et al.*, 1996).

The panel only included six replicates, including four UK vaccine strains, the *B. pertussis* type strain (NCTC 10739^T) and one clinical isolate from 1999. Thus the reproducibility value may not be representative of the whole *B. pertussis* population.

One of the replicates, UK 49, did not appear to be the same as the original strain CN5476. This strain differed from its pair in serotype, IS*1002*-RFLP and PFGE type. Serotyping of UK 49 was repeated but the results remained the same (serotype 1). The only possible explanation for this discrepancy was incorrect labelling during the coding of the panel. It is noteworthy that the panel contained an isolate UK 10 (CN 3108), which had the same profile: serotype 1, PFGE-13 and RFLP-1, as that of UK 49. For these reasons, results from UK 49 and its replicate UK 44 were excluded from the analyses. As for the epidemiological concordance, again, the calculated reproducibility value would be more ideal if a larger number of replicates were used.

The pertactin and pertussis toxin S1 gene types were completed before the isolates were selected for the panel, and are presented in Chapter 3. This gene typing data were also used to select some of the strains and were therefore excluded from this evaluation. However, this information is used in the discussion so that comparisons with other typing methods could be made. These *B. pertussis* isolates have now become a well-characterised panel, used for investigation of sequence polymorphism in other gene markers (Packard *et al.*, 2004).

5.4.2 Evaluation of the typing methods

The typing methods serotyping, PFGE typing (using *Xba*I) and IS*1002*-RFLP analysis were assessed using the *B. pertussis* panel. Typability, reproducibility, index of discrimination and epidemiological concordance values were calculated according to the Consensus Guidelines for Microbial Epidemiologic Typing Systems (Struelens *et al.*, 1996). Convenience criteria such as ease of use and interpretation, analysis time and cost of methods were also evaluated. These are summarised in Table 5.7 (page 211), which includes the pertactin and pertussis toxin S1 subunit gene typing values not generated when the isolates were coded, but are nevertheless shown to compare the various methods. The observed and the expected results of the methods IS*1002*-RFLP analysis and PFGE typing using an *in silico* digest of the *B. pertussis* Tohama-I strain and genome is also compared and described. The stability of each typing method was not assessed using the panel of isolates. According to Struelens and co-authors (1996), at least 10 strains should be studied after every fifth passage in an experiment of 50 serial passages and thus yielding 100 tests to be performed for each typing method.

There is some linkage between types of one method and types of other methods. These include: 15/16 PFGE-1 isolates being *prnA*(1) and *ptxA*(1); all *prnA*(2) isolates are *ptxA*(1); all *ptxA*(2) isolates are *prnA*(1). These groups are not mutually exclusive with other method types, but can be described as clonal. Other, smaller clonal lineages are based on PFGE types, but these also show divergence in another typing method. For example, PFGE-3 isolates are serotype 1,3, *prnA*(2) and *ptxA*(1), but generate two different RFLP types. So, although these linkages are not statistically valid, they can be viewed as true relationships, as the *B. pertussis* panel was originally chosen to comprise a large diversity.

When evaluating the various typing methods, there are several factors to consider:

1) The typability of a method is dependent on the user's skill and whether the definition of a type has a distinct end-point. Repeating a test is not reflected in the typability value, *T*, and there are no guidelines in the literature as to how many attempts should be made on an untypeable isolate (Maslow *et al.*, 1993; Struelens *et al.*, 1996; van Belkum *et al.*, 2001). This is therefore at the user's discretion, and for data shown in this study, the author attempted serotyping on some isolates from the coded panel at three separate events (5.3.1, page 192). Also, it is rare for a DNA fingerprinting technique to fail, as bands can generally be visualised even, for example, if the restriction endonuclease step is not completed. An

isolate would therefore produce a fingerprint profile and, whether it is valid or not, it will be assigned a type, and thus the typability value of these methods would be T=1.00.

2) Struelens and colleagues (1996) advise the use of a typing scheme that gives an index discrimination of D>0.95, although Hunter and Gaston (1988) recommend a more realistic value of over 0.90 as desirable. If this cannot be achieved by one method, then it is possible to combine methods to reach a higher value. This may produce a system that has too much resolving power than is actually required, and the initial purpose of the system may be over-looked (i.e., to show epidemiologically-linked isolates related by a typing method, and not to separate them by using a more discriminatory one) (Maslow and Mulligan, 1996).

3) Different typing methods may monitor different genetic events that an organism can undergo. In the case of *B. pertussis*, serotyping relies on expression of the fimbrial genes, *fim2* and *fim3*. There can be many genetic events altering the expression of these genes, but in *B. pertussis*, fimbrial variation is thought to be due to the slipped-strand mis-pairing phenomenon (Willems *et al.*, 1990). DNA sequencing elucidates any variation in the genes encoding the virulence factors pertactin and pertussis toxin S1, which may be under environmental constraint due to the introduction of pertussis vaccines (Mooi *et al.*, 1998). Point mutations and recombination are seen in these two virulence genes, which may be due to this environmental constraint (van Belkum *et al.*, 2001). Sequencing and serotyping investigate specific genes, whereas PFGE and IS*1002*-RFLP analysis are dependent on restriction endonuclease sites and/or insertion elements situated throughout the entire genome. Both methods can monitor the same genetic events such as point mutations, recombination, inversions and the acquisition of transposons (van Belkum *et al.*, 2001), but IS*1002*-RFLP analysis limits this visualisation, as the transposable element IS*1002* is hybridised to only a few fragments. It is therefore not clear whether a mutation at the *Sma*I restriction site is influencing the differences between IS*1002*-RFLP profiles (observed as differences in position of labelled fragments), or if it is the acquisition of the insertion element into a different part of the *B. pertussis* genome. Hence, there are many genetic events that can be monitored by these various typing methods, and the choice of a particular method can greatly affect how the results are interpreted for epidemiological or genetic diversity studies.

Table 5.7. Summary of the evaluation of the typing methods used on *Bordetella pertussis*

Typing method	Number of types in panel	Calculated value				Convenience criteria[1]			
		Typability	Reproducibility	Index of discrimination	Epidemiological concordance	Ease of use	Ease of interpretation	Analysis time[2]	Cost
Serotyping	4	1.00	1.00	0.70	1.00	Easy	Excellent	30 minutes	Low
PFGE using *Xba*I	23	1.00	1.00	0.89	0.67	Difficult	Fair	5 days	Medium
IS*1002*-RFLP analysis	19	1.00	1.00	0.83	0.67	Difficult	Good	5 days	Medium
Pertactin gene typing[3]	4	1.00	1.00	0.51	1.00	Moderate	Excellent	3 days	High
Pertussis toxin S1 gene typing[3]	3	1.00	1.00	0.37	1.00	Moderate	Excellent	3 days	High

[1]These were scored at the author's discretion.

[2]This is the estimated time starting from *B. pertussis* culture plate to type designation.

[3]The pertactin and pertussis toxin S1 gene typing values, although not included in the evaluation, are shown for comparison.

Serotyping was the only phenotypic method assessed, and has been widely used to type *B. pertussis* isolates to investigate population trends in terms of vaccine efficacy and epidemic investigations (Preston and Carter, 1992; Khattak *et al.*, 1992). This method is easy to use, inexpensive and the end-point (by slide-agglutination) is easy to determine (Table 5.7, page 211). Seven isolates from the panel of 53 expressed only agglutinogen 1, which is reported to be rare in isolates from humans (Preston, 1988; Roberts and Parton, 2001). The index of discrimination was also perhaps over-estimated, as circulating *B. pertussis* strains express three of the four different serotypes. Thus the index of discrimination would be much lower if the serotype 1 isolates were to be removed from the calculation (D=0.65 cf. 0.71). Recent recommendations for serotyping *B. pertussis* suggest the use of monoclonal antibodies against Fim2 and Fim3, over the use of polyclonal antisera to agglutinogens 1, 2 and 3 (Mooi *et al.*, 2000). If these recommendations had been followed, serotype 1-only isolates would not have been typed, and both the index of discrimination and the typability value would have decreased.

The DNA fingerprinting methods, PFGE (using *Xba*I) and IS*1002*-RFLP analysis gave similar scores for both the calculated value and the convenience criteria. Pulsed-field gel electrophoresis and IS*1002*-RFLP analysis are both technically difficult, labour intensive, and require at least five days to generate a type. Interpretation can also be quite problematic; sophisticated software aids this process but it still relies to some extent on the user's definition of a type. The exclusion and inclusion of bands can greatly affect the profile and the subsequent index of discrimination of the method. Interestingly, PFGE profiles contained approximately 20 bands, and IS*1002*-RFLP generated between only three to nine bands, but they score similar indices of discrimination (D=0.89 cf. 0.83, respectively). This low number of bands does not fit the recommended criteria of Struelens *et al* (1996) of about 30, but the resulting index of discrimination is high for both PFGE and IS*1002*-RFLP analysis. The method of designating types also affects the discrimination; either by building an identification library to screen unknown types: or by clustering known types with unknown ones, and using a defined similarity cut-off to group and assign types. The latter method was used for both PFGE and IS*1002*-RFLP analysis, although this step also relies on user-defined subjectivity. If time had allowed, the author would have preferred to set up an identification library for both PFGE and IS*1002*-RFLP analysis.

The epidemiological concordance value was E=0.67 for both PFGE and IS*1002*-RFLP analysis. This was dependent on three pairs of epidemioloigically-related isolates, and both typing methods generated contradictory profiles with two different pairs. The

brother/sister pair, UK 51 and UK 53, produced RFLP types that differed by the gain/loss of one band, and the father/son pair, UK 52 and UK 54, differed in three bands of their corresponding PFGE fingerprints. According to Tenover *et al.* (1995), the pair that differed with PFGE are closely related as the three band difference corresponds to one genetic event, namely a gain of a restriction site in UK 54, or *vice versa* (Figure 5.4, page 201). It was unexpected and unusual that two epidemiologically-related pairs were different by two separate methods, and this questions either the relatedness of the isolate pairs or of the typing method used. However, PFGE has, on numerous occasions, been shown to cluster epidemiologically-related strains (de Moissac *et al.*, 1994; Beall *et al.*, 1995; Bisgard *et al.*, 2001; Brennan *et al.*, 2000), whereas IS*1002*-RFLP analysis has not been used for this purpose.

The two gene targets, pertactin and pertussis toxin S1 subunit, were not included for evaluation, but the values are shown in Table 5.7 (page 211) to discuss. The two methods generated very similar values and convenience criteria scorings, because the same technique was used; PCR amplification, purification and subsequent direct sequencing of the targets. Only the discriminatory power differed (D=0.51 cf. 0.37) with *prnA* typing scoring higher than *ptxA* typing. This difference is due to the number of types seen, and the isolates chosen for the panel contained four *prnA* types and three *ptxA* types. By combining the *prnA* and *ptxA* types, the index of discrimination increases to D=0.70, but this still does not comply with the recommendations of Struelens *et al* (1996) or Hunter and Gaston (1988) of an ideal value of D>0.95 or 0.90, respectively. The typability, reproducibility and epidemiological concordance values for *prnA* and *ptxA* are excellent at 1.00 (Table 5.7, page 211), but the types were determined before coding the panel. However, the author has confidence that these values reflect the methods with equal merit, as they have been undertaken extensively throughout the thesis. Furthermore, when the determination of over 95% of the pertactin gene was performed upon six UK clinical isolates, six USA clinical isolates and three vaccine strains (Chapters 3 and 4), all sequencing was double-stranded and there were numerous occasions where three or four strands would overlap, including the polymorphic region, region 1, and were completely concordant. However, disadvantages include the number of steps involved in the methods, the high cost of the sequencing equipment and reagents and, without automated systems, it is labour intensive and takes three days to achieve a type allocation.

The availability of the annotated *B. pertussis* Tohama-I genome enabled a direct comparison between the observed results, generated using IS*1002*-RFLP analysis and PFGE, and the

predicted results of *in silico* digests using *Sma*I and *Xba*I recognition sites, respectively. Observed sizes of the IS*1002*-containing *Sma*I-digested fragments appeared larger than the predicted sizes for most of the bands (Table 5.5, page 204). This is possibly due to the difference between the commmercially available DIG-labelled DNA molecular weight marker II (Roche), which has a photodigoxigenin introduced at approximately every 200th to 300th base in the DNA fragments, and the unlabelled DNA digest investigated, resulting in a potential electrophoretic difference in the migration of fragments. Furthermore, it had been reported previously that the Tohama-I genome only had five copies of the IS*1002* element (Figures 5.1 and 5.5, pages 196 and 205, respectively) (van der Zee *et al.*, 1997). However, a sixth smaller 4070 bp weakly-hybridised band is present and is also found in the annotated Tohama-I genome (Table 1.4, page 41) (Parkhill *et al.*, 2003).

The predicted and observed PFGE results were also similar when comparing the various fragment sizes. However, the *in silico* digest of the Tohama-I genome revealed two extra bands which, in the observed profile, could have been masked by co-migration events, and that there is, in general, a greater effect of error at the top of the gel where high mass bands migrate, compared to the bottom of the gel where low mass bands migrate. This difference between the observed and expected band sizes may also be due to the algorithm used for both the reference systems for IS*1002*-RFLP analysis and PFGE in BioNumerics, which is a logarithmic-dependent cubic spline fit. Other algorithms were tried, such as a first degree fit or a pole fit, but the chosen algorithm gave less variation and created a line of best fit through the band sizes of the marker used for the reference system. This genome analysis illustrates how an entire nucleotide sequence of an organism can aid the scientist in both experimental design and in optimisation of current methodologies.

5.4.3 Recommendation of a typing system for *Bordetella pertussis*

Pulsed-field gel electrophoresis has been reported as an effective tool for the epidemiological study of *B. pertussis* (Khattak *et al.*, 1992; de Moissac *et al.*, 1994; Beall *et al.*, 1995). This method was later compared with RAPD and ERIC-PCR, using 15 clinical isolates recovered within a two-year period and in the same geographical area (Moissenet *et al.*, 1996). The authors concluded that PFGE was indeed the better discriminatory tool out of the three methods, but no replicates were added to test for reproducibility, and no isolates from different geographical areas were included to verify whether the discrimination was at a practical level to apply to larger epidemiological investigations (Moissenet *et al.*, 1996). A few years later, in May 1999, a meeting to standardise methods

for the epidemiological typing of *B. pertussis* took place, where serotyping, PFGE and pertactin and pertussis toxin S1 gene typing were agreed as the recommended methods (Mooi *et al.*, 2000). However, this was largely based on what was currently employed, and no further standardisation has been implemented within this European group. There has been a recent study in the USA, investigating the reproducibility of PFGE in five laboratories, using three strains that had been tested weekly for three weeks (Hardwick *et al.*, 2002). They assessed the variability of the molecular sizes of each DNA fragment, and found the reproducibility within a gel was very high, there were some inter-laboratory variation, but the profiles generated were identified with the correct designation.

Recommendation of a typing method based on work in this thesis is difficult. There are advantages and disadvantages for each method. The DNA fingerprinting methods PFGE and IS*1002*-RFLP analysis, although obtaining high indices of discrimination (Table 5.7, page 211), have low epidemiological concordance values and are difficult and laborious to use. However, the other methods have higher epidemiological concordance values, but their *D* values are low. Serotyping is rapid and inexpensive, but *prnA* and *ptxA* gene typing are expensive and can take up to three days for a result. Struelens *et al.* (1996) advise that if the index of discrimination power is not high enough, it may be easier to combine methods than to refine a difficult one. Although the *prnA* and *ptxA* gene typing was not performed as part of the evaluation, I could have proposed the use of these methods in conjunction with serotyping, as the index of discrimination increases to 0.89 (according to the panel results), which equals that of PFGE. This *D* value is also very close to 0.90, considered as desirable by Hunter and Gaston (1988). The option of semi-automated sequencing systems increases the ease of use of this system. The disadvantage of using these three methods together is that only defined genes are tested for variability or expression, as opposed to PFGE, which analyses the whole genome. Furthermore, there is a strong association of serotype and pertactin gene type, where recent UK islolates have been predominantly serotype 1,3 and *prnA*(2), and all recent isolates are *ptxA*(1), thus restricting the discrimination and epidemiological concordance value. Therefore, if more time was allocated to optimise and evaluate PFGE, this would have been the recommended method. Regrettably, at the time of writing, the author cannot recommend a method to epidemiologically type *B. pertussis*.

A multi-locus sequence typing (MLST) scheme has been devised recently by van Loo *et al.* (2002). This targets three surface protein genes, pertussis toxin S1, pertussis toxin S3 (*ptxC*) and tracheal colonisation factor (*tcf*) but, surprisingly, not pertactin. These additional

targets (pertussis toxin S3 and tracheal colonisation factor) have also been shown to have polymorphism in UK isolates, where two *ptxC* and five *tcf* alleles have been found between 1920 and 2002 (Packard *et al.*, 2004). These additional targets could be used in conjunction with serotyping and *prnA* and *ptxA* gene typing, although this scheme may not be adequate to discriminate for outbreak settings.

6 Novel *Bordetella pertussis* gene targets for epidemiological typing†

† **Note:** *During the course of this thesis,* Van Loo *et al.* (2002) *reported an MLST scheme based on surface protein genes, which may have affected the approach of this work.*

6.1 Strategy

The use of pertactin and pertussis toxin S1 gene sequencing for the epidemiological typing for *B. pertussis* have been used in this thesis (Chapters 3 and 4) and various countries (Mastrantonio *et al.*, 1999; Mooi *et al.*, 1998 and 1999). However, the discrimination of this typing scheme could be improved by adding other variable gene targets to the sequencing scheme. This multi-locus sequence approach has been successful for many bacteria and, although originally based on house-keeping genes (Enright and Spratt, 1998; Maiden *et al.*, 1998), other genes such as virulence factors can be targeted for inclusion of such a typing scheme (Gaia *et al.*, 2003). Two strategies were used to identify suitable polymorphic genes of *B. pertussis*.

The first strategy examined the MLEE technique used previously for *Bordetella* species, to identify any differences in electrophoretic mobilities in the enzymes tested (Musser *et al.*, 1986; van der Zee *et al.*, 1997). Only four of the 15 enzymes used in the *Bordetella* scheme showed such variation for *B. pertussis*: phosphoglucomutase, adenylate kinase, leucine aminopeptidase and glutamic oxaloacetic transaminase. The genes encoding these enzymes should therefore contain sequence polymorphisms that could be useful for typing *B. pertussis*. However, at the time of this project, sequence data were available only for adenylate kinase (*adk*), a tryptophan-containing enzyme that phosphorylates adenosine monophosphate (AMP) (GenBank accession number Z29715).

The second strategy focused on the study of virulence-associated genes for any possible polymorphisms. One virulence gene, that encoding filamentous haemagglutinin (*fhaB*), was screened for such variation. Filamentous haemagglutinin was chosen for its similarity to pertactin, as they are both adhesins, possess repeat domains and are included as important immunogens in WCVs and ACVs (1.2.3.1, page 12). Filamentous haemagglutinin could therefore contain polymorphisms similar to pertactin in either the repeat unit number or the composition of the repeats. The *fhaB* sequence, X53405, deposited in the GenBank database by Delisse-Gathoye and colleagues (1990), is derived from the Tohama-I strain.

6.2 Methods

The descriptions of PCR preparation, purification of PCR products and the sequencing method are detailed in sections 2.5.2, 2.5.5 and 2.5.6 (pages 63, 76 and 77), respectively. The primers designed for both targets are listed in Table 2.3 (page 66). Amplification and optimisation parameters specific to the *adk* and *fhaB* gene targets are described below. For

all targets identified and assessed, two temporally-different *B. pertussis* strains were used to test for polymorphism: the well-characterised type strain (NCTC 10739^T, deposited before 1947) and a clinical isolate (PICU 475) from 1999.

6.2.1 Adenylate kinase gene

The open reading frame of this gene (GenBank accession number Z29715) is 657 nucleotides and oligonucleotides (primers) were designed to amplify a larger region encompassing this (702 bp).

Amplification of *adk* was tested using the following reaction mixtures with or without 10% v/v DMSO: 2.5 mM $MgCl_2$, 10 mM Tris-HCl (pH 8.3), 2.5 mM KCl, 200 μM dNTP, 10 pmol each of adkF and adkR primers, 2.5 U *Taq* DNA polymerase, and 10 or 100 ng template DNA. The cycling conditions involved an initial denaturation at 95°C for 5 minutes, then 30 cycles of 95°C for 30 or 45 seconds, annealing at 40-70°C for 30 or 45 seconds, and extension of 72°C for 1 minute. This was completed with a final extension of 7 minutes at 72°C, and a 4°C hold. An amplicon size of 702 bp was predicted.

6.2.2 Filamentous haemagglutinin gene

The *fhaB* sequence, X53405, was used to design the fourteen primers (Table 2.3, page 66) to amplify and sequence the first 3.5 Kb of the *fhaB* gene in various overlapping amplicons.

To successfully amplify the 5'-end of the *fhaB* gene, two or three separate reactions targeting overlapping regions were attempted. The following five primer pairs were evaluated; fha-224F and fha-1889R, fha-1705F and fha-3482R, fha-224F and fha-1366R, fha-1219F and fha-2361R and fha-2223F and fha-3482R, producing amplimers of 1682, 1796, 1159, 1159, and 1277 bp, respectively (Figure 6.1, page 221). All primer pairs were optimised in the following reaction mixtures with or without 10% v/v DMSO: 1.5-5.5 mM $MgCl_2$, 10 mM Tris-HCl (pH 8.3), 2.5 mM KCl, 200 μM dNTP, 10 pmol each of forward and reverse primers, 2.5 U *Taq* DNA polymerase, and 10 or 100 ng template DNA.

All reactions were run with optimisation parameters of initial denaturation at 95°C for 5 minutes, followed by 30 cycles of 95°C for 30 or 45 seconds, annealing between 40-70°C for 30 or 45 seconds, and extension of 72°C for 1 minute, with a final extension at 72°C for 7 minutes, and a 4°C hold.

The amplification of the *fhaB* gene was not completely optimised, and details are given in section 6.3.2 (page 226).

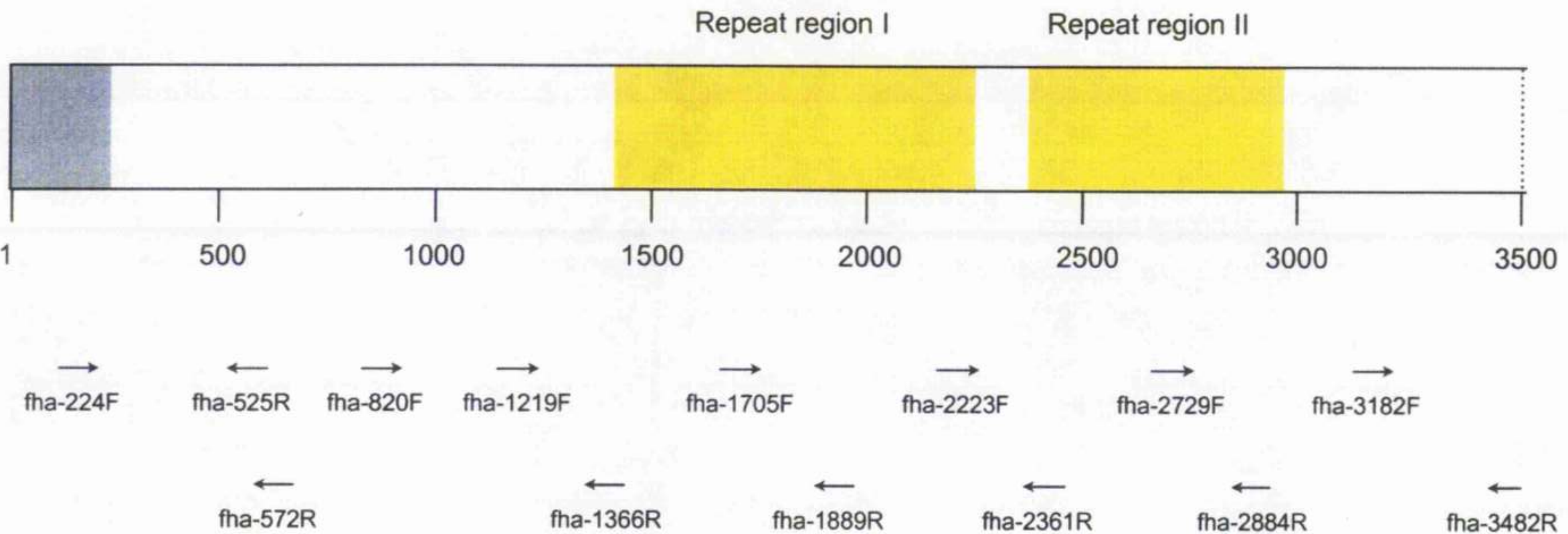
6.3 Results

6.3.1 Adenylate kinase gene

Figure 6.2 (page 222) illustrates the optimised amplification using 10 ng of DNA template, 10% DMSO, and varying the annealing temperature from 40°C to 64°C. No products were visualised when no DMSO was added and 100 ng of DNA was used. Hence, amplification of the adenylate kinase gene was successfully optimised when the reaction mixtures contained 10% DMSO, 2.5 mM $MgCl_2$, 10 mM Tris-HCl (pH 8.3), 2.5 mM KCl, 200 µM dNTP, 10 pmol of each primer, 2.5 U *Taq* DNA polymerase and 10 ng of template DNA. The optimum cycling conditions had an initial denaturation at 95°C for 5 minutes, then 30 cycles consisting of denaturation at 95°C for 45 seconds, annealing at 50°C for 45 seconds, and extension at 72°C for 1 minute, followed by a final extension of 7 minutes at 72°C, with a final 4°C hold.

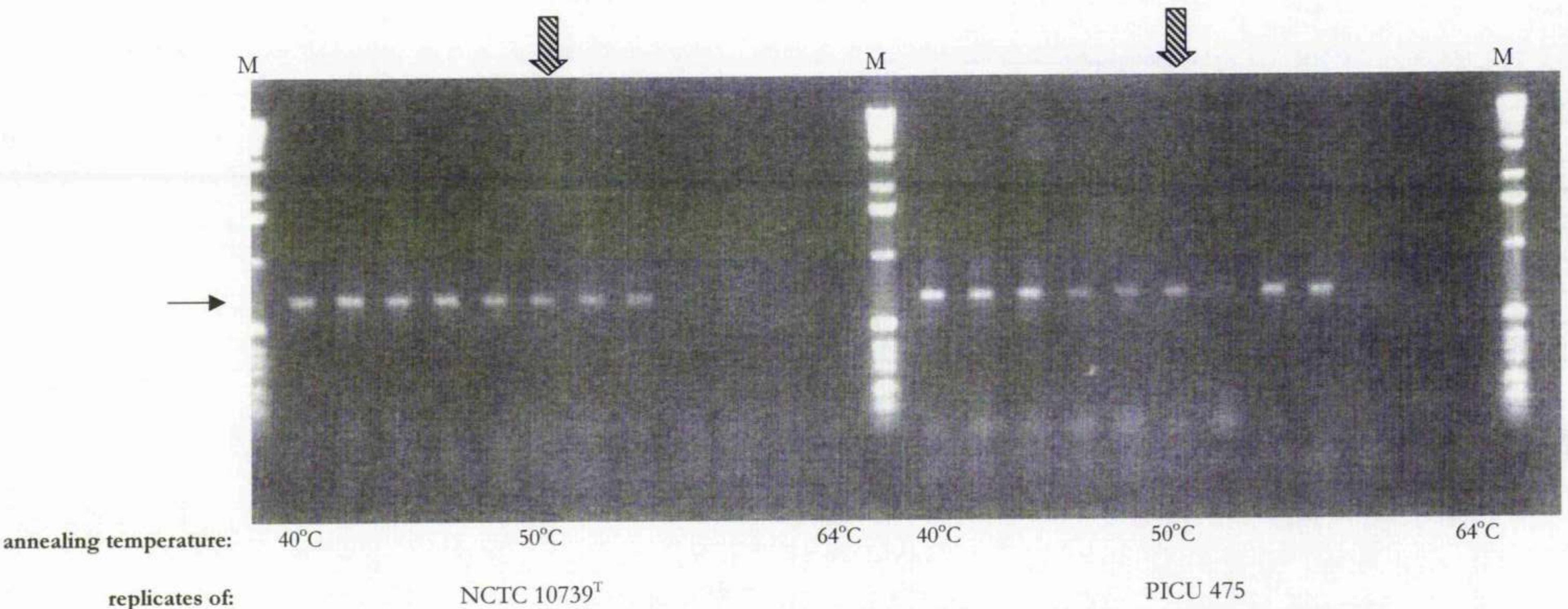
Both NCTC 10739^T and PICU 475 generated discrete bands of approximately 700 bp (Figure 6.2, page 222), and the two *adk* amplicons were purified and sequenced. The subsequent sequence data were of sufficient quality to investigate for polymorphisms. An alignment of these two *adk* sequences with that deposited in GenBank (accession number Z29715) (derived from the type strain of *B. pertussis* 18323^T), revealed a single nucleotide difference (A→G) in the generated sequence from strain NCTC 10739^T (Figure 6.3, page 223). This confers as a predicted amino acid difference: a glutamic acid (E) in the deposited sequence and a glycine (G) in the NCTC 10739^T strain sequenced. The NCTC 10739^T sequence started at position 12 and finished at position 634, according to the start codon position ATG of the deposited sequence, and was 622 nucleotides long. The deposited sequence Z29715 showed no differences with the 645 nucleotides generated from PICU 475 (start at –8, stop at 637). Therefore, there was one nucleotide difference between the *adk* sequences generated from NCTC 10739^T and PICU 475 at position 364 (Figure 6.3, page 223).

Figure 6.1. Schematic of the 5'-end of the *Bordetella pertussis* filamentous haemagglutinin gene (*fhaB*)



This schematic shows the 5'-end of the *fhaB* gene. GenBank accession sequence X53405 comprises 3514 bp and the coding region is from base 253 to >3514. The entire gene is 10.1 Kb. Repeat regions I and II, which contain 19-residue pseudo-repeats, are shown in yellow. The region encoding the mature protein is shown in white, and the grey region indicates the bases before the start codon. The arrows show the approximate position and direction of the primers used for PCR and sequencing. Details of primers can be found in Table 2.3, page 66.

Figure 6.2. PCR amplification of the *Bordetella pertussis* adenylate kinase gene with primers adkF and adkR, at various annealing temperatures



M=1 Kb ladder (Invitrogen). Reactions were performed in the presence of 10% DMSO, and 10 ng template DNA was used from either the type strain (NCTC 10739^T) or a clinical isolate (PICU 475) of *Bordetella pertussis*. The annealing temperature was varied from 40 to 64°C, from left to right on the gel. The black arrow highlights the predicted amplicon size of 702 bp and the hatched arrow indicates the optimal reaction used for subsequent experiments.

Figure 6.3. Nucleic acid alignment of the *Bordetella pertussis* adenylate kinase gene

```
    Consensus: GCCGCCGGCGTGGACACGCCAGCCGACCTGGAACGTGCCCGGGCCGCATA
       Z29715: GCCGCCGGCGTGGACACGCCAGCCGACCTGGAACGTGCCCGGGCCGCATA
NCTC 10739^T : --------------------------------------------------
     PICU 475: --------------------------------------------------

    Consensus: CACGAATCGGTTATAAGGGTTATTTCCCATGGCGCGGCCCCACATCCGGT
       Z29715: CACGAATCGGTTATAAGGGTTATTTCCCATGGCGCGGCCCCACATCCGGT
NCTC 10739^T : --------------------------------------------------
     PICU 475: --------------------------------------------------
                                      Bp-adkF
                                  ------------->
    Consensus: GCTGCATTGCGGCATAATCGCCCCGATCACAAAAAACAAACCCATCAGGA
       Z29715: GCTGCATTGCGGCATAATCGCCCCGATCACAAAAAACAAACCCATCAGGA
NCTC 10739^T : --------------------------------------------------
     PICU 475: ------------------------------------------------GA

    Consensus: GCCCCCATGCGTCTCATTCTGCTCGGACCGCCCGGAGCCGGCAAAGGCAC    44
                     M  R  L  I  L  L  G  P  P  G  A  G  K  G  T
       Z29715: GCCCCCATGCGTCTCATTCTGCTCGGACCGCCCGGAGCCGGCAAAGGCAC
NCTC 10739^T : ------------------CTGCTCGGACCGCCCGGAGCCGGCAAAGGCAC
     PICU 475: GCCCCCATGCGTCTCATTCTGCTCGGACCGCCCGGAGCCGGCAAAGGCAC

    Consensus: CCAAGCCGCCTTTCTCACCCAACACTACGGCATCCCGCAGATATCCACCG    94
                Q  A  A  F  L  T  Q  H  Y  G  I  P  Q  I  S  T
       Z29715: CCAAGCCGCCTTTCTCACCCAACACTACGGCATCCCGCAGATATCCACCG
NCTC 10739^T : CCAAGCCGCCTTTCTCACCCAACACTACGGCATCCCGCAGATATCCACCG
     PICU 475: CCAAGCCGCCTTTCTCACCCAACACTACGGCATCCCGCAGATATCCACCG

    Consensus: GTGACATGCTGCGCGCCGCCGTCAAGGCCGGCACGCCGCTGGGCCTGGAA   144
               G  D  M  L  R  A  A  V  K  A  G  T  P  L  G  L  E
       Z29715: GTGACATGCTGCGCGCCGCCGTCAAGGCCGGCACGCCGCTGGGCCTGGAA
NCTC 10739^T : GTGACATGCTGCGCGCCGCCGTCAAGGCCGGCACGCCGCTGGGCCTGGAA
     PICU 475: GTGACATGCTGCGCGCCGCCGTCAAGGCCGGCACGCCGCTGGGCCTGGAA

    Consensus: GCCAAGAAGGTCATGGACGCGGGCGGCCTGGTCTCGGACGACCTGATCAT   194
                A  K  K  V  M  D  A  G  G  L  V  S  D  D  L  I  I
       Z29715: GCCAAGAAGGTCATGGACGCGGGCGGCCTGGTCTCGGACGACCTGATCAT
NCTC 10739^T : GCCAAGAAGGTCATGGACGCGGGCGGCCTGGTCTCGGACGACCTGATCAT
     PICU 475: GCCAAGAAGGTCATGGACGCGGGCGGCCTGGTCTCGGACGACCTGATCAT

    Consensus: CGGCCTGGTGCGCGATCGCCTGACCCAGCCCGATTGCGCCAACGGCTACC   244
                 G  L  V  R  D  R  L  T  Q  P  D  C  A  N  G  Y
       Z29715: CGGCCTGGTGCGCGATCGCCTGACCCAGCCCGATTGCGCCAACGGCTACC
NCTC 10739^T : CGGCCTGGTGCGCGATCGCCTGACCCAGCCCGATTGCGCCAACGGCTACC
     PICU 475: CGGCCTGGTGCGCGATCGCCTGACCCAGCCCGATTGCGCCAACGGCTACC

    Consensus: TGTTCGACGGTTTCCCGCGCACCATCCCGCAGGCCGACGCGCTCAAGAGC   294
               L  F  D  G  F  P  R  T  I  P  Q  A  D  A  L  K  S
       Z29715: TGTTCGACGGTTTCCCGCGCACCATCCCGCAGGCCGACGCGCTCAAGAGC
NCTC 10739^T : TGTTCGACGGTTTCCCGCGCACCATCCCGCAGGCCGACGCGCTCAAGAGC
     PICU 475: TGTTCGACGGTTTCCCGCGCACCATCCCGCAGGCCGACGCGCTCAAGAGC
```

Figure 6.3. Nucleic acid alignment of the *Bordetella pertussis* adenylate kinase gene, continued

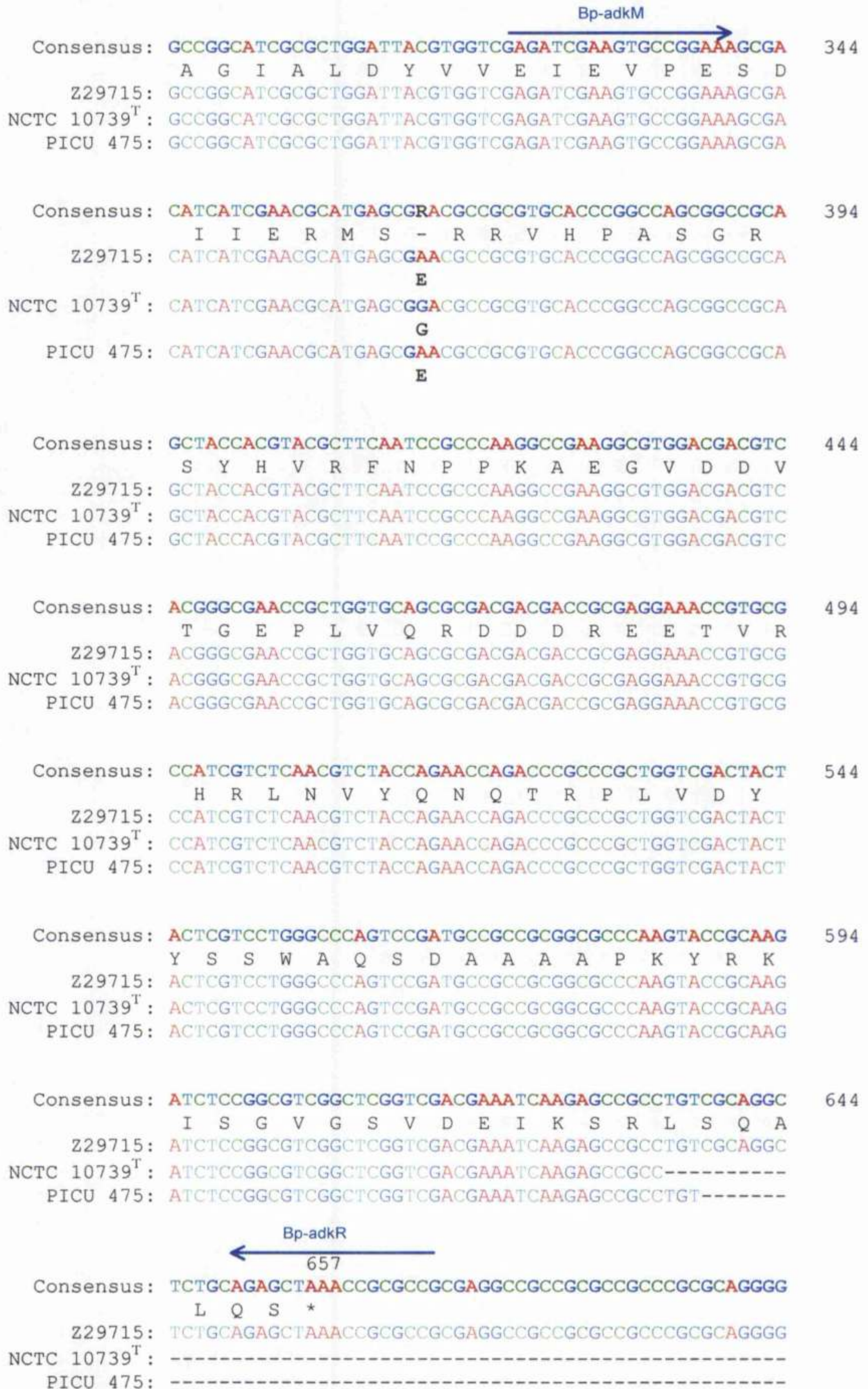


Figure 6.3. Nucleic acid alignment of the *Bordetella pertussis* adenylate kinase gene, continued

```
  Consensus: TGCGCGGCCGCGCGCGCCCTGCCCGCCTTCAT
     Z29715: TGCGCGGCCGCGCGCGCCCTGCCCGCCTTCAT
NCTC 10739T: --------------------------------
   PICU 475: --------------------------------
```

The GenBank sequence, Z29715, was used to align the sequences generated from NCTC 10739^{T} and PICU 475. The amino acids are coded as the universal single letter designations (Appendix 1), and are shown in black underneath the nucleotide coding sequence; the black horizontal line and black asterix (*) indicate the start and stop codons of adenylate kinase, respectively. The oligonucleotide positions are illustrated as blue arrows, with the direction and name of primer shown.

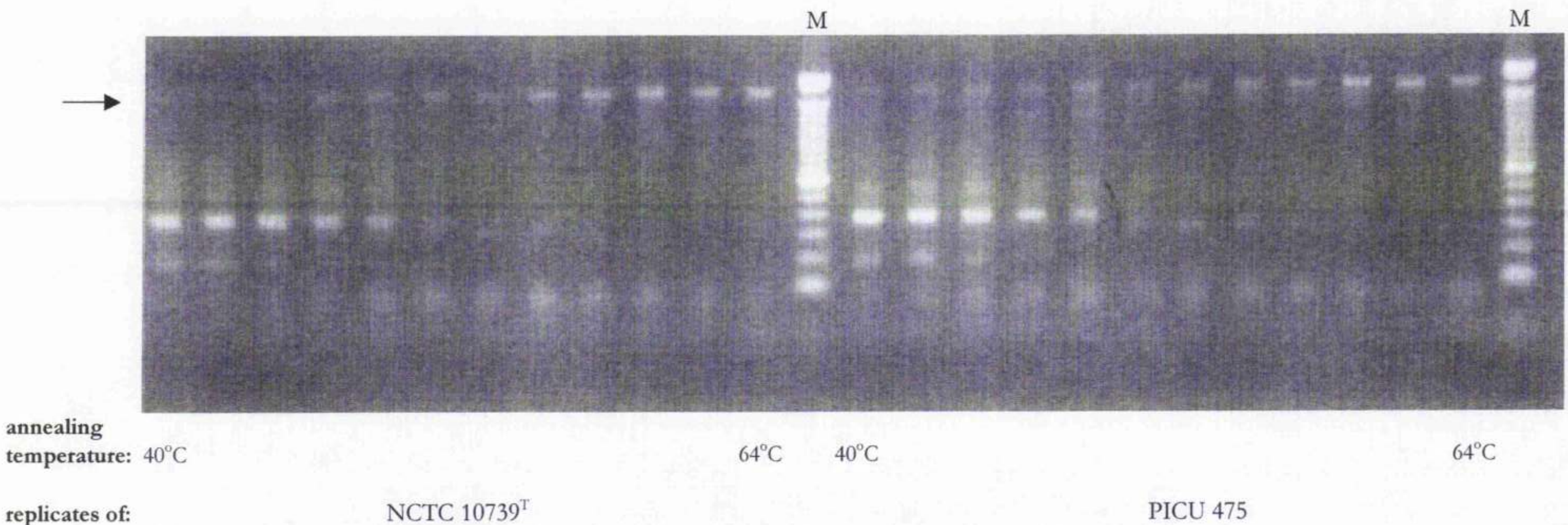
6.3.2 Filamentous haemagglutinin gene

Various optimisation experiments were performed. An initial attempt to amplify the 3514 base pair region using two overlapping fragments was largely unsuccessful. The 5'-fragment could not be amplified with primers fha-224F and fha-1889R. Primers fha-1705F and fha-3482R, although yielding a product of the predicted size of 1796 bp (Figure 6.4, page 227), other lower-sized fragments were also produced, which impeded the subsequent purification and sequencing of the targeted amplicon. Amplification of three smaller overlapping fragments was thus attempted and, during optimisation experiments, single amplicons were produced of the predicted sizes. Figures 6.5 and 6.6 (pages 228 and 229) show the optimisation with primers fha-224F and fha-1366R, and primers fha-1219F and fha-2361R respectively, by adjusting the $MgCl_2$ concentration and the annealing temperature. The optimisation experiment using primers fha-2223F and fha-3482R is shown in Figure 6.7 (page 230).

Appropriate single amplicons were purified, sequenced and assembled into larger contigs using the GenBank sequence X53405 as a reference sequence for alignment.

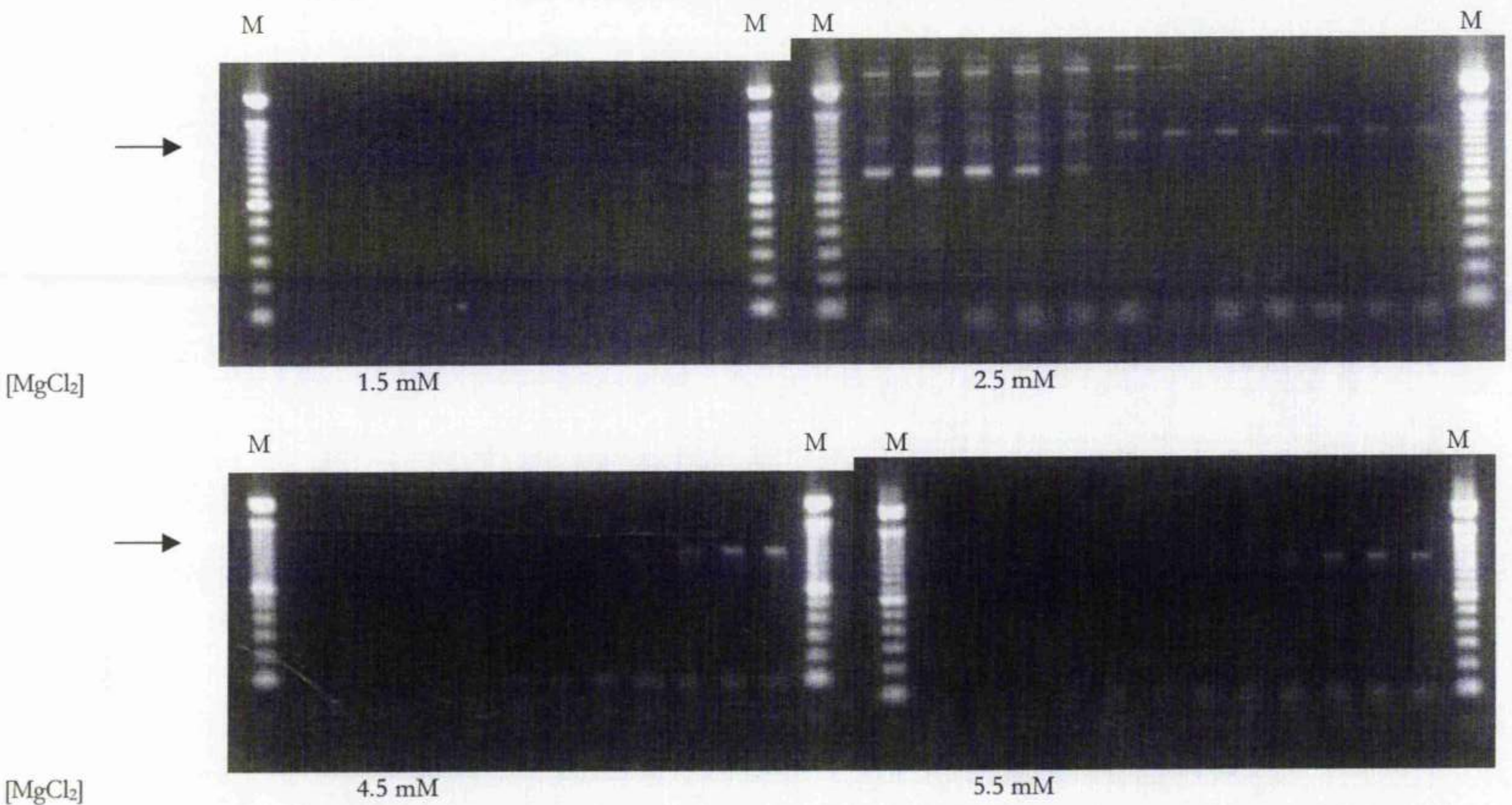
Poor quality sequence data arose from both the two *B. pertussis* isolates, NCTC 10739^T and PICU 475, but three and five contigs were produced, respectively, by aligning the data. No sequence data were available before position 1706 for NCTC 10739^T and position 1753 for PICU 475 (position according to GenBank sequence X53405). Figure 6.8 (page 231) shows the pictorial representation of the arrangement and position of the various contigs for both of the *B. pertussis* strains. Although the sequences generated were not of good quality, there were four and three definite single nucleotide polymorphisms (SNPs) in NCTC 10739^T and PICU 475, respectively, when compared with the deposited sequence X53405 from the Tohama-I strain. The other differences shown in Figure 6.8 (page 231) were not judged to be true polymorphisms. The true SNPs were at the following codons for NCTC 10739^T, with their respective amino acid changes: position 1799, GTG→GAG = valine→glutamic acid; 2745, CAA→CAC = glutamine→histidine; 3363, GTG→GTA = valine; and 3445, GCC→ACC = alanine→threonine. Although PICU 475 generated single-stranded data only, SNPs were present at: position 1799, GTG→GAG = valine→glutamic acid; 2345, TTC→TCC = phenylalanine→serine; and 3500, GGC→GAC = glycine→aspartic acid. According to findings of Delisse-Gathoye *et al.* (1990), the polymorphisms found in both strains at position 1799 correspond to the first region encoding a repetitive amino acid sequence. Another mutation at position 2745 lies within the second repetitive region.

Figure 6.4. PCR amplification of the *Bordetella pertussis* filamentous haemagglutinin gene with primers fha-1705F and fha-3482R, at various annealing temperatures



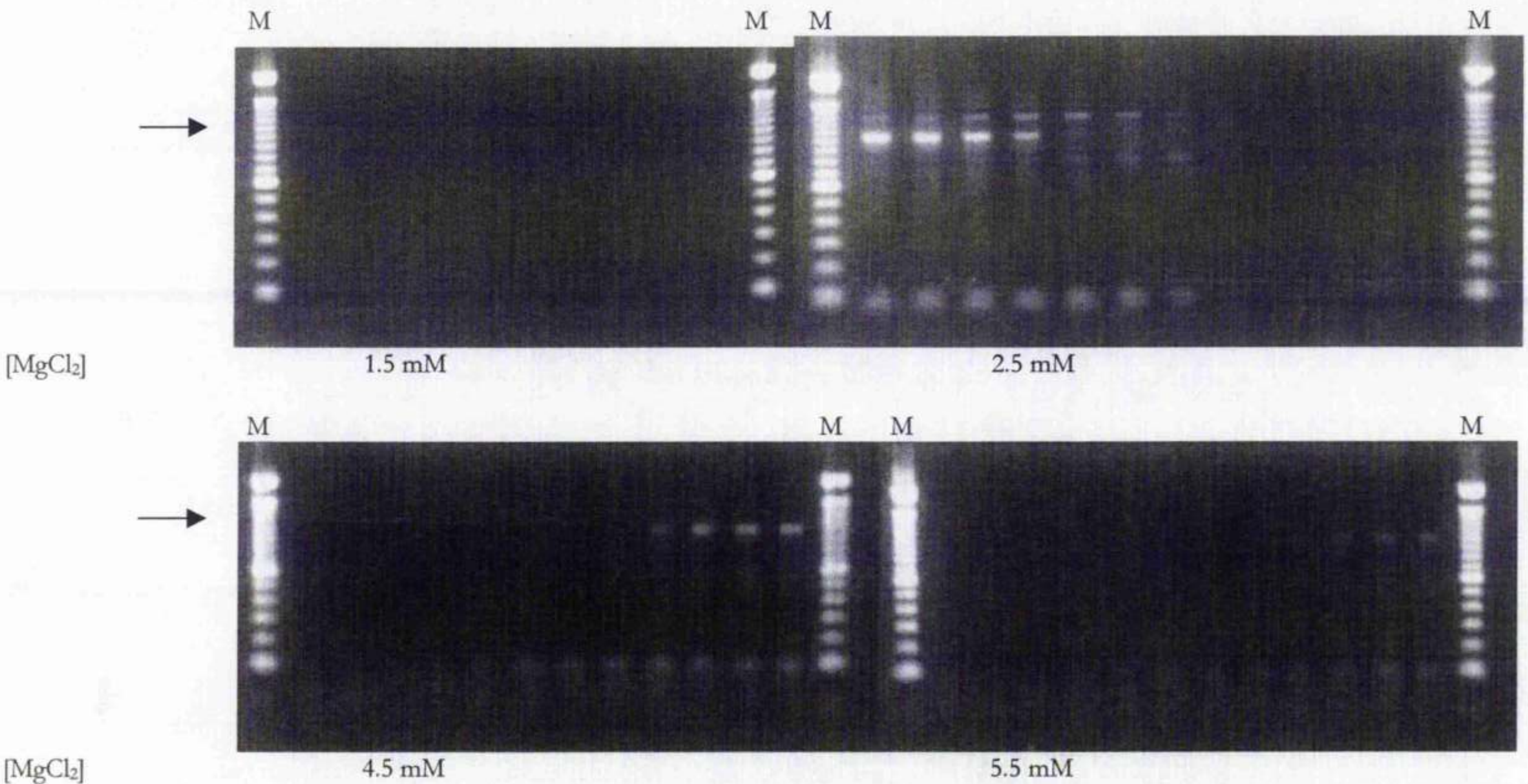
M=100 bp ladder (Invitrogen). Reactions contained 10 ng template DNA template from either the type strain (NCTC 10739^T) or a clinical isolate (PICU 475) of *Bordetella pertussis*. The annealing temperature was varied from 40 to 64°C, from left to right on the gel. The arrow highlights the predicted amplicon size of 1796 bp.

Figure 6.5. PCR amplification of the *Bordetella pertussis* filamentous haemagglutinin gene with primers fha-224F and fha-1366R, at various magnesium chloride concentrations and annealing temperatures



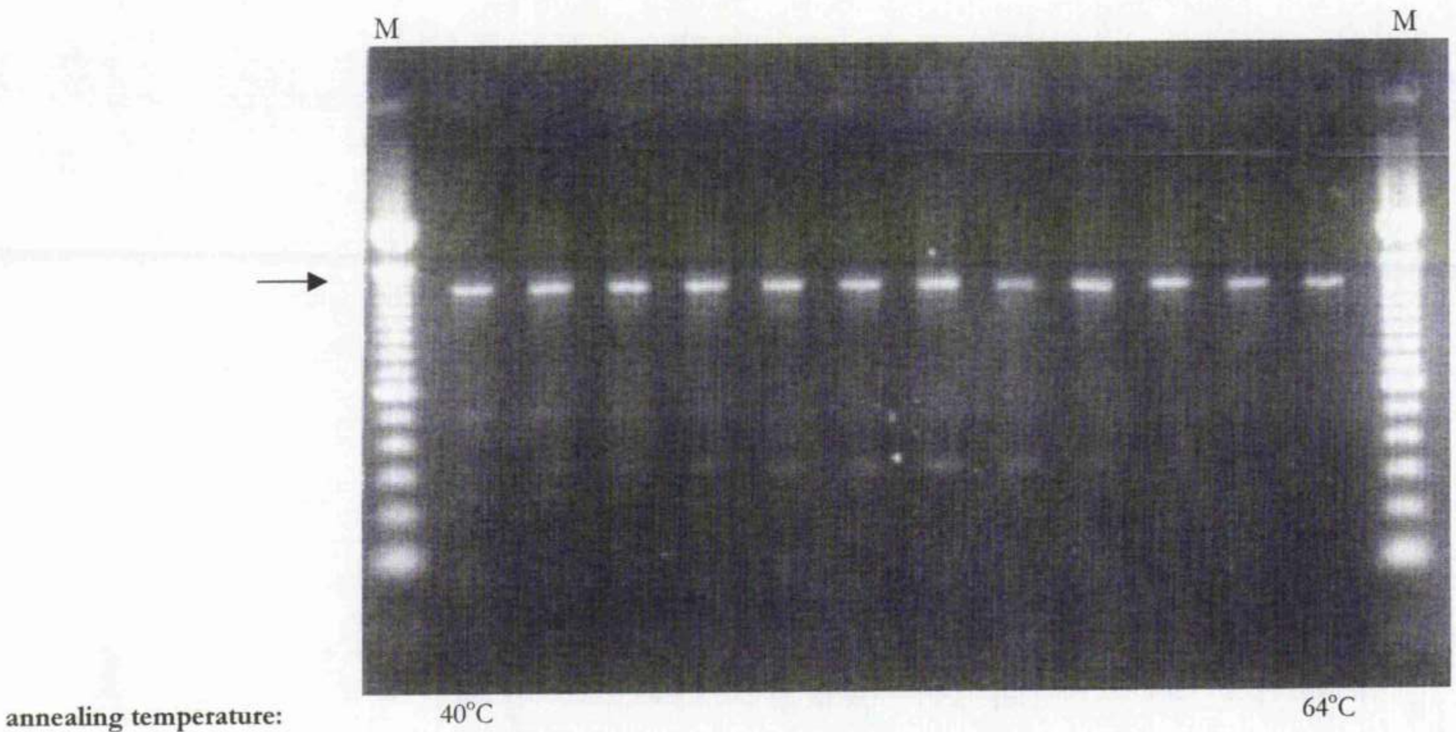
M=100 bp ladder (Invitrogen). Reactions contained 10 ng template DNA from the type strain of *Bordetella pertussis*, NCTC 10739^{T}. The annealing temperature was varied from 40 to 64°C, from left to right on the gel. The arrows highlight the predicted amplicon size of 1159 bp. The experiment using 3.5 mM $MgCl_2$ was not shown as amplicons were not visualised.

Figure 6.6. PCR amplification of the *Bordetella pertussis* filamentous haemagglutinin gene with primers fha-1219F and fha-2361R, at various magnesium chloride concentrations and annealing temperatures



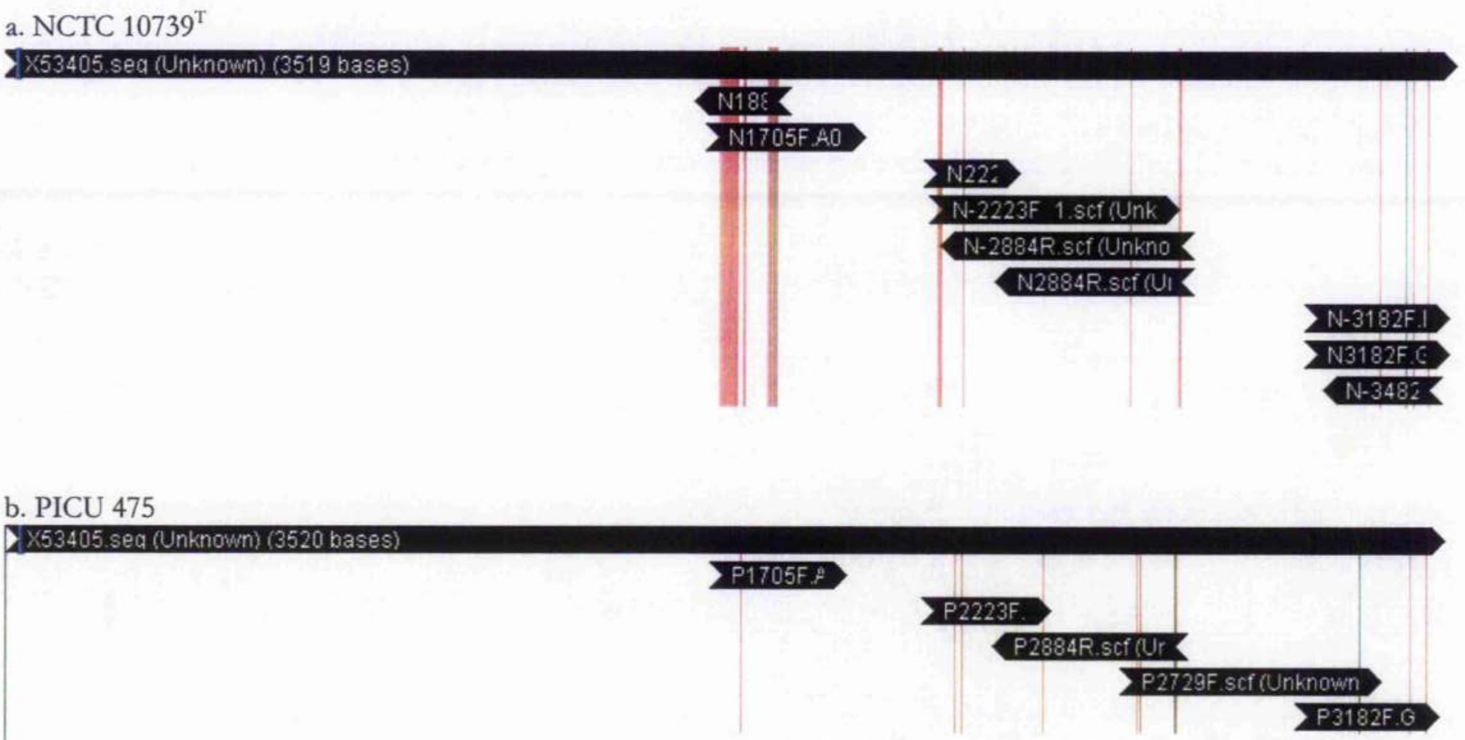
M=100 bp ladder (Invitrogen). Reactions contained 10 ng template DNA from the type strain of *Bordetella pertussis*, NCTC 10739[T]. The annealing temperature was varied from 40 to 64°C, from left to right on the gel. The arrows highlight the predicted amplicon size of 1159 bp. The experiment using 3.5 mM $MgCl_2$ was not shown as amplicons were not visualised.

Figure 6.7. PCR amplification of the *Bordetella pertussis* filamentous haemagglutinin gene with primers fha-2223F and fha-3482R, at various annealing temperatures



M=100 bp ladder (Invitrogen). Reactions contained 10 ng template DNA from the type strain of *Bordetella pertussis*, NCTC 10739^T. The annealing temperature was varied from 40 to 64°C, from left to right on the gel. The arrow highlights the predicted amplicon size of 1277 bp.

Figure 6.8. Contiguous sequences of the *Bordetella pertussis* filamentous haemagglutinin gene. Alignments are taken from the software program Gene Builder. The positions of the single-stranded raw data sequences from strains a) NCTC 10739^T and b) PICU 475 are aligned with the 5' end of sequence X53405 (GenBank), shown as the longest black arrow above the shorter forward and reverse arrows. Consequently, there are some single-stranded results which have been used for analysis.



Pink and green vertical lines correspond to nucleotide differences or an ambiguous nucleotide, respectively, between the different single-stranded sequence data.

6.4 Discussion

When investigating variability in genes for epidemiological typing, the clonality of the organism, the function of the gene and the motive of the researcher have to be considered. *Bordetella pertussis* is seen to be clonal when MLEE has been applied. Thus, variation in house-keeping genes is rare, but has been found in this study in the adenylate kinase gene. As the open reading frame of this gene is 657 bp, double-stranded consensus sequences can be produced for most of this gene using only two sequence reactions. However, variation was seen at only one nucleotide position and, although this translated to an amino acid difference, this was between the GenBank sequence Z29715 of 18323^T, and the sequence generated from this study of NCTC 10739^T. Theoretically, there should not be any differences as they are from the same strain of *B. pertussis* and, interestingly, the sequence from PICU 475 did match the deposited sequence in GenBank. The data held at GenBank may not necessarily be correct, as there are no independent checks when sequence data is submitted. Furthermore the *B. pertussis adk* deposited sequence could be wrong due to the relatively early production of the sequence (circa 1993) and the method used to generate the data (plasmid cloning and possibly a manual sequencing technique) (Gilles *et al.*, 1993). To check this discrepancy, further sequencing would be required, although the sequences generated for this study were from double-standed contigs. Sequencing of other isolates of *B. pertussis*, ideally using the panel described in Chapter 5, would help determine whether adenylate kinase could be used as an epidemiological marker.

As most of the house-keeping genes do not show variation (as seen in MLEE), and *B. pertussis* is shown to be mostly clonal, virulence genes could potentially be more polymorphic and thus would be good epidemiological markers. Indeed, pertactin and the S1 subunit of pertussis toxin exhibit polymorphisms (Mooi *et al.*, 1998), but two recent papers from van Loo *et al.* (2002) and van Loo and Mooi (2002), report variation in other virulence genes, including the S3 subunit of pertussis toxin, fimbriae for serotype 2, a porin protein, tracheal colonisation factor, virulence-activated gene-8, Bvg-intermediated phase protein and filamentous haemagglutinin. Six polymorphisms from the two strains were identified from a region in the *fhaB* gene between positions 1799 and 3500. The two strains, the type strain NCTC 10739^T, and the clinical strain PICU 475, were isolated over twenty years apart, where the type strain has been shown to be genetically distinct from other *B. pertussis* strains, and it was therefore expected to see differences between these two strains (Musser *et al.*, 1986; Gerlach *et al.*, 2001). The two SNPs that lie within regions encoding

amino acid repeat sequences translate to different amino acids. This is similar to the findings with pertactin, although the PRN protein possesses direct repeat regions (Mooi *et al.*, 1998).

The one difference found in *fhaB* by van Loo *et al.* (2002) is at the same location (position 2488) as one of the differences found in this study (position 2745, numbering notation of this study). Van Loo *et al.* (2002) initially based their data on 13 strains (origin not stated) and then confined the region investigated between positions 2250 and 2750 (their numbering commences on the first nucleotide of the coding region, which corresponds to 2507-3007 for this study) in 22 strains. The additional polymorphisms identified in this study, which code for differences in the amino acid sequence of filamentous haemagglutinin, may also have virulence and vaccine implications, and further study may be useful. Indeed, the MLST scheme of van Loo *et al.* (2002) could also be enhanced with the extended *fhaB* sequence and further optimisation may increase the quality and reliability of the data. The extent of polymorphism in *fhaB* could also be determined, as suggested for adenylate kinase, with the panel of *B. pertussis* isolates described in Chapter 5.

7 Assessment of single-enzyme Amplified Fragment Length Polymorphism analysis of *Bordetella pertussis*

7.1 Strategy

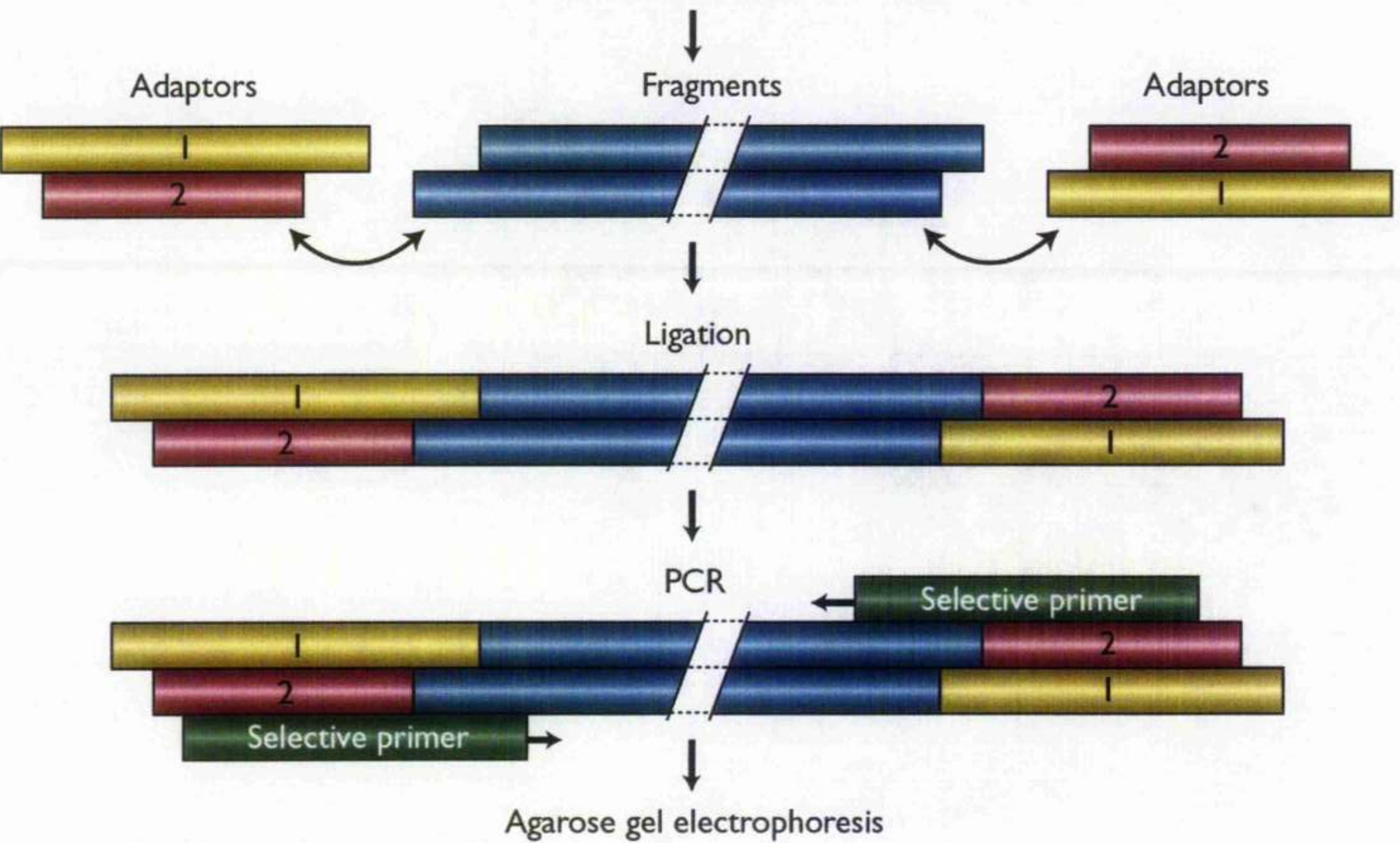
Single-enzyme AFLP analysis has been applied to several bacteria, including *Legionella pneumophila*, and this method has been chosen by the European Working Group on Legionella Infections (EWGLI) to standardise for use as an epidemiological typing scheme across 11 European countries (Fry *et al.*, 2000). As this method has been demonstrated previously to be flexible, simple, quick and relatively inexpensive, AFLP was attempted for this project on a selection of *B. pertussis* isolates from Missouri, USA. The methodology and assessment of this technique is described in this chapter only (and not in Chapter 2), as it has not been validated thoroughly to justify its use for epidemiological or population studies of *B. pertussis*.

7.2 Method

The *B. pertussis* isolates from Missouri, USA, which were previously investigated in Chapter 4, were selected to attempt AFLP because they had been sorted into three sub-sets. Nine were chosen, where three were "Early" isolates, three were "Kansas City" isolates and three were "Recent" isolates. Theoretically, the "Early" isolates would show different profiles to the "Recent" isolates, and the "Kansas City" isolates, being geographically and temporally clustered within seven weeks between June and July 1999, could be of the same AFLP profile. In addition, typing data from Chapter 4 (serotype, *prnA*, *ptxA*, IS*1002*-RFLP and PFGE) were compared with the AFLP data for any associations between types.

The basic method was adapted from Valsangiacomo *et al.* (1995) and the standard EWGLI AFLP protocol as detailed on the EWGLI *Legionella pneumophila* Typing Database, http://www.hpa.org.uk/srmd/bioinformatics/ewgli/ewglimeth.htm (Fry *et al.*, 1999). Briefly, genomic DNA is digested with a restriction enzyme and concurrently incubated with T4-DNA ligase and short adapters designed to link to the newly restricted DNA fragments (Figure 7.1, page 236). These restricted fragments are then amplified with primers that match both the adapter sequence and part of the recognition sequence of the restriction enzyme. The numbers of amplified fragments produced can be reduced by adding one or two extra bases on the primer. The amplified fragments are finally separated using conventional gel electrophoresis (Figure 7.1, page 236).

Figure 7.1. Pictorial representation of the single-enzyme AFLP method



Please refer to 7.2 (page 235) for a general description of the AFLP method.

Six enzymes and adapters were evaluated for band patterns produced with the nine isolates. Two enzymes were frequent cutters containing a four base recognition sequence (*Mse*I and *Taq*I) and the other four were rare cutters, with a six base recognition sequence (*Apa*I, *Eco*RI, *Hind*III and *Pst*I) (Table 7.1, page 238). The adapters and non-selective primers for the above enzymes were taken from previous studies (Valsangiacomo *et al.*, 1995; Janssen *et al.*, 1996). Primers were used that select one in four or one in eight of the digested and tagged fragments, by adding one or two nucleotides to the non-selective core primer sequence, respectively (Table 7.1, page 238).

7.2.1 Single step restriction-ligation of genomic DNA

Reactions were prepared in round-bottomed 1.5 ml Sarstedt tubes and consisted of 1.5 µg of genomic DNA (extracted by the method described in 2.4.1, page 60), 0.2 µg of each adapter-oligonucleotide (Table 7.1, page 238), 20 U of restriction endonuclease, 1 U of T4-DNA ligase (Roche) and 1 x ligation buffer (66 mM Tris-HCl, 5 mM $MgCl_2$, 1 mM dithioerythritol, 1 mM ATP, pH 7.5) (provided with T4-DNA ligase). Nuclease-free water was added to a final volume of 20 µl and reactions were incubated at 37°C for 3 hours.

The tagged DNA fragments were then precipitated by the addition of ammonium acetate to a final concentration of 2.5 M in 100 µl (33 µl of 7.5 M ammonium acetate, 47 µl water, and 20 µl ligation mix). One hundred microlitres of ice-cold 100% ethanol were also added, mixed gently, and the DNA allowed to precipitate for 5 minutes at room temperature. Tubes were then centrifuged at 4°C for 10 minutes at 12,000 x*g*. After the supernate was removed, 100 µl of 70% ethanol were added to wash the pellet. The tubes were centrifuged as before. The ethanol was removed from the tubes and the pellets left to air-dry on the bench for approximately 15 to 30 minutes. The DNA was re-suspended in 100 µl of 1 x TE buffer and either used immediately for the PCR amplification (7.2.2, see below) or stored at –30°C until required.

7.2.2 PCR amplification of tagged DNA fragments

Selective primers were used to amplify a sub-set of the tagged DNA fragments produced, depending on the adapter sequence and the bases of the restricted DNA fragment directly adjacent to the adapter. These are listed in Table 7.1 (page 238).

Table 7.1. Restriction endonucleases, adapters and primers used for AFLP on *Bordetella pertussis*

Enzyme/Adapter/ Primer name	Recognition sequence	Supplier/ *Reference*
*Apa*I	GGGCC ↓ C	Integra Biosciences
Apa-1	TCGTAGACTGCGTACAGGCC	*Janssen et al. (1996)*
Apa-2	CATCTGACGCATGT	*Janssen et al. (1996)*
Apa-A	GACTGCGTACAGGCC**A**	*Janssen et al. (1996)*
Apa-C	GACTGCGTACAGGCC**C**	This thesis*
Apa-G	GACTGCGTACAGGCC**G**	*Janssen et al. (1996)*
Apa-T	GACTGCGTACAGGCC**T**	This thesis*
Apa-AT	GACTGCGTACAGGCC**AT**	This thesis*
Apa-CG	GACTGCGTACAGGCC**CG**	This thesis*
Apa-GC	GACTGCGTACAGGCC**GC**	This thesis*
Apa-TA	GACTGCGTACAGGCC**TA**	This thesis*
*Eco*RI	G↓AATTC	Roche
Eco-1	CTCGTAGACTGCGTACC	*Janssen et al. (1996)*
Eco-2	AATTGGTACGCAGTC	*Janssen et al. (1996)*
Eco-C	GACTGCGTACCAATTC**C**	*Janssen et al. (1996)*
Eco-G	GACTGCGTACCAATTC**G**	This thesis*
Eco-T	GACTGCGTACCAATTC**T**	*Vos et al. (1995)*
*Hind*III	A↓AGCTT	Roche
Hind-1	CTCGTAGACTGCGTACC	*Janssen et al. (1996)*
Hind-2	AGCTGGTACGCAGTC	*Janssen et al. (1996)*
Hind-A	GACTGCGTACCAGCTT**A**	*Janssen et al. (1996)*
*Mse*I	T↓TAA	New England BioLabs
Mse-1	GACGATGAGTCCTGAG	*Janssen et al. (1996)*
Mse-2	TACTCAGGACTCATC	*Janssen et al. (1996)*
Mse-A	GATGAGTCCTGAGTAA**A**	*Janssen et al. (1996)*
Mse-C	GATGAGTCCTGAGTAA**C**	*Janssen et al. (1996)*
Mse-G	GATGAGTCCTGAGTAA**G**	*Vos et al. (1995)*
Mse-T	GATGAGTCCTGAGTAA**T**	*Janssen et al. (1996)*
Mse-AT	GATGAGTCCTGAGTAA**AT**	*Vos et al. (1995)*
*Pst*I	CTGCA↓G	Roche
Pst-Lg1	CTCGTAGACTGCGTACATGCA	*Valsangiacomo et al. (1995)*
Pst-Lg2	TGTACGCAGTCTAC	*Valsangiacomo et al. (1995)*
Pst	GACTGCGTACATGCAG	*Valsangiacomo et al. (1995)*
Pst-A	GACTGCGTACATGCAG**A**	*Valsangiacomo et al. (1995)*
Pst-C	GACTGCGTACATGCAG**C**	*Carter et al. (1998)*
Pst-G	GACTGCGTACATGCAG**G**	*Valsangiacomo et al. (1995)*
Pst-T	GACTGCGTACATGCAG**T**	*Carter et al. (1998)*
*Taq*I	T↓CGA	Roche
Taq-1	GACGATGAGTCCTGAC	*Janssen et al. (1996)*
Taq-2	CGGTCAGGACTCAT	*Janssen et al. (1996)*
Taq-A	CGATGAGTCCTGACCGA**A**	*Janssen et al. (1996)*

Enzymes are listed in red with their recognition sequence and the supplier used. The adapters named in orange have a numerical suffix, whilst the amplification primers, listed in green, have the selected base(s) added in bold. References are italicised.

*Selective primers were designed with the knowledge of the core sequence from Janssen *et al.* (1996), and the selective bases were chosen for this work.

Ready-To-Go PCR beads (Amersham) were used. The reaction mixtures consisted of 20 mM Tris-HCl (pH 8.3), 50 mM KCl, 2.5 mM $MgCl_2$, 200 μM each dNTP, 75 ng primer, 1-1.5 U *Taq* DNA polymerase, and approximately 1 ng of template DNA (5 μl of a 1 in 100 dilution of product from 7.2.1). The final volume (25 μl) was overlaid with mineral oil (Sigma) before transferring the tubes to a thermal cycler. Cycling conditions were as follows: an initial denaturation at 94°C for 4 minutes, then 33 cycles of denaturation of 94°C for 1 minute, annealing of 60°C for 1 minute, and extension of 72°C for 2.5 minutes, with a final 4°C hold. The amplified products were then separated by standard horizontal gel electrophoresis on a 1.5% agarose gel in 1 x TBE buffer for 4 hours at 3.45 V/cm. Five microlitres of each reaction mix were loaded in each lane, with 0.75 μg/lane of GeneRuler™ DNA Ladder Mix (MBI Fermentas) as the molecular size marker loaded in the first and last lanes. Gels were stained in 1 μg/ml of ethidium bromide, and a photograph was taken under UV light (305 nm).

7.2.3 Amplified Fragment Length Polymorphism analysis

Banding patterns were analysed visually. Distinct patterns were recorded, dependent on the number and position of bands present in each profile.

7.3 Results

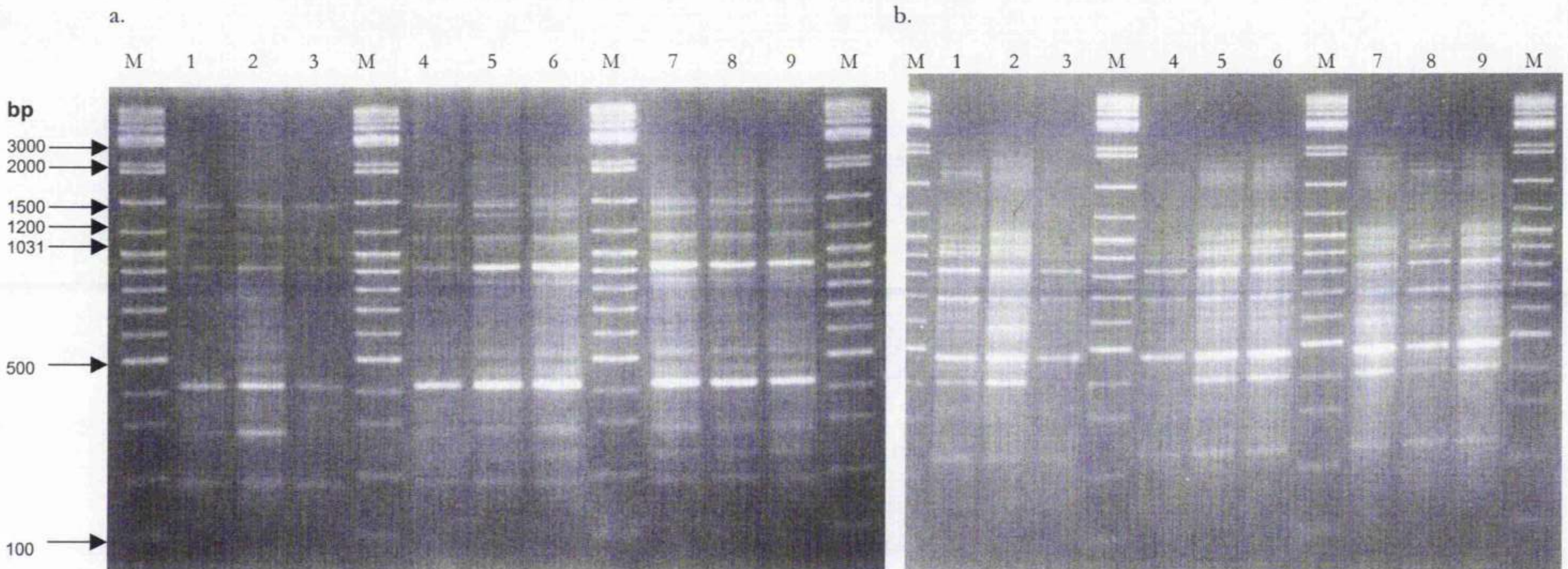
Six different enzymes and 23 primers were used to produce AFLP profiles of nine isolates (Tables 7.1 and 7.2, pages 238 and 240), resolved by standard gel electrophoresis. Eighteen of the 23 primers generated profiles, although some were faint and others revealed similar profiles for all the nine isolates tested.

The restriction endonuclease *Apa*I digest was successful, and six of the eight primers (Table 7.1, page 238) resulted in AFLP profiles. The primers Apa-T and Apa-TA did not generate any profiles, and Apa-A and Apa-AT produced only faint profiles. The other four primers (Apa-C, Apa-G, Apa-CG and Apa-GC) each gave varying intensities between the isolates, but generated the same profile for all the isolates tested. Figure 7.2 (page 241) shows the profiles produced from Apa-C and Apa-GC. Therefore there was no diversity in the profiles using a number of selective primers for *Apa*I-restricted digests.

Table 7.2. Type designations of the *Bordetella pertussis* isolates from Missouri, USA and loading order of the AFLP gels, shown in Figures 7.2 to 7.6

Lane number	Missouri isolate	Temporal sub-set	Year of isolation	Serotype	Pertactin gene type	Pertussis toxin S1 gene type	IS*1002*-RFLP type	PFGE type (*Xba*I)
1	MO-1503	"Recent"	1999	1,3	*prnA*(1)	*ptxA*(2)	RFLP-27	PFGE-21
2	MO-83	"Recent"	1999	1,3	*prnA*(2)	*ptxA*(1)	RFLP-1	PFGE-3
3	MO-234	"Recent"	1999	1,3	*prnA*(2)	*ptxA*(1)	RFLP-1	PFGE-2
4	MO-1440	"Kansas City"	1999	1,3	*prnA*(2)	*ptxA*(1)	RFLP-1	PFGE-3
5	MO-1482	"Kansas City"	1999	1,3	*prnA*(2)	*ptxA*(1)	RFLP-1	PFGE-2
6	MO-73	"Kansas City"	1999	1,3	*prnA*(2)	*ptxA*(1)	RFLP-1	PFGE-3
7	MO-20	"Early"	1984	1,3	*prnA*(6)	*ptxA*(2)	RFLP-25	PFGE-16
8	MO-814	"Early"	1984	1,2	*prnA*(1)	*ptxA*(1)	RFLP-11	PFGE-1
9	MO-908	"Early"	1988	1,3	*prnA*(9)	*ptxA*(1)	RFLP-1	PFGE-14

Figure 7.2. AFLP profiles of *Bordetella pertussis*, using *Apa*I and selective primers Apa-C (a) and Apa-GC (b)



M=GeneRuler DNA Ladder Mix. Details of lanes 1-9 are listed in Table 7.2 (page 240).

Of the three *Eco*RI selective primers attempted, only one, Eco-T, did not produce any AFLP profiles. The primers Eco-C and Eco-G (Figure 7.3, page 243) gave varying intensities for the isolates tested, and in the most intense profiles, Eco-C produced eight bands for isolate MO-1503 (lane 1) and Eco-G generated 18 bands for MO-234 (lane 3). The profiles appear different from isolate to isolate, and differences in DNA concentration may be the cause. So when analysing in more detail it is apparent that all isolates for both primers Eco-C and Eco-G generate the same AFLP profile.

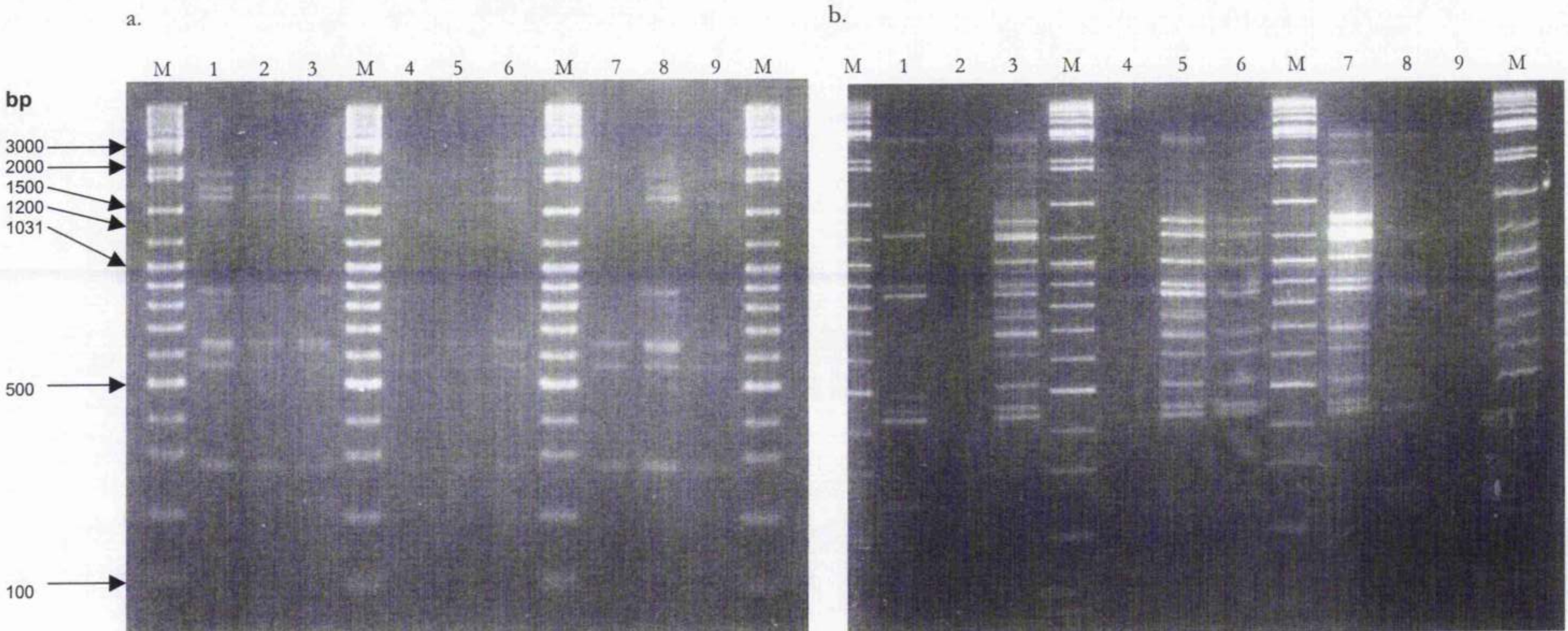
Only one primer was attempted for the enzyme *Hind*III (primer Hind-A), and either the restriction digest or the amplification with the selective primer was unsuccessful in generating any AFLP profiles.

All five selective primers for *Mse*I-restricted profiles contained between two and 15 bands. Primers Mse-C, Mse-T, and Mse-AT, whilst exhibiting between only two and seven bands in a typical AFLP profile, generated at least two different AFLP profiles. This is seen clearly when comparing two earlier Missouri isolates, MO-20 (lane 7) and MO-814 (lane 8) using Mse-AT (Figure 7.4.b, page 244). The most promising enzyme-primer combination Mse-A, amplifies up to 16 bands in an isolate, but generates only two different types of AFLP profiles (Figure 7.4.a, page 244). This difference in profiles is due to an additional band of approximately 1100 bp, visible in isolates MO-1482 (lane 5) and MO-908 (lane 9).

Four selective primers for the enzyme *Pst*I were tested. The primers Pst-A (Figure 7.5.a, page 245) and Pst-T generated only a few bands, whilst primer Pst-G produced a large number of bands. Only one AFLP profile was generated with Pst-T and Pst-G for all the isolates tested. The best candidate for further investigation for epidemiological typing of *B. pertussis* was the primer Pst-C (Figure 7.5.b, page 245). This produced many bands, and there are at least four different AFLP profiles generated. Although the differences between the profiles were due to the presence or absence of bands, the intensities of the remaining bands were equal to those in profiles with a greater number of bands. This can be seen when comparing isolates MO-20, MO-814 and MO-908 (lanes 7, 8 and 9, Figure 7.5.b, page 245).

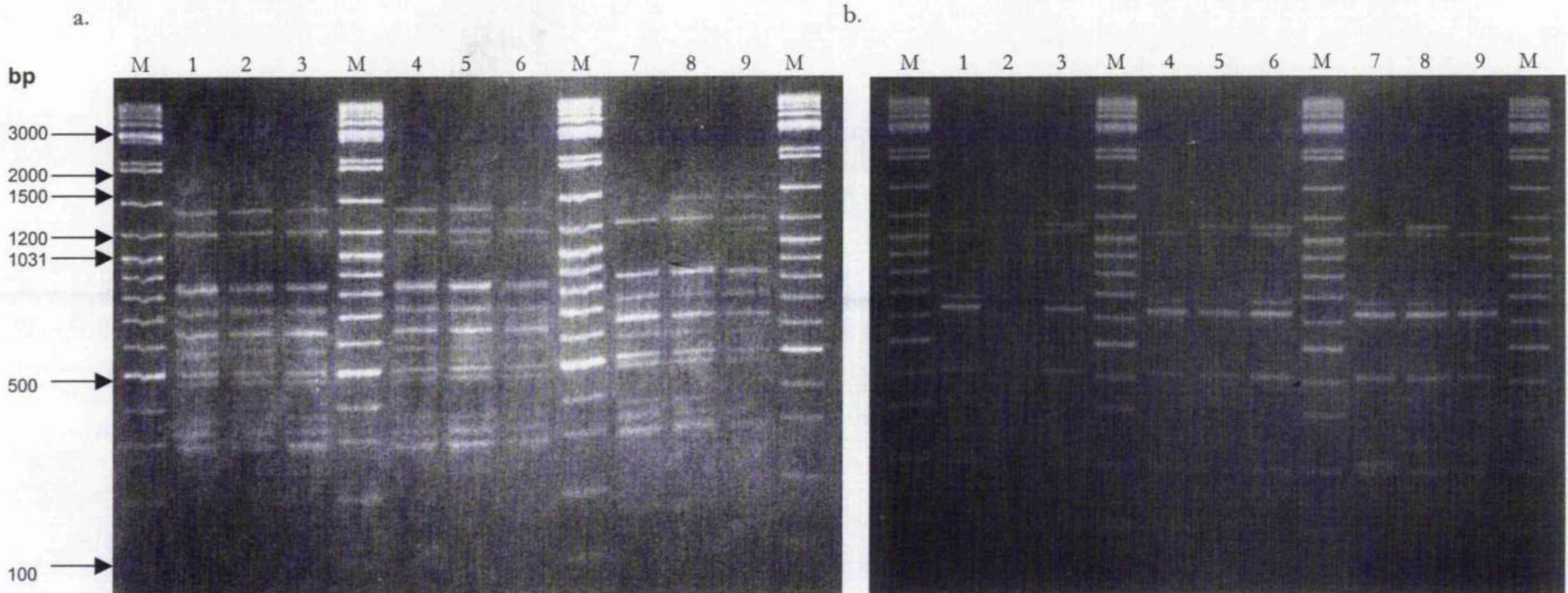
The only primer tested for *Taq*I, Taq-A, gave smeared, undefined profiles (Figure 7.6, page 246). The procedure was repeated a number of times, including an adjustment of the restriction endonuclease temperature from 37°C to 65°C, but results were not improved on the gel illustrated in Figure 7.6 (page 246).

Figure 7.3. AFLP profiles of *Bordetella pertussis*, using *Eco*RI and selective primers Eco-C (a) and Eco-G (b)



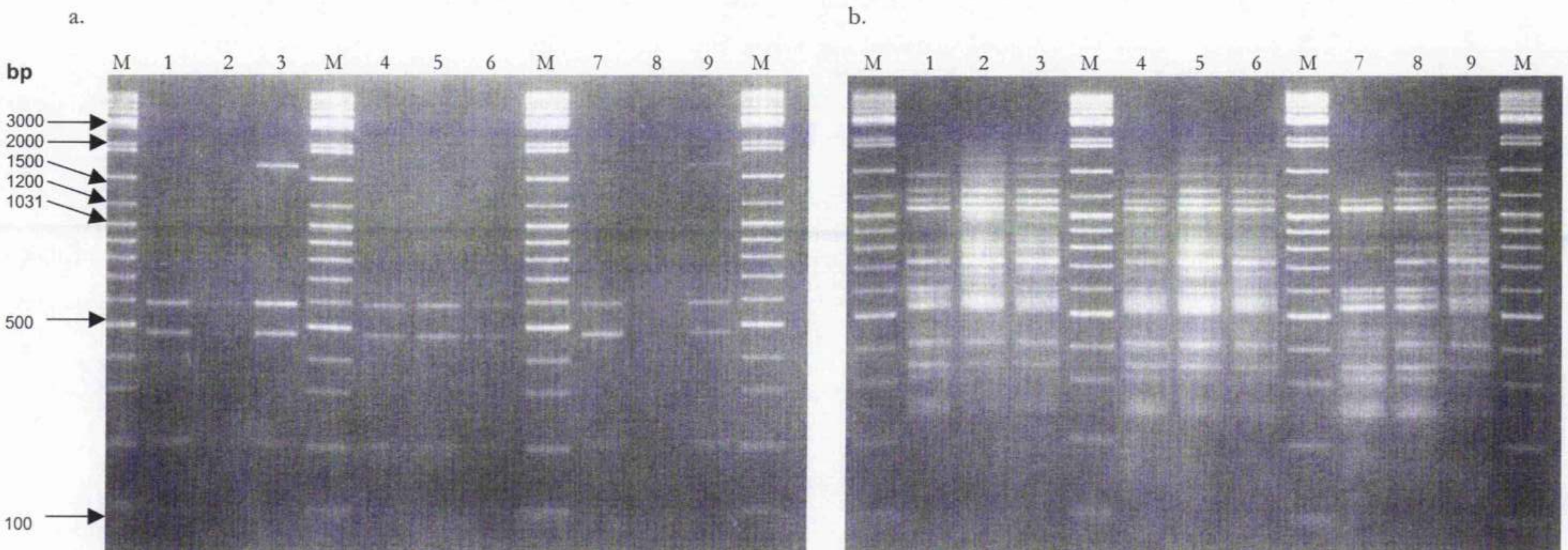
M=GeneRuler DNA Ladder Mix marker. Details of lanes 1-9 are listed in Table 7.2 (page 240).

Figure 7.4. AFLP profiles of *Bordetella pertussis*, using *Mse*I and selective primers Mse-A (a) and Mse-AT (b)



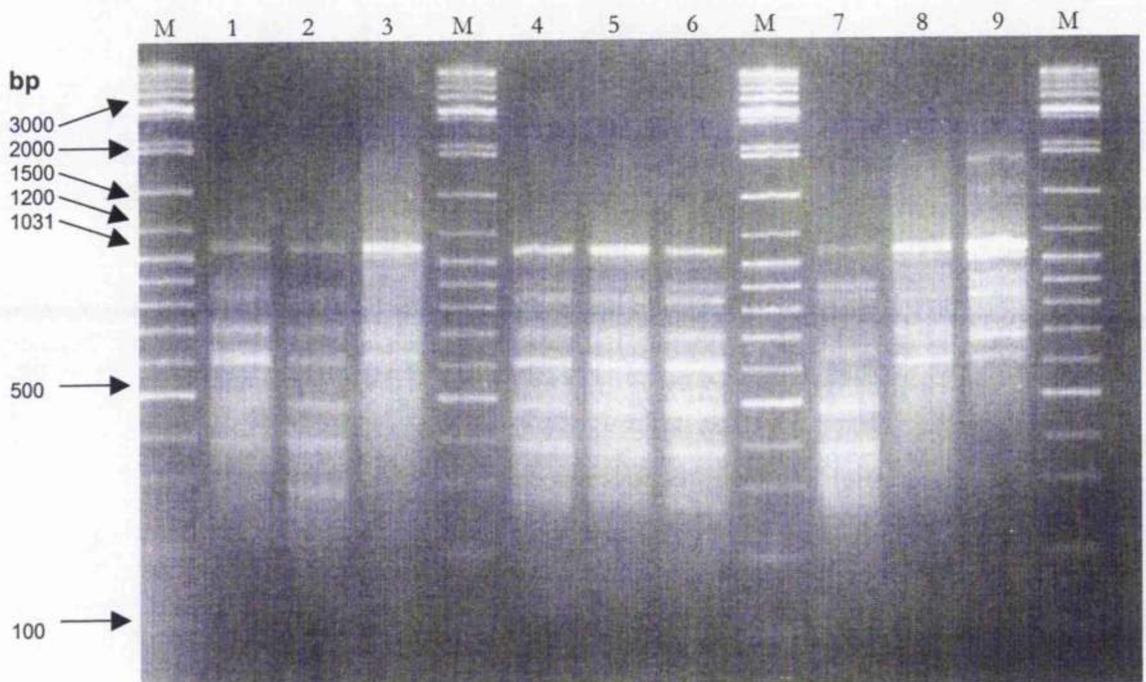
M=GeneRuler DNA Ladder Mix. Details of lanes 1-9 are listed in Table 7.2 (page 240).

Figure 7.5. AFLP profiles of *Bordetella pertussis*, using *Pst*I and selective primers Pst-A (a) and Pst-C (b)



M=GeneRuler DNA Ladder Mix. Details of lanes 1-9 are listed in Table 7.2 (page 240).

Figure 7.6. AFLP profiles of *Bordetella pertussis*, using *Taq*I and selective primer Taq-A



M=GeneRuler DNA Ladder Mix. Details of lanes 1-9 are in Table 7.2 (page 240).

The profiles from Mse-A, Mse-AT, Pst-A and Pst-C were investigated for associations with other typing data (*prnA*, *ptxA*, IS*1002*-RFLP and PFGE), as these enzyme-primer combinations generated different AFLP profiles (Table 7.2, page 240). Serotyping data were omitted from this investigation, as eight of the nine isolates were serotype 1,3.

The AFLP profiles of the two isolates MO-1482 and MO-908 shared the same AFLP type when Mse-A was used (Figure 7.4.a, page 244). They were also both *ptxA*(1) and RFLP-1, but differed in *prnA* and PFGE type and were from different time periods. Similarly, the Pst-A profile shared by MO-234 and MO-908 (Figure 7.5.a, page 245) was associated with *ptxA*(1) and RFLP-1, but differed in *prnA* and PFGE type and these two strains were from different time periods. Neither of the two Mse-AT profiles that divided the nine isolates (Figure 7.4.b, page 244), were associated with one particular *prnA*, *ptxA*, IS*1002*-RFLP or PFGE type.

The most discriminatory enzyme-primer combination, *Pst*I and Pst-C, divided the isolates into four groups; I, MO-1503 and MO-814; II, MO-83, MO-73 and MO-908; III, MO-234, MO-1440 and MO-1482 and; IV, MO-20 (Figure 7.5.b, page 245). Isolates from type I were both *prnA*(1), type II isolates were all *ptxA*(1) and RFLP-1, type III isolates were all *prnA*(2), *ptxA*(1) and RFLP-1, and the type IV isolate was *prnA*(6). All the isolates from types II and III were isolated in 1999.

7.4 Discussion

A novel typing method to *B. pertussis*, AFLP, was attempted on selected isolates from Missouri, USA. This was thought to be a potentially convenient and rapid approach, using restriction endonucleases and specific primers, for strain discrimination. Although this technique has proved useful for typing a wide range of bacteria, such as *L. pneumophila*, *Chlamydia psittaci* and *Clostridium difficile* (Fry *et al.*, 2000; Boumedine and Rodolakis, 1998; McLauchlin *et al.*, 2000), the results with six different restriction endonucleases and *B. pertussis* were largely unsatisfactory. The enzymes *Apa*I, *Pst*I and *Taq*I have G+C-rich recognition sequences, making them ideally suited as frequent cutters in the G+C-rich *B. pertussis* genome (67.7 mol %) (Parkhill *et al.*, 2003). Although the other enzymes were more suited to cleave A+T-rich organisms, frequent cutters such as *Apa*I and *Pst*I can produce too many bands, and may not give sufficient resolution for typing purposes. However, the selective primers either failed to amplify any restricted fragments (for example Apa-T, Eco-T and Hind-A), or produced identical AFLP profiles for the nine isolates tested (Apa-G,

Eco-C and Pst-G). Only a few combinations generated different AFLP profiles with a sufficient number of bands in the profile to be of any value. These included Mse-T, Mse-A and Pst-C, where the latter gave the best discrimination, producing four AFLP types in the nine isolates tested (Figure 7.5.b, page 245).

Although three of the four enzyme-primer combinations, *Mse*I and Mse-A or Mse-AT, and *Pst*I and Pst-A, generated different profiles, Mse-AT profiles were not associated with any other typing methods. Furthermore, the Mse-A and Pst-A profiles that were shared by only two isolates, were associated with *ptxA*(1) and RFLP-1, which were the two predominant types present in the nine isolates examined. Two of the four profiles produced by Pst-C were associated with *prnA* (group I and *prnA*(1), group IV and *prnA*(6)), and the three type III isolates were also *prnA*(2), *ptxA*(1) and RFLP-1. The latter Pst-C type, type III, shares these predominant types with combined type P (section 4.4.2, page 183), highlighting the success of this *B. pertussis* lineage. Although these associations are based on a small number of isolates, further work using Pst-C on a wider selection of isolates may identify more significant associations.

An *in silico* digest was performed for the techniques IS*1002*-RFLP analysis and PFGE using *Xba*I, using the recently annotated Tohama-I genome (5.3.6, page 202). This could have been performed for the single-enzyme digest and subsequent PCR amplification of each AFLP analysis, but the many enzymes and primers would have made this task arduous. Also, the Tohama-I strain was not used in the AFLP experiments in this work, so a direct comparison could not be made. However, with the appropriate *in silico* AFLP software, the Tohama-I genome could be screened for promising enzyme and primer combinations for further investigation of this technique.

Optimisation of the AFLP PCR conditions was not performed, and this may have affected the results quite significantly. The magnesium chloride concentration (2.5 mM) and the annealing temperature of 60°C were not altered. This temperature is relatively low for a *B. pertussis*-specific PCR assay due to the high G+C-rich genome. Also, most of the primers tested which had pyrimidine bases added such as Apa-AT and Pst-T gave faint profiles, and this also could be due to the lack of optimisation and the G+C-rich genome of *B. pertussis*. Other enzyme and primer combinations could have been investigated. In conclusion, single-enzyme AFLP, although flexible and simple, did not generate a substantial discrimination for *B. pertussis*, and further optimisation would be needed in order to apply this technique confidently to epidemiological investigations.

8 Epidemiological typing direct from UK clinical samples

8.1 Strategy

Although isolation of *B. pertussis* is considered to be the "gold standard" for laboratory diagnosis of pertussis, it is relatively insensitive and can take up to a week for visible colonies to appear. The detection of pertussis infection has been successful using other methods such as serology and PCR, but for epidemiological analyses, techniques such as PFGE and IS*1002*-RFLP require pure cultures of the organism to get satisfactory results. Another recent epidemiological typing method, based on pertactin and pertussis toxin S1 subunit gene variation by direct sequencing, has been described but used genomic DNA extracted from pure *B. pertussis* cultures (Mooi *et al.*, 1998; this thesis). Direct typing of these genes was therefore attempted on nasopharyngeal aspirates (NPAs) and pernasal swabs (PNSs), received by RSIL as part of the PICU study (Crowcroft *et al.*, 2003), that were known to be either culture- and/or PCR-positive for *B. pertussis*. Samples were detected as PCR-positive in the earlier PICU study, using two targets: the pertussis toxin promoter (with an internal positive control) and the IS*481* element (Fry *et al.*, 2004). To increase the sensitivity of detection of the *prnA* and *ptxA* genes, some optimisation and modifications to the published method (Mooi *et al.*, 1998) were necessary. If successful, this direct typing approach could yield new epidemiological data that was not possible with previous methods.

8.2 Methods

Extraction of DNA from all the clinical samples (NPAs and PNSs) was performed with the Qiagen Blood & Cell Culture DNA Mini Kit (Qiagen) (2.4.3, page 62). The descriptions of general methodology of PCR, purification of PCR products and the sequencing method are detailed in sections 2.5.2, 2.5.5 and 2.5.6 (pages 63, 76 and 77), respectively. DNA from the *B. pertussis* type strain (NCTC 10739^{T}) was used as the positive control and, for optimisation experiments for amplification of both the pertactin and pertussis toxin S1 genes, at 10 ng DNA per reaction, as quantified using the GeneQuant spectrophotometer (2.4.2, page 61).

8.2.1 Amplification of the pertactin gene

The amplification of the two polymorphic regions (Figure 1.2, page 16), using primers AF and BR (Table 2.3 and Figure 2.1, pages 66 and 69), was attempted (predicted amplicon size approximately 1428 bp) on: (i) four culture-positive, PCR-positive clinical samples and

(ii) 16 culture-negative, PCR-positive clinical samples, using 20 μl of clinical extract. The method is described in 2.5.4.1, page 68.

A smaller region, encompassing the major polymorphic region 1 was targeted, with primers AF and AR, to generate a product of approximately 585 bp. This was attempted on one culture-positive, PCR-positive clinical sample and 15 culture-negative, PCR-positive clinical samples. Reaction mixtures (50 μl) contained 10% v/v DMSO, 2.5 mM $MgCl_2$, 10 mM Tris-HCl (pH 8.3), 2.5 mM KCl, 200 μM dNTP, 10 pmol each of AF and AR primers, 2.5 U *Taq* DNA polymerase, and 20 μl of clinical extract. Optimisation of the cycling conditions were as follows, initial denaturation at 95°C for 5 minutes, then 40 cycles consisting of denaturation at 95°C for 15 seconds, annealing at 48-68°C for 15 seconds, and extension at 72°C for 45 seconds, followed by a final extension of 7 minutes at 72°C, with a final 4°C hold.

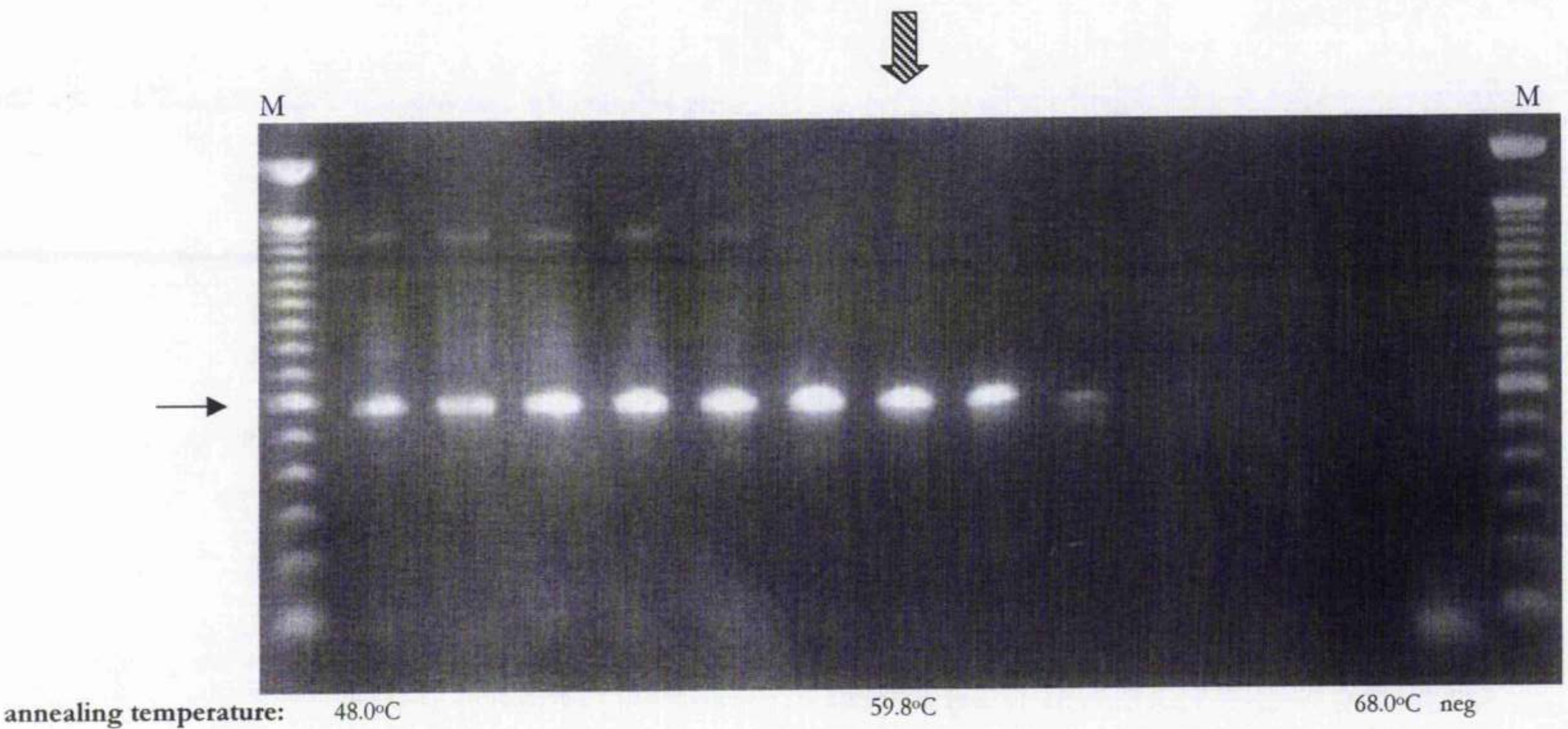
<u>Optimised conditions with primers AF and AR</u>

Successful amplification (Figure 8.1, page 253) was achieved with the above reaction mixtures and the following cycling conditions; initial denaturation at 95°C for 5 minutes, then 40 cycles consisting of denaturation at 95°C for 15 seconds, annealing at 60°C for 15 seconds, and extension at 72°C for 45 seconds, followed by a final extension of 7 minutes at 72°C, with a final 4°C hold.

8.2.2 Amplification of the pertussis toxin S1 gene

The primers SF and SR were used to generate a product of approximately 876 bp (Table 2.3, page 66), using the method described in 2.5.4.2, page 73. The amplification of this *ptxA* fragment was attempted on one culture-positive, PCR-positive clinical sample and 16 culture-negative, PCR-positive clinical samples, using 20 μl of clinical extract.

Figure 8.1. PCR amplification of the *Bordetella pertussis* pertactin gene with primers AF and AR, at various annealing temperatures



M=100 bp ladder (Invitrogen). Ten ng of DNA was used from the type strain of *Bordetella pertussis* (NCTC 10739[T]), neg=negative control. The annealing temperature was varied from 48.0 to 68.0°C, from left to right on the gel. The black arrow highlights the predicted amplicon size of 585 bp and the hatched arrow indicates the optimal reaction used for subsequent experiments.

A smaller region encompassing known polymorphic regions of the pertussis toxin S1 gene was also targeted with primers S1FM and SR (Table 2.3, page 66), to generate a smaller amplicon of approximately 404 bp (Figure 2.5, page 75). Reaction mixtures of 50 µl contained 1.5 mM $MgCl_2$, 10 mM Tris-HCl (pH 8.3), 2.5 mM KCl, 200 µM dNTP, 10 pmol each of S1FM and SR primers, 2.5 U *Taq* DNA polymerase, and 20 µl of clinical extract. Amplification parameters for optimisation consisted of initial denaturation at 95°C for 5 minutes, then 40 cycles of 95°C for 15 seconds, annealing at 48-68°C for 15 seconds, and 72°C for 45 seconds, with a final extension of 72°C for 7 minutes, and a final 4°C hold. Amplification of this smaller *ptxA* region was attempted on one culture-positive, PCR-positive clinical sample and 15 culture-negative, PCR-positive clinical samples.

Optimised conditions with primers S1FM and SR

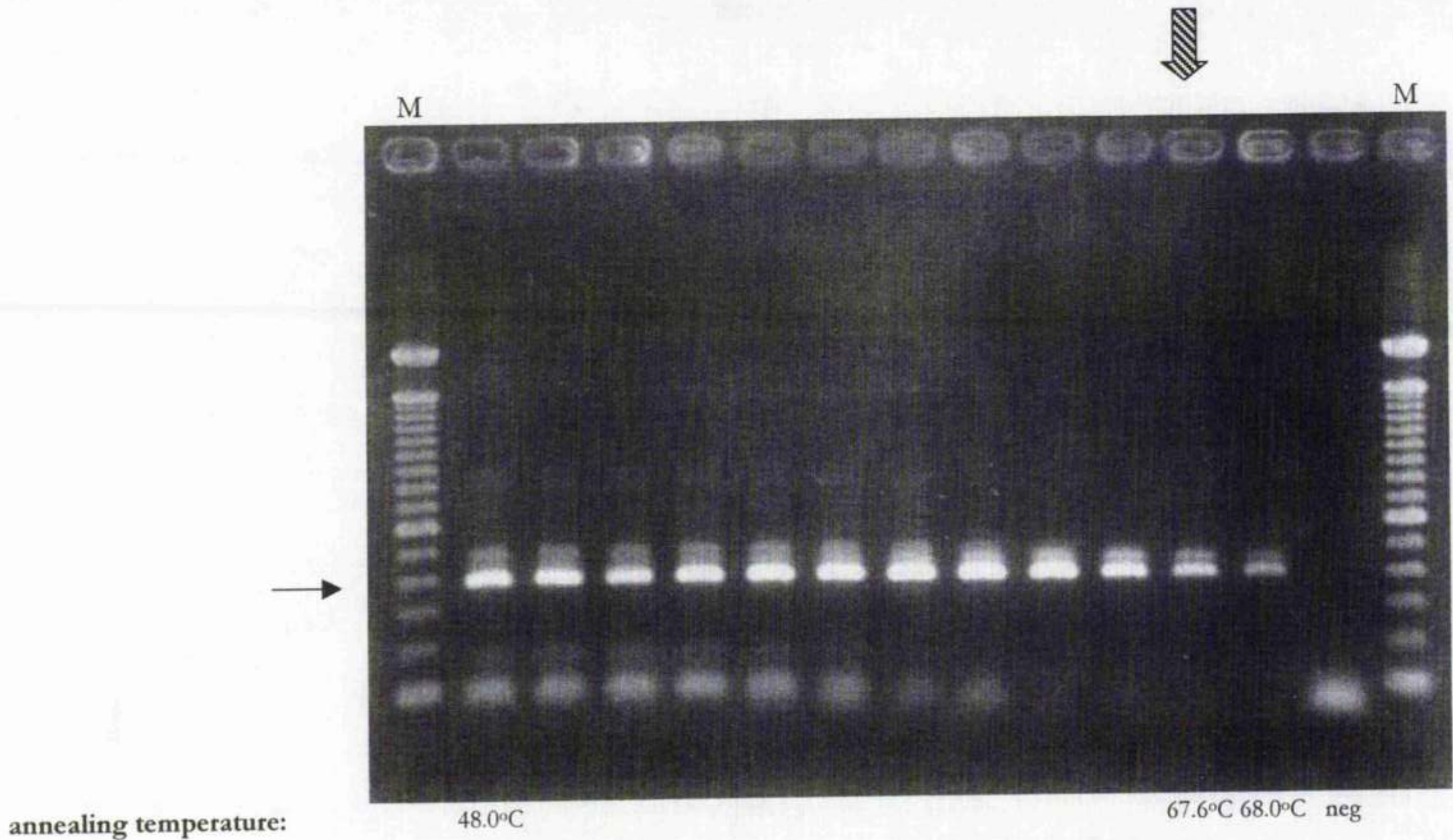
The optimised cycling conditions consisted of an initial denaturation at 95°C for 5 minutes, then 40 cycles of denaturation at 95°C for 15 seconds, annealing at 67°C for 15 seconds, and extension at 72°C for 15 seconds, followed by a final extension of 7 minutes at 72°C, with a final 4°C hold (Figure 8.2, page 255).

8.3 Results

8.3.1 Detection of the pertactin gene, with primers AF and BR, in clinical samples from the UK

An initial attempt to detect the pertactin gene was investigated with a selected group (n=5) of clinical specimens: PICU 519, PICU 521, PICU 522, PICU 523 and PICU 524 (Table 8.1, page 256). All specimens were previously shown to be culture-positive and PCR-positive for *B. pertussis*, except PICU 521, which was culture-negative and PCR-positive. The visualisation of the amplicons generated are shown in Figure 8.3 (page 257). Three of the five specimens tested were positive for pertactin (PICU 519, PICU 523 and PICU 524). Unfortunately, the PCR products from clinical samples were unavailable for further investigation to determine the nucleotide sequence, although sequences were determined previously using DNA extracts from pure culture isolates from the same patient.

Figure 8.2. PCR amplification of the ***Bordetella pertussis*** pertussis toxin S1 gene with primers S1FM and SR, at various annealing temperatures



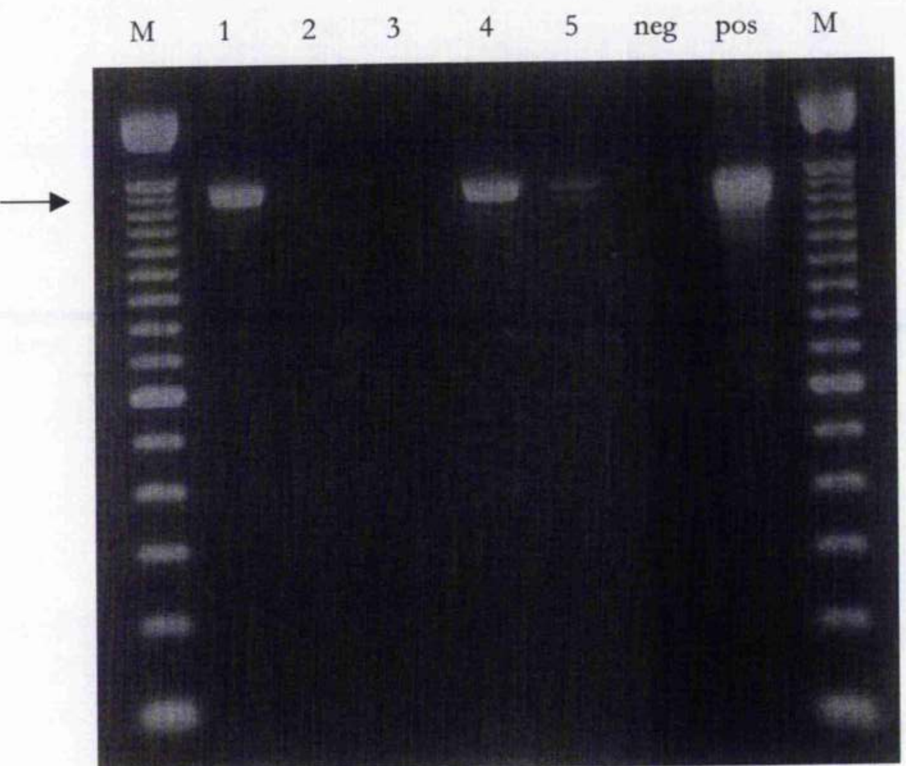
M=100 bp ladder (Invitrogen). Ten ng of DNA was used from the type strain of *Bordetella pertussis* (NCTC 10739^T), neg=negative control. The annealing temperature was varied from 48.0 to 68.0°C, from left to right on the gel. The black arrow highlights the predicted amplicon size of 404 bp and the hatched arrow indicates the optimal reaction used for subsequent experiments.

Table 8.1. ***Bordetella pertussis*** **clinical specimens investigated for PCR amplification of the pertactin gene, as visualised in Figure 8.3**

		Results from PICU study[1]			Results from this study
Lane number	**Clincial specimen**	**Culture result**	**Pertussis toxin promoter PCR**	**IS*481* PCR**	**Pertactin PCR (AF-BR)**
1	PICU 519	Positive	Positive	Positive	Positive
2	PICU 521	Negative	Positive	Positive	Negative
3	PICU 522	Positive	Positive	Positive	Negative
4	PICU 523	Positive	Positive	Positive	Positive
5	PICU 524	Positive	Positive	Positive	Positive

[1]Results obtained from Fry *et al.* (2004)

Figure 8.3. Detection of the *Bordetella pertussis* pertactin gene with primers AF and BR, in UK clinical specimens



M=100 bp ladder (Invitrogen), neg=negative control, pos=10 ng DNA from the type strain of *Bordetella pertussis*, NCTC 10739^T. The specimens loaded in lanes from 1 to 5 are listed in Table 8.1 (page 256). The black arrow highlights the predicted amplicon size of 1428 bp.

As amplification of the pertactin gene was shown to be worthwhile, a second round was attempted on clinical specimens that were previously culture-negative, but PCR-positive (n=15) (Table 8.2 and Figure 8.4, pages 259 and 260). Only two of the 15 specimens were positive (PICU 705 and PICU 708) with primers AF and BR, and nucleotide sequences were determined as described in 2.5.6.2 (page 79).

8.3.2 Detection of the pertussis toxin S1 gene, using primers SF and SR, in UK clinical samples

The detection of *ptxA* was performed with the same clinical specimens as above (n=15) (Table 8.2 and Figure 8.5, pages 259 and 261). Amplification of *ptxA* was marginally more successful than *prnA*, i.e., six of the 15 samples were positive, which included the two clinical extracts that were also positive for *prnA* (PICU 705 and PICU 708). Nucleotide sequences were determined for these six products as described in 2.5.6.2 (page 79).

8.3.3 Detection of the pertactin and pertussis toxin S1 genes, with primers AF and AR, and S1FM and SR, respectively, in UK clinical samples

To try to increase the sensitivity of the two targets pertactin and pertussis toxin S1, smaller regions were attempted and optimised, as described earlier (8.2, page 251). The amplification of pertactin with primers AF and AR, was applied to previously culture-negative and PCR-positive clinical specimens (n=15), and one culture- and PCR-positive clinical specimen (Table 8.3, page 262). Figure 8.6 (page 263) shows the successful *prnA* amplification of the positive control and six of the 16 clinical extracts tested. Amplification of the pertactin gene from specimens PICU 643, PICU 644 and PICU 666 with primers AF and BR was previously unsuccessful. Nucleotide sequences were determined as described in 2.5.6.2 (page 79).

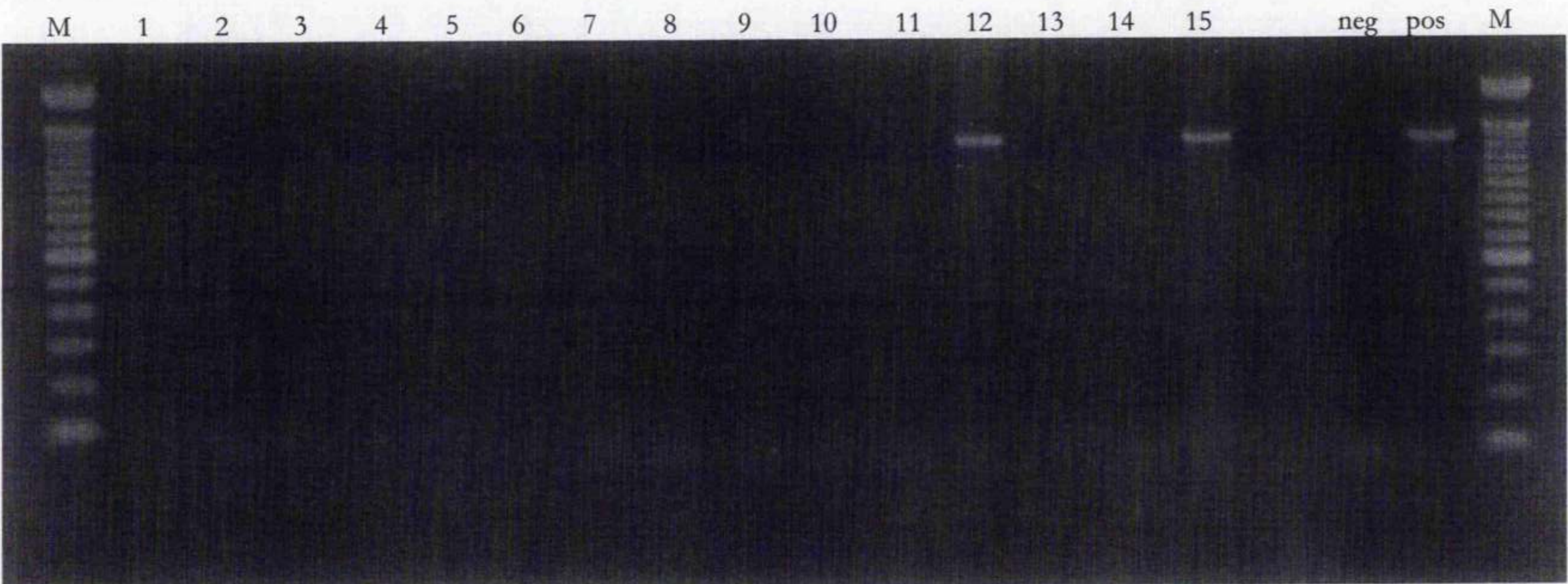
The optimisation of *ptxA* using primers S1FM and SR was successful, but the PCR failed when it was applied on the same 16 clinical specimens used for the AF-AR *prnA* amplification (Table 8.3, page 262). A further identical optimisation experiment was successful and identified the same optimum annealing temperature, but when the clinical specimens were tested a second time, PCR amplification was unsuccessful. Investigations using the same clinical extracts could not be attempted a third time, due to insufficient material.

Table 8.2. *Bordetella pertussis* clinical specimens investigated for PCR amplification of the pertactin and pertussis toxin S1 genes, as visualised in Figure 8.4 and 8.5

		Results from PICU study[1]		Results from this study	
Lane number	**Clinical specimen**	**Pertussis toxin promoter PCR**	**IS*481* PCR**	**Pertactin PCR (AF-BR)**	**Pertussis toxin S1 PCR (SF-SR)**
1	PICU 617	Positive	Positive	Negative	Negative
2	PICU 642	Positive	Positive	Negative	Negative
3	PICU 643	Positive	Positive	Negative	Positive
4	PICU 644	Positive	Positive	Negative	Positive
5	PICU 649	Positive	Positive	Negative	Negative
6	PICU 650	Positive	Positive	Negative	Negative
7	PICU 654	Negative	Positive	Negative	Negative
8	PICU 655	Negative	Positive	Negative	Negative
9	PICU 656	Positive	Positive	Negative	Negative
10	PICU 666	Positive	Positive	Negative	Positive
11	PICU 667	Positive	Positive	Negative	Positive
12	PICU 705	Positive	Positive	Positive	Positive
13	PICU 706	Positive	Positive	Negative	Negative
14	PICU 707	Positive	Positive	Negative	Negative
15	PICU 708	Positive	Positive	Positive	Positive

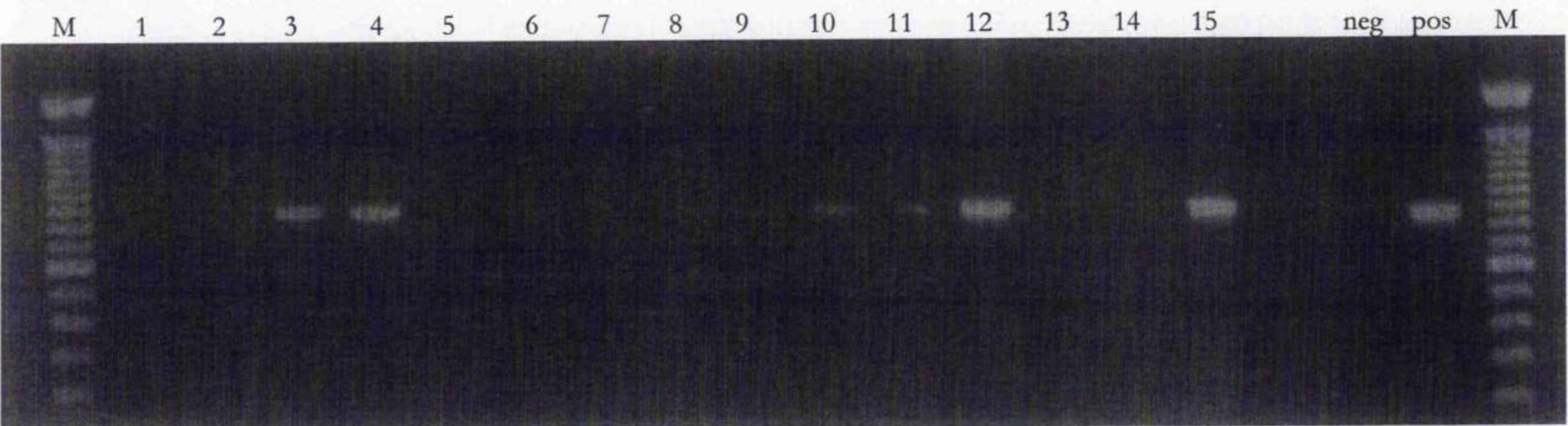
[1]Results obtained from Fry *et al.* (2004). Isolates were previously culture negative.

Figure 8.4. Detection of the *Bordetella pertussis* pertactin gene with primers AF and BR, in UK clinical specimens



M=100 bp ladder (Invitrogen), neg=negative control, pos=10 ng DNA from the type strain of *Bordetella pertussis*, NCTC 10739^{T}. The specimens loaded in lanes from 1 to 15 are listed in Table 8.2 (page 259). The black arrow highlights the predicted amplicon size of 1428 bp.

Figure 8.5. Detection of the *Bordetella pertussis* pertussis toxin S1 gene with primers SF and SR, in UK clinical specimens



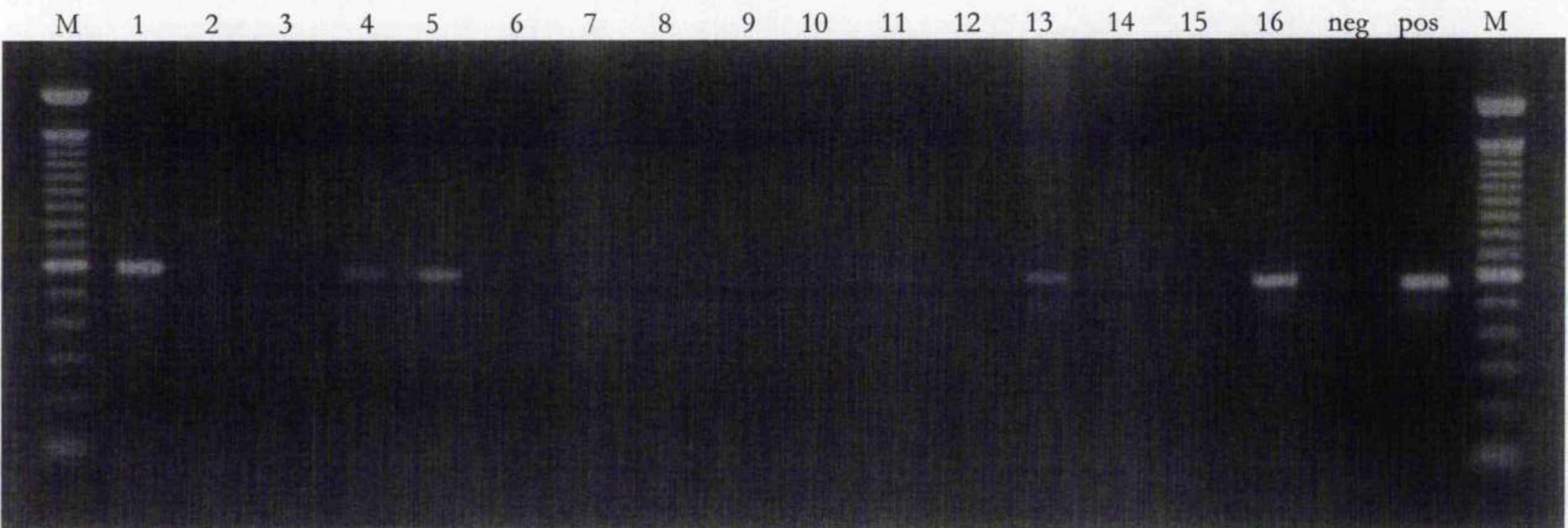
M=100 bp ladder (Invitrogen), neg=negative control, pos=10 ng DNA from the type strain of *Bordetella pertussis*, NCTC 10739T. The specimens loaded in lanes from 1 to 15 are listed in Table 8.2 (page 259). The black arrow highlights the predicted amplicon size of 876 bp.

Table 8.3. *Bordetella pertussis* clinical DNA extracts investigated for detection of the pertactin gene, as visualised in Figure 8.6

		Results from PICU study[1]		**Results from this project**
Lane number	**Clinical specimen**	**Pertussis toxin promoter PCR**	**IS*481* PCR**	**Pertactin PCR (AF-AR)**
1	PICU 519[2]	Positive	Positive	Positive
2	PICU 617	Positive	Positive	Negative
3	PICU 642	Positive	Positive	Negative
4	PICU 643	Positive	Positive	Positive
5	PICU 644	Positive	Positive	Positive
6	PICU 649	Positive	Positive	Negative
7	PICU 650	Positive	Positive	Negative
8	PICU 654	Negative	Positive	Negative
9	PICU 655	Negative	Positive	Negative
10	PICU 656	Positive	Positive	Negative
11	PICU 666	Positive	Positive	Positive
12	PICU 667	Positive	Positive	Negative
13	PICU 705	Positive	Positive	Positive
14	PICU 706	Positive	Positive	Negative
15	PICU 707	Positive	Positive	Negative
16	PICU 708	Positive	Positive	Positive

[1]Results obtained from Fry *et al.* (2004).
[2]Isolate PICU 519 was also previously culture-positive.

Figure 8.6. Detection of the *Bordetella pertussis* pertactin gene with primers AF and AR, in UK clinical specimens



M=100 bp ladder (Invitrogen), neg=negative control, pos=10 ng DNA from the type strain of *Bordetella pertussis*, NCTC 10739^{T}. The specimens loaded in lanes from 1 to 16 are listed in Table 8.3 (page 262). The black arrow highlights the predicted amplicon size of 585 bp.

8.3.4 Pertactin and pertussis S1 gene types generated from PCR-positive clinical specimens

All available PCR products generated were investigated further to determine their nucleotide sequence. The pertactin and pertussis toxin S1 gene types are recorded in Table 8.4 (page 265). The two clinical specimens, PICU 705 and PICU 707, generated amplicons using both AF and BR, and AF and AR, and produced the same pertactin gene type, *prnA*(2) when sequenced. The remaining three AF-AR *prnA* amplicons from clinical specimens PICU 643, PICU 644 and PICU 666, were also *prnA*(2), and these specimens, as well as PICU 705 and PICU 707, all generated the same pertussis toxin S1 type, *ptxA*(1). These five clinical specimens, previously non-culturable, generated *prnA* and *ptxA* results, which were then used for the UK *B. pertussis* population dynamics study, detailed in Chapter 3.

The pertactin gene type had been determined previously for the three clinical specimens, PICU 519, PICU 523 and PICU 524, using DNA extracted from pure cultures. Unfortunately, when investigated directly from the clinical specimen with primers AF and BR, the pertactin gene type was not identified and therefore could not be compared. The pertactin gene type, *prnA*(2), was however, generated from the clinical extract of PICU 519 using the primer pair AF and AR, and this matched the *prnA* type as identified using AF and BR on the DNA extract from culture.

Although the clinical specimen PICU 667 was faintly positive for pertussis toxin S1 amplification (Figure 8.5, page 261), the purified amplicon failed to give a nucleotide sequence when analysed on the CEQ™ 2000 (Beckman Coulter). This negative result was probably due to the insufficient concentration of amplicon, as visualised in Figure 8.5 (page 261, lane 11).

Table 8.4. Pertactin and pertussis toxin S1 gene types from clinical specimens

Clinical specimen	Pertactin type from AF-BR PCR	Pertactin type from AF-AR PCR	Pertussis toxin S1 type from SF-SR PCR
PICU 519[1]	Positive, but amplicon not available	*prnA*(2)	n/a
PICU 521[1]	Amplicon not generated	n/a	n/a
PICU 522[1]	Amplicon not generated	n/a	n/a
PICU 523[1]	Positive, but amplicon not available	n/a	n/a
PICU 524[1]	Positive, but amplicon not available	n/a	n/a
PICU 617	Amplicon not generated	Amplicon not generated	Amplicon not generated
PICU 642	Amplicon not generated	Amplicon not generated	Amplicon not generated
PICU 643	Amplicon not generated	*prnA*(2)	*ptxA*(1)
PICU 644	Amplicon not generated	*prnA*(2)	*ptxA*(1)
PICU 649	Amplicon not generated	Amplicon not generated	Amplicon not generated
PICU 650	Amplicon not generated	Amplicon not generated	Amplicon not generated
PICU 654	Amplicon not generated	Amplicon not generated	Amplicon not generated
PICU 655	Amplicon not generated	Amplicon not generated	Amplicon not generated
PICU 656	Amplicon not generated	Amplicon not generated	Amplicon not generated
PICU 666	Amplicon not generated	*prnA*(2)	*ptxA*(1)
PICU 667	Amplicon not generated	Amplicon not generated	failed[2]
PICU 705	*prnA*(2)	*prnA*(2)	*ptxA*(1)
PICU 706	Amplicon not generated	Amplicon not generated	Amplicon not generated
PICU 707	Amplicon not generated	Amplicon not generated	Amplicon not generated
PICU 708	*prnA*(2)	*prnA*(2)	*ptxA*(1)

[1]Pertactin and pertussis toxin S1 gene types had been determined previously using DNA extracts from pure culture.
[2]Denotes that although PICU 667 is positive in Figure 8.5 (page 261), there was insufficient amplified DNA to generate adequate sequence data.
n/a=not attempted.

8.4 Discussion

Simultaneous detection and epidemiological typing has proved valuable for bacteria, such as *N. meningitidis* (Molling *et al.*, 2002), and other organisms, such as HIV and hepatitis C virus (HCV) (Re *et al.*, 2001; Gargiulo *et al.*, 2003) that are difficult to isolate. To date, this direct typing approach has not been performed for *B. pertussis*. Previous epidemiological methods applied to *B. pertussis*, such as serotyping and PFGE, require bacterial cells and large amounts of DNA, respectively. However, determining the genotypic variation of virulence factors, such as the pertactin and pertussis toxin S1 subunit by PCR detection and DNA sequencing, should be possible, as this utilises the PCR technique, which amplifies the target gene required to determine the nucleotide sequence.

Work presented in this thesis demonstrates that direct typing of *B. pertussis* in clinical specimens is possible and worthwhile. Three of the four culture-positive specimens were typed directly from the clinical extract. Furthermore, two of the 15 culture-negative clinical specimens gave *prnA* types and five generated *ptxA* types, producing *B. pertussis* genotypes that would normally not be deduced due to the lack of isolate material. To increase the sensitivity, smaller fagments which still encompassed the polymorphic region of interest, were targeted for both *prnA* and *ptxA*. A further three *prnA* types were generated from culture-negative clinical extracts (n= 5/15 clinical extracts) but, unfortunately, the respective *ptxA* amplicon failed to generate any gene type for the clinical extracts tested.

These two methods could be improved further to increase the sensitivity when using clinical extracts directly. For example, by using nested PCR, this would involve a first round amplification for both *prnA* and *ptxA* genes then, using internal primers, a second amplification that encompasses the polymorphic regions would be done. Amplification problems may occur due to the rich GC content of the *B. pertussis* genome, which could affect the annealing temperature. Hence, the second amplification may require a higher annealing temperature than the first round PCR, to prevent any first round primers from annealing, and this higher temperature could initiate the extension of DNA. Therefore, various optimisation and titration experiments would have to be performed. Another approach would be real-time PCR technology with specific hybridisation probes for each *prnA* and *ptxA* type. This has been developed for both *prnA* and *ptxA* by Makinen *et al.* (2001 and 2002), although hybridisation probes have not been designed for each known *prnA* and *ptxA* type, and is reliant on a negative result as representing certain types, including *ptxA*(1). This would not be adequate when testing clinical extracts that may have

insufficient DNA to amplify; if they are negative, they would be reported as *ptxA*(1) isolates, although they could also harbour a novel *ptxA* type, or that there are no *B. pertussis* cells in the clinical specimen to detect. Therefore, the suggested amendments of these two methods of nested PCR or real-time PCR with hybridisation probes, could potentially identify *prnA* and *ptxA* gene variation from previously non-culturable *B. pertussis.*

Bordetella pertussis is a slow-growing, fastidious organism, and isolation rates are as low as 20-40% (Heininger, 2001). Recent sensitive techniques such as PCR and serology are only detecting the presence of the organism or past infection, respectively. More information can be gained by combining PCR and subsequent sequencing and/or differential probe hybridisation, and these direct typing approaches should be actively pursued to broaden the knowledge of *B. pertussis* population dynamics. This kind of direct typing has also been successful using MLST in *N. meningitidis* (Diggle *et al.*, 2003), and additional targets could be added to *prnA* and *ptxA* to form a similar direct typing scheme. However, the need to attempt *B. pertussis* isolation from these clinical specimens should still be encouraged, as serotyping and PFGE have become established epidemiological typing techniques, and these rely on isolates. Furthermore, the lack of isolation could lead to the loss of information, despite future typing methods promising detailed data using unculturable organisms.

9 CONCLUSIONS

There has been a resurgence of pertussis in several countries despite high vaccination coverage. The use of various regimens of different pertussis vaccines available (WCV and ACV) has made it difficult to compare these countries, and each one should be investigated individually, as pertussis resurgence may be due to a combination of factors. The resurgence of pertussis has, in part, coincided with the introduction of more sensitive diagnostic techniques, such as serology and PCR detection. Thus, pertussis cases are now ascertained in babies, adolescents and adults in larger numbers than before, so that many countries including the USA and Australia (Guris *et al.*, 1999b; Andrews *et al.*, 1997) are finding an increase in pertussis notification rates. In the UK, *B. pertussis* is detected routinely in our national laboratories only by culture isolation. However, since April 2002, PCR (for infants under six months of age with suspected pertussis) and ELISA serology (for adults with a cough persisting for over 21 days and children with a cough persisting for over 14 days), have been available through RSIL, CPHL, London (Anon, 2002). Therefore, the true extent of pertussis incidence in the UK has yet to be established, although at the time of writing, 35 laboratory confirmations of pertussis infection were seen between January 2003 and March 2003, compared to 74 within the same quarter period in 2002 (Health Protection Agency, 2003a and 2003b). This reduction in laboratory-confirmed pertussis cases could be due to the effect of the ACV booster, introduced in November 2001, or as part of the three-yearly cycle seen in the notification of pertussis cases.

Although there is no apparent resurgence in the UK, studies from The Netherlands, Finland and Italy have stimulated the investigation of the genotypic diversity of the *B. pertussis* population in the UK (Mastrantonio *et al.*, 1999; Mooi *et al.*, 1998 and 1999). Various well-established typing methods were used for this, and to directly compare with other countries. Results indicated that, between 1940 and 1999, *B. pertussis* changed in its diversity of *prnA*, *ptxA* and IS*1002*-RFLP types. The *prnA* and *ptxA* data showed a shift from predominantly *prnA*(1) and *ptxA*(1) or *ptxA*(2) isolates, to mostly *prnA*(2) and *ptxA*(1) isolates. This shift was probably not due to vaccine-driven evolution, as suggested for The Netherlands (Mooi *et al.*, 1998), because the earlier UK WCV, manufactured by Wellcome, contains *ptxA*(1) strains, and also there has not been a resurgence of pertussis in the UK. The IS*1002*-RFLP data also support this view, as the UK Wellcome WCV strains are RFLP-1, RFLP-2 and RFLP-3, and match the current predominant RFLP types in UK *B. pertussis* isolates. The low number of IS*1002*-RFLP types seen in the 1990s, is also reflected in the low genotypic diversity value. A resurgence of pertussis would likely be seen as either a rise in genotypic diversity, due to an increase of different *B. pertussis* strains, or a drop in genotypic diversity, as successful lineages of *B. pertussis* expand. Therefore,

studying the genotypic diversity alone does not give a realistic depiction of *B. pertussis* population dynamics, and information on pertussis incidence, how pertussis cases are ascertained, and the schedule and type of pertussis vaccine used, can greatly enhance the data interpreted.

Combining laboratory, clinical and epidemiological data provides a rich data-set, and is invaluable to the epidemiologist in tracking serious disease traits in outbreaks and epidemics. Work in this thesis has emphasised that the under two month age group is most at risk of pertussis infection, and that cases admitted to hospital are more likely to be infected with a serotype 1,2/*prnA*(1) strain. The detection of these significant associations has been unexpected, as they are based on just 90 isolates from a six month period in the UK. With the introduction of the ACV pre-school booster in the UK, further analyses could reveal whether these risks have decreased, and if any other associations are evident. This could prove difficult however, as the UK is experiencing a historically low reported incidence of pertussis. Nevertheless, this type of study would be of benefit in identifying new severe markers or predicting the outcome of pertussis disease.

The characterisation of a small selection of isolates from Missouri, USA has strengthened many of the conclusions reached on the UK data. These include the suggestion that RFLP-1 and RFLP-2 are successful lineages of *B. pertussis*, and that the diversity of *B. pertussis* has dropped from the 1980s to the 1990s. Pertussis resurgence has occurred in Missouri, USA, but this study indicated no apparent reason for this. Although there has been a shift in isolates towards *prnA*(2)/*ptxA*(1) types, no conclusion can also be made regarding shifts to non-vaccine types, as the earliest isolate tested is from 1984, and vaccination was introduced in the USA in the 1950s. This small sub-set of *B. pertussis* isolates also generated a novel pertactin gene type, *prnA*(9), found in two of the 43 isolates obtained, and subsequently found in Canada (Peppler *et al.*, 2003). Whether this *prnA* type, like *prnA*(6), has geographical limitations, has yet to be seen. The Canadian study also found *prnA*(3) in circulation, and this type was originally thought to occur only in Europe. The ten isolates from Kansas City were also studied to elucidate whether they could be linked by epidemiological typing methods. These isolates did indeed share the same serotype, *prnA* and *ptxA* type, and IS*1002*-RFLP type, but were divided into two groups when investigated by PFGE. This highlights the difficulty in deciding on the level of strain discrimination, which is deemed necessary for epidemiological investigations.

Thus, when comparing epidemiological studies of *B. pertussis* from other countries, the use of the same typing method is ideal. In addition, the ability to standardise a typing scheme between laboratories is essential, and standardised methods for serotyping, pertactin and pertussis toxin S1 gene typing and PFGE for *B. pertussis* have been recommended (Mooi *et al.*, 2000). Pertactin and pertussis toxin S1 gene typing have become ubiquitous methods, as the results are easily portable, and direct comparisions can be performed. Various studies generating *prnA* and *ptxA* data from countries both with and without pertussis resurgence have been exhaustively compared. Overall, there has been the expansion of new *prnA* types such as *prnA*(2) and *prnA*(3), and the reduction of *ptxA* types to predominantly *ptxA*(1). Whether this is due to vaccine-driven evolution is unclear, but this has not halted the acceptance of this hypothesis. In order to clarify this hypothesis, the *prnA* and *ptxA* diversity is required in areas or countries where there is a very low or negligible pertussis vaccine coverage, so that a set of "control" data can be compared with those countries with high vaccination coverage.

Although pertactin and pertussis toxin S1 gene typing have become popular, other methods used to type *B. pertussis* include serotyping and PFGE for epidemiological investigations (Beall *et al.*, 1995; Preston, 1988) and IS*1002*-RFLP analysis for population genetic studies (van Loo *et al.*, 1999). These methods, except for pertactin and pertussis toxin S1 gene typing, were evaluated in this thesis, according to the guidelines of Struelens *et al.* (1996). This study revealed that, although PFGE had the greatest index of discrimination, a combination of serotyping and *prnA* and *ptxA* gene typing could equal this discrimination. In addition, these latter methods are more available, easier to perfom and the interpretation of results is less subjective. However, the recommendation of a method could not be given, as the pertactin and pertussis toxin S1 gene typing data was not generated from the blinded panel of isolates, and further work is required to optimise the PFGE technique.

Other approaches to type *B. pertussis* were also attempted, namely single-enzyme AFLP and adenylate kinase and filamentous haemagglutinin gene typing. Unfortunately, the AFLP technique did not generate sufficiently good results to warrant further investigation, and would require the optimisation of each primer for amplification. Amplified fragment length polymorphism is also dependent on the availability of a large amount of DNA from the *B. pertussis* isolate, and the subjectivity of gel profile analysis. The variations found in both *adk* and *fhaB* does, however, merit further investigation, and these two targets could be included in a DNA sequence-based scheme with *prnA* and *ptxA*. This type of scheme,

focusing on virulence factors, was first devised for *B. pertussis* by van Loo *et al.* (2002). The additional gene targets, the S3 subunit of pertussis toxin, the fimbrial gene for serotype 2, and the tracheal colonisation factor, have also been evaluated by Packard *et al.* (2004), using UK isolates collected for this thesis. When coded into the multi-locus sequence types (MLST's), all but one of UK clinical isolates from 2000 to 2002 were identified as MLST-5, which differed from the three Wellcome WCV strains (MLST-2 or MLST-3), and the ACV strain Tohama-I (MLST-2).

All of the typing methods described above rely on the isolation of the organism *B. pertussis*. Therefore, any past pertussis infection detected only by PCR and serology, would not be subjected to further characterisation, as the organism will not be available. I therefore investigated the possibility of typing *prnA* and *ptxA* directly from a clinical extract. The results were promising but further work is required to confidently obtain the *prnA* and *ptxA* types from cases where no isolate was obtained and is thus untypeable; e.g., from babies or adults presenting atypical symptoms. This sub-group could be harbouring different genotypes to those found in culture-positive cases.

Thanks to Dr J Parkhill, the annotated *B. pertussis* Tohama-I genome was available for this thesis, and was invaluable for comparing the observed and expected results of both IS*1002*-RFLP analysis and PFGE with *Xba*I. The genome could also have been used for predicting single-enzyme AFLP profiles (Chapter 7) and to ascertain other potential genes targets for epidemiological typing (Chapter 6). Indeed, several workers have used the genome to screen for virulence factors and other putatitive BvgAS-regulated genes (Antoine *et al.*, 2000; Passerini de Rossi *et al.*, 2003). The availability and publication of the annotated *B. pertussis* genome will therefore become an important tool for diagnostic and pathogenic research (Parkhill *et al.*, 2003).

10 FUTURE WORK

10.1 Pertactin and pertussis toxin gene sequencing of UK isolates

The analysis of *prnA* and *ptxA* sequences of current and future UK isolates will verify whether *prnA*(2) and *prnA*(3) types are increasing in the circulating population. Such strains could cause epidemics, as in The Netherlands (Mooi *et al.*, 1998). Also, the introduction of the ACV in November 1999, which is formulated from purified products of one strain (compared to three strains in the WCV), may have an effect on *prnA* type distribution. By combining this data with the epidemiological data available from CDSC, a larger data-set can be generated to measure any significant severe markers of pertussis disease, using multiple logistic regression analysis.

10.2 Pertactin and pertussis toxin gene sequencing in other countries

There have been many studies looking at the temporal distribution of *prnA* and *ptxA* types in vaccinated countries. Several studies have reported a shift in types and it has been hypothesised that the *B. pertussis* circulating population is undergoing antigenic shift due to vaccine pressure. In order to evaluate this hypothesis without bias, a country with no vaccination programme, or little vaccine coverage (e.g., Sweden, or a developing country) could be tested for their *prnA* and *ptxA* variants.

10.3 Pulsed-Field Gel Electrophoresis

Epidemiological typing of *B. pertussis* has been studied using PFGE in a number of countries. This method, after evalution in Chapter 5, was found to be sub-optimal. Consequently, further work is required to improve this established method and to re-examine the reproducibility, as PFGE can produce good discrimination, and could possibly be invaluable in an outbreak setting.

10.4 Amplified Fragment Length Polymorphism typing

Following a small-scale study investigating the potential of AFLP in typing *B. pertussis*, this could be adapted into a fluorescent-AFLP protocol. The use of *Pst*I in a single digest or *Eco*RI and *Mse*I in a double digest, could be evaluated with the well-characterised strains described in Chapter 7. With the availability of the *B. pertussis* Tohama-I genome, AFLP profiles could be predicted for other enzyme and primer combinations.

10.5 Epidemiological typing of *Bordetella pertussis* direct from clinical samples

Pertactin and pertussis toxin S1 subunit gene typing may be optimised using either nested PCR or real-time PCR techniques, as already mentioned in 8.4, page 266. An attempt to amplify other genes such as pertussis toxin S3 subunit, tracheal colonisation factor, etc, directly from clinical samples could be advantageous, as these have been shown to exhibit polymorphisms (Packard *et al.*, 2004; van Loo *et al.*, 2002). Therefore, these clinical samples could be used to determine the type characteristics pertactin, pertussis toxin, and tracheal colonisation factor of the *B. pertussis* organism present. Real-time PCR technology could also be used on clinical samples to target pertactin and other genes, and evaluate whether the resulting products could be sequenced.

11 REFERENCES

Abbott JD, Gillespie EH, Higgins PG, Polakoff S, Pollock TM. (1973). Efficacy of whooping-cough vaccines used in the United Kingdom before 1968. *Br Med J* **1,** 259-262.

Akerley BJ, Monack DM, Falkow S, Miller JF. (1992). The *bvgAS* locus negatively controls motility and synthesis of flagella in *Bordetella bronchiseptica*. *J Bacteriol* **174,** 980-990.

Andrews R, Herceg A, Roberts C. (1997). Pertussis notifications in Australia, 1991 to 1997. *Commun Dis Intell* **21,** 145-148.

Anon. (1970). Diagnosis of whooping cough: comparison of serological tests with isolation of *Bordetella pertussis*. A combined Scottish study. *Br Med J* **4,** 637-639.

Anon. (1999a). Pertussis vaccines- WHO position paper. *Wkly Epid Rec- WHO* **74,** 137-143.

Anon. (1999b). Special issue on pertussis. *Eurosurveillance* **4,** 127-134.

Anon. (2002). Whooping cough: enhanced laboratory surveillance of pertussis, first quarter and statutory notifications to week 26: 2002. *CDR Weekly* **12,** 12-14.

Anon. (2003a). German vaccination recommendations. Taken from http://www.rki.de/GESUND/IMPFEN/IMPFEN.HTM?/GESUND/IMPFEN/STI_NEU.HTM&1, last accessed 22nd March 2004.

Anon. (2003b). Canadian vaccination recommendations. Taken from http://www.hc-sc.gc.ca/pphb-dgspsp/publicat/cig-gci/pdf/cdn_immuniz_guide-2002-6.pdf, last accessed 22nd March 2004.

Antoine R, Alonso S, Raze D, Coutte L, Lesjean S, Willery E, Locht C, Jacob-Dubuisson F. (2000). New virulence-activated and virulence-repressed genes identified by systematic gene inactivation and generation of transcriptional fusions in *Bordetella pertussis*. *J Bacteriol* **182,** 5902-5905.

Arbeit RD, Arthur M, Dunn R, Kim C, Selander RK, Goldstein R. (1990). Resolution of recent evolutionary divergence among *Escherichia coli* from related lineages: the application of pulsed field electrophoresis to molecular epidemiology. *J Infect Dis* **161,** 230-235.

Arbeit RD, Slutsky A, Barber TW, Maslow JN, Niemczyk S, Falkinham JO, O'Connor GT, von Reyn CF. (1993). Genetic diversity among strains of *Mycobacterium avium* causing monoclonal and polyclonal bacteremia in patients with AIDS. *J Infect Dis* **167,** 1384-1390.

Arico B, Scarlato V, Monack DM, Falkow S, Rappuoli R. (1991). Structural and genetic analysis of the *bvg* locus in *Bordetella species*. *Mol Microbiol* **5,** 2481-2491.

Arico B, Nuti S, Scarlato V, Rappuoli R. (1993). Adhesion of *Bordetella pertussis* to eukaryotic cells requires a time-dependent export and maturation of filamentous hemagglutinin. *Proc Natl Acad Sci U S A* **90,** 9204-9208.

Arnold C, Metherell L, Willshaw G, Maggs A, Stanley J. (1999). Predictive fluorescent amplified-fragment length polymorphism analysis of *Escherichia coli*: high-resolution typing method with phylogenetic significance. *J Clin Microbiol* **37,** 1274-1279.

Ashworth LA, Robinson A, Funnell S, Gorringe AR, Irons LI, Seabrook RN. (1988). Agglutinogens and fimbriae of *Bordetella pertussis*. *Tokai J Exp Clin Med* **13 Suppl,** 203-210.

Beall B, Cassiday PK, Sanden GN. (1995). Analysis of *Bordetella pertussis* isolates from an epidemic by pulsed-field gel electrophoresis. *J Clin Microbiol* **33,** 3083-3086.

Beard SM, Finn A. (2000). Do we need to boost pertussis immunization within the existing UK vaccination schedule? *J Public Health Med* **22,** 349-356.

Berggard K, Johnsson E, Mooi FR, Lindahl G. (1997). *Bordetella pertussis* binds the human complement regulator C4BP: role of filamentous hemagglutinin. *Infect Immun* **65,** 3638-3643.

Birkeback NH. (2001). *Bordetella pertussis* in the aetiology of chronic cough in adults. Diagnostic methods and clinic. *Dan Med Bull* **48,** 77-80.

Bisgard KM, Christie CD, Reising SF, Sanden GN, Cassiday PK, Gomersall C, Wattigney WA, Roberts NE, Strebel PM. (2001). Molecular epidemiology of *Bordetella pertussis* by pulsed-field gel electrophoresis profile: Cincinnati, 1989-1996. *J Infect Dis* **183,** 1360-1367.

Blackburn P. (2000). Characterisation of the virulence-related, outer-membrane proteins of *Bordetella pertussis*. University of Glasgow.

Bokhari H. (2002). Characterisation and secretion mechanisms of *Bordetella pertussis* autotransporter proteins. University of Glasgow.

Bordet J, Gengou O. (1906). Le microbe de la coqueluche. *Ann Inst Pasteur* **20,** 731-741.

Boumedine KS, Rodolakis A. (1998). AFLP allows the identification of genomic markers of ruminant *Chlamydia psittaci* strains useful for typing and epidemiological studies. *Res Microbiol* **149,** 735-744.

Boursaux-Eude C, Thiberge S, Carletti G, Guiso N. (1999). Intranasal murine model of *Bordetella pertussis* infection: II. Sequence variation and protection induced by a tricomponent acellular vaccine. *Vaccine* **17,** 2651-2660.

Boursaux-Eude C, Guiso N. (2000). Polymorphism of repeated regions of pertactin in *Bordetella pertussis*, *Bordetella parapertussis*, and *Bordetella bronchiseptica*. *Infect Immun* **68,** 4815-4817.

Brennan MJ, Hannah JH, Leininger E. (1991). Adhesion of *Bordetella pertussis* to sulfatides and to the GalNAc beta 4Gal sequence found in glycosphingolipids. *J Biol Chem* **266,** 18827-18831.

Brennan M, Strebel P, George H, Yih WK, Tachdjian R, Lett SM, Cassiday P, Sanden G, Wharton M. (2000). Evidence for transmission of pertussis in schools, Massachusetts, 1996: epidemiologic data supported by pulsed-field gel electrophoresis studies. *J Infect Dis* **181,** 210-215.

Carter MW, Harrison TG, Shuja SM, George RC. (1998). Typing strains of *Chlamydia pneumoniae* by amplified fragment length polymorphism typing. *Clin Microbiol Infect* **4,** 663-664.

Cassiday P, Sanden G, Heuvelman K, Mooi F, Bisgard KM, Popovic T. (2000). Polymorphism in *Bordetella pertussis* pertactin and pertussis toxin virulence factors in the United States, 1935-1999. *J Infect Dis* **182,** 1402-1408.

Chaby R, Caroff M. (1988). Lipopolysaccharides of *Bordetella pertussis* Endotoxin. In: Wardlaw AC, Parton R (eds) *Pathogenesis and Immunity in Pertussis.* Chichester; John Wiley & Sons, pp 247-272.

Charles IG, Dougan G, Pickard D, Chatfield S, Smith M, Novotny P, Morrissey P, Fairweather NF. (1989). Molecular cloning and characterization of protective outer membrane protein P.69 from *Bordetella pertussis*. *Proc Natl Acad Sci U S A* **86,** 3554-3558.

Charles IG, Li JL, Roberts M, Beesley K, Romanos M, Pickard DJ, Francis M, Campbell D, Dougan G, Brennan MJ, Manclark CR, Jensen MA, Heron I, Chubb A, Novotny P, Fairweather NF. (1991). Identification and characterization of a protective immunodominant B cell epitope of pertactin (P.69) from *Bordetella pertussis*. *Eur J Immunol* **21,** 1147-1153.

Charles I, Fairweather N, Pickard D, Beesley J, Anderson R, Dougan G, Roberts M. (1994). Expression of the *Bordetella pertussis* P.69 pertactin adhesin in *Escherichia coli*: fate of the carboxy-terminal domain. *Microbiology* **140,** 3301-3308.

Chen I, Finn TM, Yanqing L, Guoming Q, Rappuoli R, Pizza M. (1998). A recombinant live attenuated strain of *Vibrio cholerae* induces immunity against tetanus toxin and *Bordetella pertussis* tracheal colonization factor. *Infect Immun* **66,** 1648-1653.

Cherry JD. (1996). Historical review of pertussis and the classical vaccine. *J Infect Dis* **174 Suppl 3,** S259-S263.

Chhatwal GS, Walker MJ, Yan H, Timmis KN, Guzman CA. (1997). Temperature dependent expression of an acid phosphatase by *Bordetella bronchiseptica*: role in intracellular survival. *Microb Pathog* **22,** 257-264.

Christie CD, Garrison KM, Kiely L, Gupta RK, Heubi J, Marchant CD. (2001). A trial of acellular pertussis vaccine in hospital workers during the Cincinnati pertussis epidemic of 1993. *Clin Infect Dis* **33,** 997-1003.

Clabots CR, Johnson S, Olson MM, Peterson LR, Gerding DN. (1992). Acquisition of *Clostridium difficile* by hospitalized patients: evidence for colonized new admissions as a source of infection. *J Infect Dis* **166,** 561-567.

Cleary PP, Kaplan EL, Livdahl C, Skjold S. (1988). DNA fingerprints of *Streptococcus pyogenes* are M type specific. *J Infect Dis* **158,** 1317-1323.

Cookson BT, Vandamme P, Carlson LC, Larson AM, Sheffield JV, Kersters K, Spach DH. (1994). Bacteremia caused by a novel *Bordetella* species, "*B. hinzii*". *J Clin Microbiol* **32,** 2569-2571.

CPHL, PHLS. (2000a). Opening of Freeze-dried ampoule cultures- COSHH assessment number M34. Central Public Health Laboratory, Public Health Laboratory Service.

CPHL, PHLS. (2000b). Preservation of *Legionella* and other bacteria on beads- SOP R-6211/01-00. Central Public Health Laboratory, Public Health Laboratory Service.

Crowcroft NS, Andrews N, Rooney C, Brisson M, Miller E. (2002). Deaths from pertussis are underestimated in England. *Arch Dis Child* **86,** 336-338.

Crowcroft NS, Britto J. (2002). Whooping cough-a continuing problem. *BMJ* **324,** 1537-1538.

Crowcroft NS, Booy R, Harrison T, Spicer L, Britto J, Mok Q, Heath P, Murdoch I, Zambon M, George R, Miller E. (2003). Severe and unrecognised: pertussis in UK infants. *Arch Dis Child* **88,** 802-806.

Cullinane LC, Alley MR, Marshall RB, Manktelow BW. (1987). *Bordetella parapertussis* from lambs. *N Z Vet J* **35,** 175.

de Ley J, Spegers P, Kersters K, Mannheim W, Lievens A. (1986). Intra- and intergenic similarities of the *Bordetella* Ribosomal Ribonucleic acid cistrons: proposal for a new family, *Alcaligenaceae*. *Int J Syst Bacteriol* **36,** 405-414.

de Melker HE, Conyn-van Spaendonck MA, Rumke HC, van Wijngaarden JK, Mooi FR, Schellekens JF. (1997). Pertussis in The Netherlands: an outbreak despite high levels of immunization with whole-cell vaccine. *Emerg Infect Dis* **3,** 175-178.

de Melker HE, Schellekens JF, Neppelenbroek SE, Mooi FR, Rumke HC, Conyn-van Spaendonck MA. (2000a). Reemergence of pertussis in the highly vaccinated population of the Netherlands: observations on surveillance data. *Emerg Infect Dis* **6,** 348-357.

de Melker HE, Versteegh FG, Conyn-van Spaendonck MA *et al.* (2000b). Specificity and sensitivity of high levels of immunoglobulin G antibodies against pertussis toxin in a single serum sample for diagnosis of infection with *Bordetella pertussis*. *J Clin Microbiol* **38,** 800-806.

de Moissac YR, Ronald SL, Peppler MS. (1994). Use of pulsed-field gel electrophoresis for epidemiological study of *Bordetella pertussis* in a whooping cough outbreak. *J Clin Microbiol* **32,** 398-402.

de Schutter I, Malfroot A, Dab I, Hoebrekx N, Muyldermans G, Pierard D, Lauwers S. (2003). Molecular typing of *Bordetella pertussis* isolates recovered from Belgian children and their household members. *Clin Infect Dis* **36,** 1391-1396.

De Serres G, Boulianne N, Douville Fradet M., Duval B. (1995). Pertussis in Quebec: ongoing epidemic since the late 1980s. *Can Commun Dis Rep* **21,** 45-48.

Delisse-Gathoye AM, Locht C, Jacob F, Raaschou-Nielsen M, Heron I, Ruelle JL, de Wilde M, Cabezon T. (1990). Cloning, partial sequence, expression, and antigenic analysis of the filamentous hemagglutinin gene of *Bordetella pertussis. Infect Immun* **58,** 2895-2905.

Department of Health and Social Security. (1977). Whooping Cough Vaccination- Review of the Evidence on Whooping Cough Vaccination by the Joint Committee on Vaccination and Immunization. Department of Health and Social Security. London, Her Majesty's Stationary Office.

Diggle MA, Bell CM, Clarke SC. (2003). Nucleotide sequence-based typing of meningococci directly from clinical samples. *J Med Microbiol* **52,** 505-508.

Dodhia H, Crowcroft NS, Bramley JC, Miller E. (2002). UK guidlines for use of erythromycin chemoprophylaxis in persons exposed to pertussis. *J Public Health Med* **24,** 200-206.

Edmunds WJ, Brisson M, Melegaro A, Gay NJ. (2002). The potential cost-effectiveness of acellular pertussis booster vaccination in England and Wales. *Vaccine* **20,** 1316-1330.

Edwards KM, Decker MD, Mortimer EA. (1999). Pertussis vaccine. In: Plotkin SA, Orenstein WA (eds) *Vaccines.* Philadelphia; W.B. Saunders Company, pp 293-344.

Edwards KM. (2001). Is pertussis a frequent cause of cough in adolescents and adults? Should routine pertussis immunization be recommended? *Clin Infect Dis* **32,** 1698-1699.

el Baya A, Linnemann R, von Olleschik-Elbheim L, Robenek H, Schmidt MA. (1997). Endocytosis and retrograde transport of pertussis toxin to the Golgi complex as a prerequisite for cellular intoxication. *Eur J Cell Biol* **73,** 40-48.

Eldering G, Kendrick P. (1938). *Bacillus parapertussis*: A species resembling both *Bacillus pertussis* and *Bacillus bronchisepticus* but identical with neither. *J Bacteriol* **35,** 561-572.

Eldering G, Hornbeck C, Baker J. (1957). Serological study of *Bordetella pertussis* and related species. *J Bacteriol* **74,** 133-136.

Elnifro EM, Ashshi AM, Cooper RJ, Klapper PE. (2000). Multiplex PCR: optimization and application in diagnostic virology. *Clin Microbiol Rev* **13,** 559-570.

Emsley P, Charles IG, Fairweather NF, Isaacs NW. (1996). Structure of *Bordetella pertussis* virulence factor P.69 pertactin. *Nature* **381,** 90-92.

Enright MC, Spratt BG. (1998). A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. *Microbiology* **144,** 3049-3060.

Enright MC, Day NP, Davies CE, Peacock SJ, Spratt BG. (2000). Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus. J Clin Microbiol* **38,** 1008-1015.

Everest P, Li J, Douce G, Charles I, De Azavedo J, Chatfield S, Dougan G, Roberts M. (1996). Role of the *Bordetella pertussis* P.69/pertactin protein and the P.69/pertactin RGD motif in the adherence to and invasion of mammalian cells. *Microbiology* **142,** 3261-3268.

Farfel Z, Konen S, Wiertz E, Klapmuts R, Addy PA, Hanski E. (1990). Antibodies to *Bordetella pertussis* adenylate cyclase are produced in man during pertussis infection and after vaccination. *J Med Microbiol* **32,** 173-177.

Farizo KM, Fiddner S, Cheung AM, Burns DL. (2002). Membrane localization of the S1 subunit of pertussis toxin in *Bordetella pertussis* and implications for pertussis toxin secretion. *Infect Immun* **70,** 1193-1201.

Feavers IM, Gray SJ, Urwin R, Russell JE, Bygraves JA, Kaczmarski EB, Maiden MC. (1999). Multilocus sequence typing and antigen gene sequencing in the investigation of a meningococcal disease outbreak. *J Clin Microbiol* **37,** 3883-3887.

Fernandez RC, Weiss AA. (1994). Cloning and sequencing of a *Bordetella pertussis* serum resistance locus. *Infect Immun* **62,** 4727-4738.

Fernandez RC, Weiss AA. (1998). Serum resistance in *bvg*-regulated mutants of *Bordetella pertussis. FEMS Microbiol Lett* **163,** 57-63.

Ferry NS. (1912). *Bacillus bronchisepticus* (*bronchicanis*): The cause of distemper in dogs and a similar disease in other animals. *Vet J* **68,** 376-391.

Finn TM, Stevens LA. (1995). Tracheal colonization factor: a *Bordetella pertussis* secreted virulence determinant. *Mol Microbiol* **16,** 625-634.

Finn TM, Amsbaugh DF. (1998). Vag8, a *Bordetella pertussis bvg*-regulated protein. *Infect Immun* **66,** 3985-3989.

Flak TA, Goldman WE. (1999). Signalling and cellular specificity of airway nitric oxide production in pertussis. *Cell Microbiol* **1,** 51-60.

Fletcher MA, Saliou P, Ethevenaux C, Plotkin SA. (2001). The efficacy of whole cell pertussis immunisation: collected data on a vaccine produced in France. *Public Health* **115,** 119-129.

Forde CB, Parton R, Coote JG. (1998). Bioluminescence as a reporter of intracellular survival of *Bordetella bronchiseptica* in murine phagocytes. *Infect Immun* **66,** 3198-3207.

Foreman-Wykert AK, Deora R, Mattoo S, Miller JF. (2002). An autotransporter is required for secretion of type III effector proteins in *Bordetella bronchiseptica.* [7], American Society for Microbiology. Seventh International Symposium on Pertussis: Genome, Pathogenesis and Immunity. 18-9-2002.

Foxman B, Riley L. (2001). Molecular epidemiology: focus on infection. *Am J Epidemiol* **153,** 1135-1141.

Friedman RL, Nordensson K, Wilson L, Akporiaye ET, Yocum DE. (1992). Uptake and intracellular survival of *Bordetella pertussis* in human macrophages. *Infect Immun* **60,** 4578-4585.

Fry NK, Alexiou-Daniel S, Bangsborg JM, Bernander S, Pastoris MC, Etienne J, Forsblom B, Gaia V, Helbig JH, Lindsay D, Luck PC, Pelaz C, Uldum SA, Harrison TG. (1999). A multicenter evaluation of genotypic methods for the epidemiologic typing of *Legionella pneumophila* serogroup 1: results of a pan-European study. *Clin Microbiol Infect* **5,** 462-477.

Fry NK, Bangsborg JM, Bernander S, Etienne J, Forsblom B, Gaia V, Hasenberger P, Lindsay D, Papoutsi A, Pelaz C, Struelens M, Uldum SA, Visca P, Harrison TG. (2000). Assessment of intercentre reproducibility and epidemiological concordance of *Legionella pneumophila* serogroup 1 genotyping by amplified fragment length polymorphism analysis. *Eur J Clin Microbiol Infect Dis* **19,** 773-780.

Fry NK, Tzivra O, Li YTL, McNiff A, Doshi N, Maple PAC, Crowcroft NS, Miller E, George R, Harrison TG. (2004). Laboratory diagnosis of pertussis infections: the role of PCR and serology. *J Med Microbiol* **53**, 519-525.

Funke G, Hess T, von Graevenitz A, Vandamme P. (1996). Characteristics of *Bordetella hinzii* strains isolated from a cystic fibrosis patient over a 3-year period. *J Clin Microbiol* **34**, 966-969.

Gaia V, Fry NK, Harrison T, Peduzzi R. (2003). Sequence-based typing of *Legionella pneumophila* serogroup 1 offers the potential for true portability in legionellosis outbreak investigation. *J Clin Microbiol* **41**, 2932-2939.

Gangarosa EJ, Galazka AM, Wolfe CR, Phillips LM, Gangarosa RE, Miller E, Chen RT. (1998). Impact of anti-vaccine movements on pertussis control: the untold story. *Lancet* **351**, 356-361.

Gargiulo F, De Francesco MA, Pinsi G, Pollara C, Terlenghi L, Perandin F, Manca N. (2003). Determination of HCV genotype by direct sequence analysis of quantitative PCR products. *J Med Virol* **69**, 202-206.

Gentry-Weeks CR, Cookson BT, Goldman WE, Rimler RB, Porter SB, Curtiss R. (1988). Dermonecrotic toxin and tracheal cytotoxin, putative virulence factors of *Bordetella avium*. *Infect Immun* **56**, 1698-1707.

Gentry-Weeks CR, Provence DL, Keith JM, Curtiss R. (1991). Isolation and characterization of *Bordetella avium* phase variants. *Infect Immun* **59**, 4026-4033.

Gerlach G, von Wintzingerode F, Middendorf B, Gross R. (2001). Evolutionary trends in the genus *Bordetella*. *Microbes Infect* **3**, 61-72.

Geuijen CA, Willems RJ, Hoogerhout P, Puijk WC, Meloen RH, Mooi FR. (1998). Identification and characterization of heparin binding regions of the Fim2 subunit of *Bordetella pertussis*. *Infect Immun* **66**, 2256-2263.

Giardina PC, Foster LA, Musser JM, Akerley BJ, Miller JF, Dyer DW. (1995). *bvg* repression of alcaligin synthesis in *Bordetella bronchiseptica* is associated with phylogenetic lineage. *J Bacteriol* **177**, 6058-6063.

Gilles AM, Sismeiro O, Munier H, Fabian H, Mantsch HH, Surewicz WK, Craescu CC, Barzu O, Danchin A. (1993). Structural and physico-chemical characteristics of *Bordetella pertussis* adenylate kinase, a tryptophan-containing enzyme. *Eur J Biochem* **218**, 921-927.

Gilson E, Clement JM, Brutlag D, Hofnung M. (1984). A family of dispersed repetitive extragenic palindromic DNA sequences in *E. coli*. *EMBO J* **3**, 1417-1421.

Gladbach S, Hanauer S, Reischl U, Wilson K, Sanden G. (2002). Identification of IS*481* in *Bordetella bronchiseptica*: Implications for *Bordetella* spp phylogeny and diagnosis of *Bordetella pertussis* infections by polymerase chain reaction assays. *7th International Symposium on Pertussis: Genome, Pathogenesis, and Immunity, Hinxton, U.K.*

Glare EM, Paton JC, Premier RR, Lawrence AJ, Nisbet IT. (1990). Analysis of a repetitive DNA sequence from *Bordetella pertussis* and its application to the diagnosis of pertussis using the polymerase chain reaction. *J Clin Microbiol* **28**, 1982-1987.

Goldman WE, Klapper DG, Baseman JB. (1982). Detection, isolation, and analysis of a released *Bordetella pertussis* product toxic to cultured tracheal cells. *Infect Immun* **36**, 782-794.

Goldman WE. (1988). Tracheal Cytotoxin of *Bordetella pertussis*. In: Wardlaw AC, Parton R (eds) *Pathogenesis and Immunity in Pertussis*. Chichester; John Wiley & Sons, pp 231-246.

Grant CC, Cherry JD. (2002). Keeping Pace with the Elusive *Bordetella pertussis*. *J Infect* **44**, 7-12.

Greig JR, Gunda SS, Kwan JTC. (2001). *Bordetella holmesii* bacteraemia in an individual on haemodialysis. *Scand J Infect Dis* **33,** 716-717.

Guiso N, Szatanik M, Rocancourt M. (1991). Protective activity of *Bordetella* adenylate cyclase-hemolysin against bacterial colonization. *Microb Pathog* **11,** 423-431.

Guiso N, Boursaux-Eude C, Weber C, Hausman SZ, Sato H, Iwaki M, Kamachi K, Konda T, Burns DL. (2001a). Analysis of *Bordetella pertussis* isolates collected in Japan before and after introduction of acellular pertussis vaccines. *Vaccine* **19,** 3248-3252.

Guiso N, von Konig CH, Becker C, Hallander H. (2001b). Fimbrial typing of *Bordetella pertussis* isolates: agglutination with polyclonal and monoclonal antisera. *J Clin Microbiol* **39,** 1684-1685.

Guris D, Martin R, Wharton M. (1999a). Pertussis. In: Wharton M, Roush S (eds) *Manual for the Surveillance of Vaccine-Preventable Diseases.* Centres of Disease Control and Prevention and the Council of State and Territorial Epidemiologists, pp 8-1-8.

Guris D, Strebel PM, Bardenheier B, Brennan M, Tachdjian R, Finch E, Wharton M, Livengood JR. (1999b). Changing epidemiology of pertussis in the United States: increasing reported incidence among adolescents and adults, 1990-1996. *Clin Infect Dis* **28,** 1230-1237.

Gurtler V, Mayall BC. (2001). Genomic approaches to typing, taxonomy and evolution of bacterial isolates. *Int J Syst Evol Microbiol* **51,** 3-16.

Gzyl A, Augustynowicz E, van Loo I, Slusarczyk J. (2002). Temporal nucleotide changes in pertactin and pertussis toxin genes in *Bordetella pertussis* strains isolated from clinical cases in Poland. *Vaccine* **20,** 299-303.

Hardwick TH, Plikaytis B, Cassiday PK, Cage G, Peppler MS, Shea D, Boxrud D, Sanden GN. (2002). Reproducibility of *Bordetella pertussis* genomic DNA fragments generated by *Xba*I restriction and resolved by pulsed-field gel electrophoresis. *J Clin Microbiol* **40,** 811-816.

Hazenbos WL, van den Berg BM, van Furth R. (1993). Very late antigen-5 and complement receptor type 3 cooperatively mediate the interaction between *Bordetella pertussis* and human monocytes. *J Immunol* **151,** 6274-6282.

He Q, Makinen J, Berbers G, Mooi FR, Viljanen MK, Arvilommi H, Mertsola J. (2003). *Bordetella pertussis* protein pertactin induces type-specific antibodies: one possible explanation for the emergence of antigenic variants? *J Infect Dis* **187,** 1200-1205.

Health Protection Agency. (2003a). Laboratory Confirmed Cases of Pertussis Infection England and Wales: January to December 2002. Taken from http://www.hpa.org.uk/cdr/archive03/immunisation03.htm#whooping, last accessed 22nd March 2004.

Health Protection Agency. (2003b). Laboratory confirmed cases of pertussis infection in England and Wales by age group January to March 2003. Taken from http://www.hpa.org.uk/cdr/archive03/immunisation03b.htm#pert, last accessed 22nd March 2004.

Health Protection Agency. (2003c). Pertussis General Information on the Health Protection Agency web-site. Taken from http://www.hpa.org.uk/infections/topics_az/whoopingcough/gen_info.htm, last accessed 22nd March 2004.

Heininger U. (2001). Pertussis: an old disease that is still with us. *Current Opinion in Infectious Diseases* **14,** 329-335.

Heiss LN, Lancaster JR, Corbett JA, Goldman WE. (1994). Epithelial autotoxicity of nitric oxide: role in the respiratory cytopathology of pertussis. *Proc Natl Acad Sci U S A* **91,** 267-270.

Henderson IR, Navarro-Garcia F, Nataro JP. (1998). The great escape: structure and function of the autotransporter proteins. *Trends Microbiol* **6,** 370-378.

Hewlett E, Gordon VM. (1988). Adenylate cyclase toxin of *Bordetella pertussis*. In: Wardlaw AC, Parton R (eds) *Pathogenesis and Immunity in Pertussis*. Chichester; John Wiley & Sons, pp 193-209.

Hewlett EL. (2000). *Bordetella* species. In: Mandell GL, Bennet JE, Dolin R (eds) *Mandell, Douglas and Bennett's principles and practice of infectious diseases*. Philadelphia; Churchill Livingstone, pp 2414-2422.

Holmes WH. Bacillary and Rickettsial Infections. New York, Macmillan. 1940: 395-398.

Hormozi K, Parton R, Coote J. (1999). Adjuvant and protective properties of native and recombinant *Bordetella pertussis* adenylate cyclase toxin preparations in mice. *FEMS Immunol Med Microbiol* **23,** 273-282.

Hot D, Antoine R, Renauld-Mongenie G, Caro V, Hennuy B, Levillain E, Huot L, Wittmann G, Poncet D, Jacob-Dubuisson F, Guyard C, Rimlinger F, Aujame L, Godfroid E, Guiso N, Quentin-Millet MJ, Lemoine Y, Locht C. (2003). Differential modulation of *Bordetella pertussis* virulence genes as evidenced by DNA microarray analysis. *Mol Genet Genomics* **269,** 475-486.

Houard S, Hackel C, Herzog A, Bollen A. (1989). Specific identification of *Bordetella pertussis* by the polymerase chain reaction. *Res Microbiol* **140,** 477-487.

Hulton CS, Higgins CF, Sharp PM. (1991). ERIC sequences: a novel family of repetitive elements in the genomes of *Escherichia coli, Salmonella typhimurium* and other enterobacteria. *Mol Microbiol* **5,** 825-834.

Hunter PR, Gaston MA. (1988). Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J Clin Microbiol* **26,** 2465-2466.

Hunter PR. (1990). Reproducibility and indices of discriminatory power of microbial typing methods. *J Clin Microbiol* **28,** 1903-1905.

Immunisation Division, Communicable Disease Surveillance Centre, Health Protection Agency. (2004a). Pertussis incidence and vaccine coverage in the United Kingdom, from 1940 to 2003.

Immunisation Division, Communicable Disease Surveillance Centre, Health Protection Agency. (2004b). *Bordetella pertussis* serotype distribution in England and Wales, from 1995 to 2002.

Janssen P, Coopman R, Huys G, Swings J, Bleeker M, Vos P, Zabeau M, Kersters K. (1996). Evaluation of the DNA fingerprinting method AFLP as an new tool in bacterial taxonomy. *Microbiology* **142,** 1881-1893.

Jungnitz H, West NP, Walker MJ, Chhatwal GS, Guzman CA. (1998). A second two-component regulatory system of *Bordetella bronchiseptica* required for bacterial resistance to oxidative stress, production of acid phosphatase, and *in vivo* persistence. *Infect Immun* **66,** 4640-4650.

Kerr JR, Matthews RC. (2000). *Bordetella pertussis* infection: pathogenesis, diagnosis, management, and the role of protective immunity. *Eur J Clin Microbiol Infect Dis* **19,** 77-88.

Kersters K, Hinz KH, Hertle A, Segers P, Lievens A, Siegmann O, de Ley J. (1984). *Bordetella avium* sp. nov., isolated from the respiratory tracts of turkeys and other birds. *Int J Syst Bacteriol* **34,** 56-70.

Khattak MN, Matthews RC, Burnie JP. (1992). Is *Bordetella pertussis* clonal? *BMJ* **304,** 813-815.

King AJ, Berbers G, van Oirschot HF, Hoogerhout P, Knipping K, Mooi FR. (2001). Role of the polymorphic region 1 of the *Bordetella pertussis* protein pertactin in immunity. *Microbiology* **147,** 2885-2895.

Kinnear SM, Boucher PE, Stibitz S, Carbonetti NH. (1999). Analysis of BvgA activation of the pertactin gene promoter in *Bordetella pertussis*. *J Bacteriol* **181,** 5234-5241.

Knapp S, Mekalanos JJ. (1988). Two *trans*-acting regulatory genes (*vir* and *mod*) control antigenic modulation in *Bordetella pertussis*. *J Bacteriol* **170,** 5059-5066.

Kosters K, Riffelmann M, von Konig CH. (2001). Evaluation of a real-time PCR assay for detection of *Bordetella pertussis* and *B.parapertussis* in clinical samples. *J Med Microbiol* **50,** 436-440.

Kourova N, Caro V, Weber C, Thiberge S, Chuprinina R, Tseneva G, Guiso N. (2003). Comparison of the *Bordetella pertussis* and *Bordetella parapertussis* isolates circulating in Saint Petersburg between 1998 and 2000 with Russian vaccine strains. *J Clin Microbiol* **41,** 3706-3711.

Lacey BW. (1960). Antigenic modulation of *Bordetella pertussis*. *J Hygiene* **58,** 57-93.

Lee CK, Roberts AL, Finn TM, Knapp S, Mekalanos JJ. (1990). A new assay for invasion of HeLa 229 cells by *Bordetella pertussis*: effects of inhibitors, phenotypic modulation, and genetic alterations. *Infect Immun* **58,** 2516-2522.

Leslie PH, Gardner AD. (1931). The phases of *Haemophilus pertussis*. *J Hygiene* **31,** 423-434.

Linnemann CC, Ramundo N, Perlstein PH, Minton SD, Englender GS. (1975). Use of pertussis vaccine in an epidemic involving hospital staff. *Lancet* **2,** 540-543.

Locht C. (1999). Molecular aspects of *Bordetella pertussis* pathogenesis. *Int Microbiol* **2,** 137-144.

Locht C, Antoine R. (1999). *Bordetella pertussis* protein toxins. In: Alouf JE, Freer JH (eds) *The Comprehensive Sourcebook of Protein Toxins*. Academic Press, pp 130-146.

Locht C, Antoine R, Jacob-Dubuisson F. (2001). *Bordetella pertussis*, molecular pathogenesis under multiple aspects. *Curr Opin Microbiol* **4,** 82-89.

Louie M, Louie L, Simor AE. (2000). The role of DNA amplification technology in the diagnosis of infectious diseases. *CMAJ* **163,** 301-309.

Mahillon J, Chandler M. (1998). Insertion sequences. *Microbiol Mol Biol Rev* **62,** 725-774.

Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG. (1998). Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A* **95,** 3140-3145.

Makhov AM, Hannah JH, Brennan MJ, Trus BL, Kocsis E, Conway JF, Wingfield PT, Simon MN, Steven AC. (1994). Filamentous hemagglutinin of *Bordetella pertussis*. A bacterial adhesin formed as a 50-nm monomeric rigid rod based on a 19-residue repeat motif rich in beta strands and turns. *J Mol Biol* **241,** 110-124.

Makinen J, Viljanen MK, Mertsola J, Arvilommi H, He Q. (2001). Rapid Identification of *Bordetella pertussis* Pertactin Gene Variants Using LightCycler Real-Time Polymerase Chain Reaction Combined with Melting Curve Analysis and Gel Electrophoresis. *Emerg Infect Dis* **7,** 952-958.

Makinen J, Mertsola J, Viljanen MK, Arvilommi H, He Q. (2002). Rapid typing of *Bordetella pertussis* pertussis toxin gene variants by LightCycler real-time PCR and fluorescence resonance energy transfer hybridization probe melting curve analysis. *J Clin Microbiol* **40,** 2213-2216.

Martin B, Humbert O, Camara M, Guenzi E, Walker J, Mitchell T, Andrew P, Prudhomme M, Alloing G, Hakenbeck R. (1992). A highly conserved repeated DNA element located in the chromosome of *Streptococcus pneumoniae*. *Nucleic Acids Res* **20,** 3479-3483.

Maslow JN, Mulligan ME, Arbeit RD. (1993). Molecular epidemiology: application of contemporary techniques to the typing of microorganisms. *Clin Infect Dis* **17,** 153-162.

Maslow J, Mulligan ME. (1996). Epidemiologic typing systems. *Infect Control Hosp Epidemiol* **17,** 595-604.

Mastrantonio P, Spigaglia P, van Oirschot H, van der Heide HG, Heuvelman K, Stefanelli P, Mooi FR. (1999). Antigenic variants in *Bordetella pertussis* strains isolated from vaccinated and unvaccinated children. *Microbiology* **145,** 2069-2075.

Maynard Smith J, Smith NH, O'Rourke M, Spratt BG. (1993). How clonal are bacteria? *Proc Natl Acad Sci U S A* **90,** 4384-4388.

Mazengia E, Silva EA, Peppe JA, Timperi R, George H. (2000). Recovery of *Bordetella holmesii* from patients with pertussis-like symptoms: use of pulsed-field gel electrophoresis to characterize circulating strains. *J Clin Microbiol* **38,** 2330-2333.

McLauchlin J, Ripabelli G, Brett MM, Threlfall EJ. (2000). Amplified fragment length polymorphism (AFLP) analysis of *Clostridium perfringens* for epidemiological typing. *Int J Food Microbiol* **56,** 21-28.

McMillan DJ, Shojaei M, Chhatwal GS, Guzman CA, Walker MJ. (1996). Molecular analysis of the *bvg*-repressed urease of *Bordetella bronchiseptica. Microb Pathog* **21,** 379-394.

Meade BD, Bollen A. (1994). Recommendations for use of the polymerase chain reaction in the diagnosis of *Bordetella pertussis* infections. *J Med Microbiol* **41,** 51-55.

Menozzi FD, Mutombo R, Renauld G, Gantiez C, Hannah JH, Leininger E, Brennan MJ, Locht C. (1994). Heparin-inhibitable lectin activity of the filamentous hemagglutinin adhesin of *Bordetella pertussis. Infect Immun* **62,** 769-778.

Michaels RH. (1998). Pertussis: recent resurgence and advances in diagnostic techniques and control. *Sem Ped Infect Dis* **9,** 120-125.

Miller E, Vurdien JE, White JM. (1992). The epidemiology of pertussis in England and Wales. *Commun Dis Rep CDR Rev* **2,** R152-R154.

Miller E, White JM, Fairley CK. (1994). Pertussis vaccination. *Lancet* **344,** 1575-1576.

Mills KH. (2001). Immunity to *Bordetella pertussis. Microbes Infect* **3,** 655-677.

Missouri Department of Health and Senior Services. (2002). Summary of Notifiable Diseases, Missouri 2001. Walker, P. and Kahn, F. Jefferson City, MO, USA.

Moissenet D, Valcin M, Marchand V, Grimprel E, Begue P, Garbarg-Chenon A, Vu-Thien H. (1996). Comparative DNA analysis of *Bordetella pertussis* clinical isolates by pulsed-field gel electrophoresis, randomly amplified polymorphism DNA, and ERIC polymerase chain reaction. *FEMS Microbiol Lett* **143,** 127-132.

Molling P, Jacobsson S, Backman A, Olcen P. (2002). Direct and rapid identification and genogrouping of meningococci and *porA* amplification by LightCycler PCR. *J Clin Microbiol* **40,** 4531-4535.

Mooi FR. (1994). *Bordetella pertussis* Fimbrae. In: Per Klemm (ed) *Fimbrae- Adhesion, Genetics, Biogenesis and Vaccines.* CRC Press Inc., pp 115-126.

Mooi FR, van Oirschot H, Heuvelman K, van der Heide HG, Gaastra W, Willems RJ. (1998). Polymorphism in the *Bordetella pertussis* virulence factors P.69/pertactin and pertussis toxin in The Netherlands: temporal trends and evidence for vaccine-driven evolution. *Infect Immun* **66,** 670-675.

Mooi FR, He Q, van Oirschot H, Mertsola J. (1999). Variation in the *Bordetella pertussis* virulence factors pertussis toxin and pertactin in vaccine strains and clinical isolates in Finland. *Infect Immun* **67,** 3133-3134.

Mooi FR, Hallander H, von Konig CH, Hoet B, Guiso N. (2000). Epidemiological typing of *Bordetella pertussis* isolates: recommendations for a standard methodology. *Eur J Clin Microbiol Infect Dis* **19,** 174-181.

Munoz JJ. (1985). Biological activities of pertussigen (pertussis toxin). In: Sekura RD, Moss J, Vaughn M (eds) *Pertussis Toxin.* Orlando, Florida; Academic Press, pp 1-18.

Musser JM, Hewlett EL, Peppler MS, Selander RK. (1986). Genetic diversity and relationships in populations of *Bordetella* spp. *J Bacteriol* **166,** 230-237.

Nakase Y, Endoh M. (1988). Heat-liable toxin of *Bordetella pertussis.* In: Wardlaw AC, Parton R (eds) *Pathogenesis and Immunity in Pertussis.* Chichester; John Wiley & Sons, pp 211-229.

Nardone A, Pebody RG, Maple PAC, Andrews N, Gay NJ, Miller E. (2004). Sero-epidemiology of *Bordetella pertussis* in England and Wales. *Vaccine* **22,** 1314-1319.

Nei M, Tajima F. (1981). DNA polymorphism detectable by restriction endonucleases. *Genetics* **97,** 145-163.

Njamkepo E, Delisle F, Hagege I, Gerbaud G, Guiso N. (2000). *Bordetella holmesii* isolated from a patient with sickle cell anemia: analysis and comparison with other *Bordetella holmesii* isolates. *Clin Microbiol Infect* **6,** 131-136.

Njamkepo E, Rimlinger F, Thiberge S, Guiso N. (2002). Thirty-five years' experience with the whole-cell pertussis vaccine in France: vaccine strains analysis and immunogenicity. *Vaccine* **20,** 1290-1294.

Notifications of Infectious Diseases. (2003). Notifications of pertussis, England and Wales, by age group, 1980 - 2002.

Packard ER, Parton R, Coote JG, Fry NK. (2004). Sequence variation and conservation in virulence-related genes of *Bordetella pertussis* isolates from the UK. *J Med Microbiol* **53,** 355-365.

Parkhill J, Sebaihia M, Preston A, Murphy LD, Thomson N, Harris DE, Holden MT, Churcher CM, Bentley SD, Mungall KL, Cerdeno-Tarraga AM, Temple L, James K, Harris B, Quail MA, Achtman M, Atkin R, Baker S, Basham D, Bason N, Cherevach I, Chillingworth T, Collins M, Cronin A, Davis P, Doggett J, Feltwell T, Goble A, Hamlin N, Hauser H, Holroyd S, Jagels K, Leather S, Moule S, Norberczak H, O'Neil S, Ormond D, Price C, Rabbinowitsch E, Rutter S, Sanders M, Saunders D, Seeger K, Sharp S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Unwin L, Whitehead S, Barrell BG, Maskell DJ. (2003). Comparative analysis of the genome sequences of *Bordetella pertussis, Bordetella parapertussis* and *Bordetella bronchiseptica. Nat Genet* **35,** 32-40.

Parton R. (1998). *Bordetella.* In: Collier L, Balows A, Sussman M (eds) *Topley & Wilson's Microbiology and Microbial Infections.* London; Arnold, pp 901-918.

Passerini de Rossi BN, Friedman LE, Belzoni CB, Savoni S, Arico B, Rappouli R, Masigani V, Franco MA. (2003). *vir90*, a virulence-activated gene coding for a *Bordetella pertussis* iron-regulated outer membrane protein. *Res Microbiol* **154,** 443-450.

Peppler MS, Kuny S, Nevesinjac A, Rogers C, de Moissac YR, Knowles K, Lorange M, De Serres G, Talbot J. (2003). Strain variation among *Bordetella pertussis* isolates from Quebec and Alberta provinces of Canada from 1985 to 1994. *J Clin Microbiol* **41,** 3344-3347.

Peppoloni S, Pizza M, De Magistris MT, Bartoloni A, Rappuoli R. (1995). Acellular pertussis vaccine composed of genetically inactivated pertussis toxin. *Physiol Chem Phys Med NMR* **27,** 355-361.

Pittman M, Cox CB. (1965). Pertussis vaccine testing for freedom from toxicity. *J Appl Microbiol* **13,** 447-456.

Pittman M. (1984). Genus *Bordetella.* In: Krieg NR, Holt JG (eds) *Bergey's Manual of Systematic Bacteriology.* Baltimore; The Williams and Wilkins Co., pp 388-393.

Porter JF, Parton R, Wardlaw AC. (1991). Growth and survival of *Bordetella bronchiseptica* in natural waters and in buffered saline without added nutrients. *Appl Environ Microbiol* **57,** 1202-1206.

Porter JF, Wardlaw AC. (1993). Long-term survival of *Bordetella bronchiseptica* in lakewater and in buffered saline without added nutrients. *FEMS Microbiol Lett* **110,** 33-36.

Poynten M, McIntyre PB, Mooi FR, Heuvelman CJ, Gilbert GL. (2004). Temporal trends in circulating *Bordetella pertussis* strains: a contributor to the re-emergence of pertussis in Australia? *Epidemiol Infect* **132,** 185-193.

Preston NW. (1963). Type-specific immunity against whooping-cough. *BMJ* **2,** 724-726.

Preston NW. (1965). Effectiveness of Pertussis Vaccines. *BMJ* **2,** 11-13.

Preston NW. (1988). Pertussis Today. In: Wardlaw AC, Parton R (eds) *Pathogenesis and immunity in pertussis.* Chichester; John Wiley and Sons, pp 1-18.

Preston NW, Carter EJ. (1992). Serotype specificity of vaccine-induced immunity to pertussis. *Commun Dis Rep CDR Rev* **2,** R155-R156.

Prevost G, Jaulhac B, Piemont Y. (1992). DNA fingerprinting by pulsed-field gel electrophoresis is more effective than ribotyping in distinguishing among methicillin-resistant *Staphylococcus aureus* isolates. *J Clin Microbiol* **30,** 967-973.

Ranganathan S, Tasker R, Booy R, Habibi P, Nadel S, Britto J. (1999). Pertussis is increasing in unimmunized infants: is a change in policy needed? *Arch Dis Child* **80,** 297-299.

Re MC, Monari P, Bon I, Gibellinni D, Vitone F, Borderi M, La Placa M. (2001). Analysis of HIV-1 drug resistant mutations by line probe assay and direct sequencing in a cohort of therapy naive HIV-1 infected Italian patients. *BMC Microbiol* **1,** 30.

Reischl U, Lehn N, Sanden GN, Loeffelholz MJ. (2001). Real-time PCR assay targeting IS*481* of *Bordetella pertussis* and molecular basis for detecting *Bordetella holmesii. J Clin Microbiol* **39,** 1963-1966.

Roberts M, Fairweather NF, Leininger E, Pickard D, Hewlett EL, Robinson A, Hayward C, Dougan G, Charles IG. (1991). Construction and characterization of *Bordetella pertussis* mutants lacking the *vir*-regulated P.69 outer membrane protein. *Mol Microbiol* **5,** 1393-1404.

Roberts M, Parton R. (2001). *Bordetella pertussis.* In: Sussman M (ed) *Molecular Medical Microbiology.* Academic Press, pp 1565-1618.

Robinson A, Ashworth LA, Irons LI. (1989). Serotyping *Bordetella pertussis* strains. *Vaccine* **7,** 491-494.

Rohani P, Earn DJ, Grenfell BT. (2000). Impact of immunisation on pertussis transmission in England and Wales. *Lancet* **355,** 285-286.

Rosenthal S, Strebel P, Cassiday P, Sanden G, Brusuelas K, Wharton M. (1995). Pertussis infection among adults during the 1993 outbreak in Chicago. *J Infect Dis* **171,** 1650-1652.

Sambrook J, Russell DW. (2001). Molecular Cloning- a Laboratory Manual. New York, Cold Spring Harbour Laboratory Press.

Sanger F, Nicklen S, Coulson AR. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74,** 5463-5467.

Sato H, Sato Y. (1984). *Bordetella pertussis* infection in mice: correlation of specific antibodies against two antigens, pertussis toxin, and filamentous hemagglutinin with mouse protectivity in an intracerebral or aerosol challenge system. *Infect Immun* **46,** 415-421.

Scarlato V, Arico B, Prugnola A, Rappuoli R. (1991). Sequential activation and environmental regulation of virulence genes in *Bordetella pertussis*. *EMBO J* **10,** 3971-3975.

Schmidt JE, Tozzi AE, Rava L, Glismann S, on behalf of the EUVAC-NET country contacts. (2001). The EUVAC-NET survey: national pertussis surveillance systems in the European Union, Switzerland, Norway, and Iceland. *Eurosurveillance* **6,** 98-104.

Selander RK, Caugant DA, Ochman H, Musser JM, Gilmour MN, Whittam TS. (1986). Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. *Appl Environ Microbiol* **51,** 873-884.

Simmons DG, Gore AR, Hodgin EC. (1980). Altered immune function in turkey poults infected with *Alcaligenes faecalis*, the etiologic agent of turkey rhinotracheitis (coryza). *Avian Dis* **24,** 702-714.

Skowronski DM, De Serres G, MacDonald D, Wu W, Shaw C, Macnabb J, Champagne S, Patrick DM, Halperin SA. (2002). The changing age and seasonal profile of pertussis in Canada. *J Infect Dis* **185,** 1448-1453.

Sloan LM, Hopkins MK, Mitchell PS, Vetter EA, Rosenblatt JE, Harmsen WS, Cockerill FR, Patel R. (2002). Multiplex LightCycler PCR assay for detection and differentiation of *Bordetella pertussis* and *Bordetella parapertussis* in nasopharyngeal specimens. *J Clin Microbiol* **40,** 96-100.

Smith AM, Guzman CA, Walker MJ. (2001). The virulence factors of *Bordetella pertussis*: a matter of control. *FEMS Microbiol Rev* **25,** 309-333.

Southern EM. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* **98,** 503-517.

Stein PE, Boodhoo A, Armstrong GD, Cockle SA, Klein MH, Read RJ. (1994). The crystal structure of pertussis toxin. *Structure* **2,** 45-57.

Steven AC, Bisher ME, Trus BL, Thomas D, Zhang JM, Cowell JL. (1986). Helical structure of *Bordetella pertussis* fimbriae. *J Bacteriol* **167,** 968-974.

Stevenson M, Beard S, Finn A, Brennan A. (2002). Estimating the potential health gain and cost consequences of introducing a pre-school DTPa pertussis booster into the UK child vaccination schedule. *Vaccine* **20,** 1778-1786.

Stibitz S, Yang MS. (1999). Genomic plasticity in natural populations of *Bordetella pertussis*. *J Bacteriol* **181,** 5512-5515.

Stockbauer KE, Fuchslocher B, Miller JF, Cotter PA. (2001). Identification and characterization of BipA, a *Bordetella* Bvg-intermediate phase protein. *Mol Microbiol* **39,** 65-78.

Struelens MJ, the Members of the European Study Group on Epidemiological Markers. (1996). Consensus guidelines for appropriate use and evaluation of microbial epidemiologic typing systems. *Clin Microbiol Infect* **2,** 2-11.

Struelens MJ. (1998). Molecular epidemiologic typing systems of bacterial pathogens: current issues and perspectives. *Mem Inst Oswaldo Cruz* **93,** 581-585.

Swaminathan B, Barrett TJ, Hunter SB, Tauxe RV, the CDC PulseNet Task Force. (2001). PulseNet: the molecular subtyping network for foodborne bacterial disease surveillance, United States. *Emerg Infect Dis* **7,** 382-389.

Syedabubakar SN, Matthews RC, Preston NW, Owen D, Hillier V. (1995). Application of pulsed field gel electrophoresis to the 1993 epidemic of whooping cough in the UK. *Epidemiol Infect* **115,** 101-113.

Tenover FC, Arbeit RD, Goering RV, Mickelsen PA, Murray BE, Persing DH, Swaminathan B. (1995). Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. *J Clin Microbiol* **33,** 2233-2239.

Therre H, Baron S. (2000). Pertussis immunisation in Europe- the situation in late 1999. *Eurosurveillance* **5,** 6-10.

Tiru M, Askelof P, Granstrom M, Hallander H. (1997). *Bordetella pertussis* serotype of clinical isolates in Sweden during 1970-1995 and influence of vaccine efficacy studies. *Dev Biol Stand* **89,** 239-245.

Uhl MA, Miller JF. (1994). Autophosphorylation and phosphotransfer in the *Bordetella pertussis* BvgAS signal transduction cascade. *Proc Natl Acad Sci U S A* **91,** 1163-1167.

Uhl MA, Miller JF. (1995). BvgAS is sufficient for activation of the *Bordetella pertussis ptx* locus in *Escherichia coli*. *J Bacteriol* **177,** 6477-6485.

Valsangiacomo C, Baggi F, Gaia V, Balmelli T, Peduzzi R, Piffaretti JC. (1995). Use of amplified fragment length polymorphism in molecular typing of *Legionella pneumophila* and application to epidemiological studies. *J Clin Microbiol* **33,** 1716-1719.

van Belkum A, Kluytmans J, van Leeuwen W, Bax R, Quint W, Peters E, Fluit A, Vandenbroucke-Grauls C, van den Brule A, Koeleman H, Melchers W, Meis J, Elaichouni A, Vaneechoutte M, Moonens F, Struelens M, Tenover FC, Verbrugh H. (1995). Multicenter evaluation of arbitrarily primed PCR for typing of *Staphylococcus aureus* strains. *J Clin Microbiol* **33,** 1537-1547.

van Belkum A, Struelens M, de Visser A, Verbrugh H, Tibayrenc M. (2001). Role of genomic typing in taxonomy, evolutionary genetics, and microbial epidemiology. *Clin Microbiol Rev* **14,** 547-560.

van Buynder PG, Owen D, Vurdien JE, Andrews NJ, Matthews RC, Miller E. (1999). *Bordetella pertussis* surveillance in England and Wales: 1995-7. *Epidemiol Infect* **123,** 403-411.

van der Zee A, Groenendijk H, Peeters M, Mooi FR. (1996a). The differentiation of *Bordetella parapertussis* and *Bordetella bronchiseptica* from humans and animals as determined by DNA polymorphism mediated by two different insertion sequence elements suggests their phylogenetic relationship. *Int J Syst Bacteriol* **46,** 640-647.

van der Zee A, Vernooij S, Peeters M, van Embden J, Mooi FR. (1996b). Dynamics of the population structure of *Bordetella pertussis* as measured by IS*1002*-associated RFLP: comparison of pre- and post-vaccination strains and global distribution. *Microbiology* **142,** 3479-3485.

van der Zee A, Mooi F, van Embden J, Musser J. (1997). Molecular evolution and host adaptation of *Bordetella* spp.: phylogenetic analysis using multilocus enzyme electrophoresis and typing with three insertion sequences. *J Bacteriol* **179,** 6609-6617.

van Embden JD, Cave MD, Crawford JT, Dale JW, Eisenach KD, Gicquel B, Hermans P, Martin C, McAdam R, Shinnick TM, Small PM. (1993). Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. *J Clin Microbiol* **31,** 406-409.

van Loo IH, van der Heide HG, Nagelkerke NJ, Verhoef J, Mooi FR. (1999). Temporal trends in the population structure of *Bordetella pertussis* during 1949-1996 in a highly vaccinated population. *J Infect Dis* **179,** 915-923.

van Loo IH, Heuvelman KJ, King AJ, Mooi FR. (2002). Multilocus sequence typing of *Bordetella pertussis* based on surface protein genes. *J Clin Microbiol* **40,** 1994-2001.

van Loo IH, Mooi FR. (2002). Changes in the Dutch *Bordetella pertussis* population in the first 20 years after the introduction of whole-cell vaccines. *Microbiology* **148,** 2011-2018.

Vandamme P, Hommez J, Vancanneyt M, Monsieurs M, Hoste B, Cookson B, von Konig CH, Kersters K, Blackall PJ. (1995). *Bordetella hinzii* sp. nov., isolated from poultry and humans. *Int J Syst Bacteriol* **45,** 37-45.

Vandamme P, Heyndrickx M, Vancanneyt M, Hoste B, De Vos P, Falsen E, Kersters K, Hinz KH. (1996). *Bordetella trematum* sp. nov., isolated from wounds and ear infections in humans, and reassessment of *Alcaligenes denitrificans* Ruger and Tan 1983. *Int J Syst Bacteriol* **46,** 849-858.

Versalovic J, Koeuth T, Lupski JR. (1991). Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Res* **19,** 6823-6831.

Viljanen MK, Ruuskanen O, Granberg C, Salmi TT. (1982). Serological diagnosis of pertussis: IgM, IgA and IgG antibodies against *Bordetella pertussis* measured by enzyme-linked immunosorbent assay (ELISA). *Scand J Infect Dis* **14,** 117-122.

von Konig CH, Postels-Multani S, Bock HL, Schmitt HJ. (1995). Pertussis in adults: frequency of transmission after household exposure. *Lancet* **346,** 1326-1329.

von Konig CH, Halperin S, Riffelmann M, Guiso N. (2002). Pertussis of adults and infants. *Lancet Infect Dis* **2,** 744-750.

von Wintzingerode F, Schattke A, Siddiqui RA, Rosick U, Gobel UB, Gross R. (2001). *Bordetella petrii* sp. nov., isolated from an anaerobic bioreactor, and emended description of the genus *Bordetella. Int J Syst Evol Microbiol* **51,** 1257-1265.

von Wintzingerode F, Gerlach G, Schneider B, Gross R. (2002). Phylogenetic relationships and virulence evolution in the genus *Bordetella. Curr Top Microbiol Immunol* **264,** 177-199.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* **23,** 4407-4414.

Walker E. (1988). Clinical Aspects of Pertussis. In: Wardlaw AC, Parton R (eds) *Pathogenesis and Imunity in Pertussis.* Chichseter; John Wiley & Sons, pp.

Weber C, Boursaux-Eude C, Coralie G, Caro V, Guiso N. (2001). Polymorphism of *Bordetella pertussis* isolates circulating for the last 10 years in France, where a single effective whole-cell vaccine has been used for more than 30 years. *J Clin Microbiol* **39,** 4396-4403.

Weiss AA, Falkow S. (1984). Genetic analysis of phase change in *Bordetella pertussis. Infect Immun* **43,** 263-269.

Weiss AA, Johnson FD, Burns DL. (1993). Molecular characterization of an operon required for pertussis toxin secretion. *Proc Natl Acad Sci U S A* **90,** 2970-2974.

Weiss A. (1997). Mucosal immune defenses and the response of *Bordetella pertussis. ASM News* **63,** 22-28.

Welsh J, McClelland M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res* **18,** 7213-7218.

Weyant RS, Hollis DG, Weaver RE, Amin MF, Steigerwalt AG, O'Connor SP, Whitney AM, Daneshvar MI, Moss CW, Brenner DJ. (1995). *Bordetella holmesii* sp. nov., a new gram-negative species associated with septicemia. *J Clin Microbiol* **33,** 1-7.

Willems R, Paul A, van der Heide HG, ter Avest AR, Mooi FR. (1990). Fimbrial phase variation in *Bordetella pertussis*: a novel mechanism for transcriptional regulation. *EMBO J* **9,** 2803-2809.

Willems RJ, van der Heide HG, Mooi FR. (1992). Characterization of a *Bordetella pertussis* fimbrial gene cluster which is located directly downstream of the filamentous haemagglutinin gene. *Mol Microbiol* **6,** 2661-2671.

Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* **18,** 6531-6535.

World Health Organisation. (1979). WHO Expert Committee on Biological Standardisation. **638,** 61-65 Geneva, Switzerland, WHO.

World Health Organisation. (1991). WHO meeting on case definition of pertussis. **MIM/EPI/PERT/1991,** Geneva, World Health Organisation.

World Health Organisation. (2001). Pertussis surveillance- A global meeting. Geneva, Switzerland, Department of Vaccines and Biologicals, WHO. 16-10-2000.

Yih WK, Silva EA, Ida J, Harrington N, Lett SM, George H. (1999). *Bordetella holmesii*-like organisms isolated from Massachusetts patients with pertussis-like symptoms. *Emerg Infect Dis* **5,** 441-443.

Yuk MH, Harvill ET, Miller JF. (1998). The BvgAS virulence control system regulates type III secretion in *Bordetella bronchiseptica*. *Mol Microbiol* **28,** 945-959.

12 APPENDICES

12.1 Appendix 1. Amino acid abbreviations

NAME	3 Letter Abbreviation	1 Letter Abbreviation
Alanine	Ala	A
Cysteine	Cys	C
Aspartic Acid	Asp	D
Glutamic Acid	Glu	E
Phenylalanine	Phe	F
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Lysine	Lys	K
Leucine	Leu	L
Methionine	Met	M
Asparagine	Asn	N
Proline	Pro	P
Glutamine	Gln	Q
Arginine	Arg	R
Serine	Ser	S
Threonine	Thr	T
Valine	Val	V
Tryptophan	Trp	W
Tyrosine	Tyr	Y

12.2 Appendix 2. Address list of suppliers

Applied Biosystems, Lingley House, 120 Birchwood Boulevard, Warrington, Cheshire, WA3 7QH.

Applied Maths BVBA, Keistraat 120, 9830 Sint-Martens-Latem, Belguim.

Amersham Biosciences UK Ltd, Amersham Place, Little Chalfont, Buckinghamshire, HP7 9NA.

Aventis Pasteur MSD Ltd, Mallards Reach, Bridge Avenue, Maidenhead, Berkshire, SL6 1QP.

bioMérieux UK Limited, Grafton Way, Basingstoke, Hampshire, RG22 6HY.

BioWhittaker Molecular Applications (BMA), 1 Ashville Way, Wokingham, Berkshire, RG41 2PL.

BDH Laboratory Supplies, Poole, Dorset, BH15 1TD.

Beckman Coulter UK Ltd, Oakley Court, Kingsmead Business Park, High Wycombe, Buckinghamshire, HP11 1JU.

Bio-Rad Laboratories Ltd, Bio-Rad House, Maylands Avenue, Hemel Hempstead, Hertfordshire, HP2 7TD.

Camlab Ltd, Nuffield Road, Cambridge, Cambridgeshire, CB4 1TH.

Centre for Applied Microbiology and Research (CAMR), Porton Down, Salisbury, Wiltshire, SP4 OJG.

Creative Beadcraft Ltd, Denmark Works, Beamond End, Nr Amersham, Buckinghamshire, HP7 0RX.

Cruachem Ltd, Todd Campus, West of Scotland Science Park, Acre Road, Glasgow, G20 0UA.

Eppendorf UK Ltd, 10 Signet Court, Swanns Road, Cambridge, Cambridgeshire, CB5 8LA.

Genetic Research Instrumentation Limited (GRI), Gene House, Queensborough Lane, Rayne, Braintree, Essex, CM7 8TF.

Grant Instruments (Cambridge) Ltd, Shepreth, Cambridgeshire, SG8 6GB.

Hewlett Packard, CC Centre, Cain Road, Bracknell, Berkshire, RG12 1HU.

Hybaid Limited, Action Court, Ashford Road, Ashford, Middlesex, TW15 1XB.

Institut Pasteur, 25-28, rue du Dr Roux, 75015 Paris, France.

Integra Biosciences, Northumbria Biologicals Ltd, Nelson Industrial Estate, Cramlington, Northumberland, NE23 9BL.

Invitrogen Ltd[6], Inchinnan Business Park, 3 Fountain Drive, Paisley, PA4 9RF.

MBI Fermentas, Fermentas UAB, V. Graiciuno 8, Vilnius 2028, Lithuania.

---

[6]Life Technologies has now merged with Invitrogen

Media Services, Public Health Laboratory Service (PHLS), 61 Colindale Avenue, London, NW9 5HT.

Missouri State Public Health Lab, MO DOH, 307 W. McCarty Street, Jefferson City, MO 65102.

MWG-BIOTECH (UK) Ltd, Mill Court, Featherstone Road, Wolverton Mill South, Milton Keynes, Bedfordshire, MK12 5RD.

National Center for Biotechnology Information (NCBI), National Library of Medicine, Building 38A, Bethesda, MD 20894.

National Collection of Type Cultures (NCTC), Health Protection Agency (HPA), 61 Colindale Avenue, London, NW9 5HT.

National Institute for Biological Standards and Control (NIBSC), Blanche Lane, South Mimms, Potters Bar, Hertfordshire, EN6 3QG.

National Institute of Public Health and the Environment in The Netherlands (RIVM), Antonie van Leeuwenhoeklaan 9, 3721 MA Bilthoven, The Netherlands.

New England BioLabs (UK) Ltd, 73 Knowl Piece, Wilbury Way, Hitchin, Hertfordshire, SG4 0TY.

Numerical Algorithms Group Ltd, Wilkinson House, Jordan Hill Road, Oxford, OX2 8DR.

Pertussis Reference Laboratory (PRL), University of Manchester, Central Manchester Healthcare NHS Trust, Oxford Road, Manchester, M13 9WL.

Promega UK Ltd, Delta House, Chilworth Research Centre, Southampton, SO16 7NS.

Qiagen Ltd, Boundary Court, Gatwick Road, Crawley, West Sussex, RH10 9AX.

Roche Diagnostics Ltd, Bell Lane, Lewes, East Sussex, BN7 1LG.

The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA.

Sigma-Aldrich Company Ltd, Fancy Road, Poole, Dorset, BH12 4QH.

Stata Corporation, 4905 Lakeway Drive, College Station, Texas, 77845, USA.

Thermo Life Sciences, Unit 5, The Ringway Centre, Edison Road, Basingstoke, Hampshire, RG21 6YH.

Unicam Limited, PO Box 209, York Street, Cambridge, Cambridgeshire, CB1 2SX.

12.3 Appendix 3. Blue shading corresponds to those strains which were not used as part of this thesis.

Part A. Historical *Bordetella pertussis* isolates (see 2.1.1, page 55)

Source culture collection[1]	Strain number	Other strain designations	Further details	Original source of isolate	Earliest year known	Reasons for exclusion
NCTC	10901	Hektoen; NCTC 364; Phase IV		St Mary's Hospital London	1920	
CN	128			Dr IE MacLean, St Mary's Hospital Paddington	1941	
CN	129	White's' strain		Dr IE MacLean, St Mary's Hospital Paddington	1941	
CN	130	Eos' strain		Dr IE MacLean, St Mary's Hospital Paddington	1941	
CN	132	21389/41		Dr Cruickshank, NW Hospital Group Lab, Hampstead NW3	1941	
CN	133	21453/41		Dr Cruickshank, NW Hospital Group Lab, Hampstead NW3	1941	
CN	134	21015/41		Dr Cruickshank, NW Hospital Group Lab, Hampstead NW3	1941	
CN	135	25167/41		Dr Cruickshank, NW Hospital Group Lab, Hampstead NW3	1941	
CN	136	25341/41		Dr Cruickshank, NW Hospital Group Lab, Hampstead NW3	1941	
CN	137	26381/41		Dr Cruickshank, NW Hospital Group Lab, Hampstead NW3	1941	
CN	138	26540/41		Dr Cruickshank, NW Hospital Group Lab, Hampstead NW3	1941	
CN	127	Wilson strain			1942	Unable to grow from ampoule
CN	140	28036/41		Dr Cruickshank, NW Hospital Group Lab, Hampstead NW3	1942	
CN	141	27746/42		Dr Cruickshank, NW Hospital Group Lab, Hampstead NW3	1942	
CN	142	27902/41		Dr Cruickshank, NW Hospital Group Lab, Hampstead NW3	1942	
CN	143	27913/41		Dr Cruickshank, NW Hospital Group Lab, Hampstead NW3	1942	
CN	350	E Knight		Dr Wright, NE Group Hospitals Laboratory	1942	
CN	351	M Newton		Dr Wright, NE Group Hospitals Laboratory	1942	
CN	352	T Newton		Dr Wright, NE Group Hospitals Laboratory	1942	
CN	453	Strain E		Dr DG Evans, MRC, Hampstead	1942	
CN	454	Strain OE		Dr DG Evans, MRC, Hampstead	1942	
CN	455	OJ		Dr DG Evans, MRC, Hampstead	1942	Unable to grow from ampoule
CN	537	A.1	Isolated from blood from tail of mouse injected at WRL ip with mixed CN132, CN133, CN137	WRL	1942	Laboratory strain
CN	538	B.1	Isolated from blood from tail of mouse injected at WRL ip with mixed CN132, CN133, CN137	WRL	1942	Laboratory strain

Source culture collection[1]	Strain number	Other strain designations	Further details	Original source of isolate	Earliest year known	Reasons for exclusion
CN	741	B.667		Dr Ungar, Glaxo Ltd	1943	
CN	742	B.668		Dr Ungar, Glaxo Ltd	1943	
CN	743	B.634		Dr Ungar, Glaxo Ltd	1943	
CN	908	B.26659	Atypical D Parker	Prof GS Wilson, Oxford	1944	
CN	909	B.26408	Atypical K Knibbs	Prof GS Wilson, Oxford	1944	
CN	910	B.31122	Atypical M Smith	Prof GS Wilson, Oxford	1944	
CN	911	B.26406	Atypical B Preston	Prof GS Wilson, Oxford	1944	
CN	1005	P10646		E Topley	1944	
CN	1006	P10650		E Topley	1944	
CN	1102	Phase I strain		Dr Evans	1944	
CN	1260		Virulent by intra-cerebral route	Dr Margaret Pittman, National Institute of Health, USA	1945	USA laboratory strain
CN	1262	5-27	Pearl Kendrick strain 18-323 Passaged through brain	Dr Margaret Pittman, National Institute of Health, USA	1945	USA laboratory strain
CN	1407	G.173		E Topley, EPHLS, Oxford	1946	
CN	1529	CN137	Adapted to grow on nutrient agar	WRL	1946	Laboratory strain
CN	1791		Trained to grow on modified Hornibrook's medium with charcoal	Dr Pollock, Lister Institute	1947	Laboratory strain
CN	2055			London School of Hygiene and Tropical Medicine	1947	
CN	2128	CN134	Passaged through mouse and re-isolated from lung	WRL	1948	Laboratory strain
CN	2216	CN134	Adapted to grow on nutrient agar. Phase IV		1948	Laboratory strain
CN	2217	CN1262	Adapted to grow on nutrient agar		1948	Laboratory strain
CN	2374	659; Green 51	Phase I	Dr A Kimball, Minnesota Dept of Health	1948	USA strain
CN	2376	538; Green 9	Phase III	Dr A Kimball, Minnesota Dept of Health	1948	USA strain
CN	2377	587; Green 42	Phase III	Dr A Kimball, Minnesota Dept of Health	1948	USA strain
CN	2420	L.12		Dr Standfast, Lister Institute	1948	
CN	2421	L.11		Dr Standfast, Lister Institute	1948	
CN	2422	C.5146		Dr Standfast, Lister Institute	1948	
CN	2423	C.5476		Dr Standfast, Lister Institute	1948	
CN	2424	863/P3		Dr Standfast, Lister Institute	1948	
CN	2454	CN1262	Deposited as brain-passaged	Chemotherapy Dept, WRL	1948	Laboratory strain

Source culture collection[1]	Strain number	Other strain designations	Further details	Original source of isolate	Earliest year known	Reasons for exclusion
CN	2455		Isolated from sputum of whooping cough case 'Mason', Farnborough Hospital	Chemotherapy Dept, WRL	1948	
CN	2598	P.61		Dr Keagh, Commonwealth Laboratories, Australia	1948	Australian strain
CN	2808	83E	Typical	Glaxo Laboratories	1949	
CN	2809	154E	Atypical	Glaxo Laboratories	1949	
CN	2991	25456		Dr Holt, CPHL, Colindale	1949	
CN	2993	25966		Dr Holt, CPHL, Colindale	1949	
CN	2994	26602		Dr Holt, CPHL, Colindale	1949	Unable to grow from ampoule
CN	2995	25302		Dr Holt, CPHL, Colindale	1949	
CN	2996	25654		Dr Holt, CPHL, Colindale	1949	
CN	2998	26898		Dr Holt, CPHL, Colindale	1949	
CN	2999	23228		Dr Holt, CPHL, Colindale	1949	
CN	3000	25500		Dr Holt, CPHL, Colindale	1949	
CN	3001	26190		Dr Holt, CPHL, Colindale	1949	
CN	3011	27143		Dr Holt, CPHL, Colindale	1949	
CN	3012	23180		Dr Holt, CPHL, Colindale	1949	Unable to grow from ampoule
CN	3013	26897		Dr Holt, CPHL, Colindale	1949	
CN	3014	26191		Dr Holt, CPHL, Colindale	1949	
CN	3015	23182		Dr Holt, CPHL, Colindale	1949	Unable to grow from ampoule
CN	3016	26646		Dr Holt, CPHL, Colindale	1949	
CN	3100	260E/260E	Vaccine strain	Dr Ungar, Glaxo Laboratories	1950	Laboratory strain
CN	3108	L.46C.21/1; AFES/EE 26.1.50		Dr Standfast, Lister Institute, Elstree	1950	
CN	3109	C.26573E.24/1; AFES/EE 26.1.50		Dr Standfast, Lister Institute, Elstree	1950	
CN	3208	C.26573E.24/1; AFES/EE 26.1.50	Duplicate culture of CN3109	Dr Standfast, Lister Institute, Elstree	1950	Duplicate strain
NCTC	8189	2048		HD Holt, PHLS Colindale in 1950	1950	
NCTC	8605	18234		HD Holt, PHLS Colindale in 1950	1950	
NCTC	8616	18482		HD Holt, PHLS Colindale in 1950	1950	
NCTC	10902	L84 Phase IV		MRC trial, Leeds	1950	

Source culture collection[1]	Strain number	Other strain designations	Further details	Original source of isolate	Earliest year known	Reasons for exclusion
NCTC	11089	L84 Phase I	CRC Harrow in 1976		1950	
NCTC	8474	Kendrick 17921; L4,ATCC9340,NIH5		M Pittman,Maryland in 1951/P Kendrick	1951	USA strain
NCTC	10905	3747 variant	JM Dolby, LIPM in 1973; Serotype 1.0.0.0.0.0.	EK Andersen, Copenhagen	1953	Danish strain
NCTC	10906	3865	JM Dolby, LIPM in 1973; Serotype 1.2.0.4.0.0.	EK Andersen, Copenhagen	1953	Danish strain
NCTC	10908	3747	JM Dolby, LIPM in 1973; Serotype 1.2.0.0.5.6.	EK Andersen, Copenhagen	1953	Danish strain
CN	3700	BLO.54		Prof CL Oakley, Medical School, Leeds	1954	
CN	3701	JEF.54		Prof CL Oakley, Medical School, Leeds	1954	
CN	3704	MCL.54		Prof CL Oakley, Medical School, Leeds	1954	
CN	3705	SMI.54		Prof CL Oakley, Medical School, Leeds	1954	
CN	4131	332E		Lister Institute	1956	
CN	4154			Dr G Turner, School of Medicine, Leeds	1956	
CN	4158	GL.291ESCI(1+4-12)		Dr Standfast, Lister Institute, Elstree	1956	
CN	4159	GL.332ESCI(2-3+5+8+10)		Dr Standfast, Lister Institute, Elstree	1956	
CN	4160	GL.343ESCI(1-4)(2)		Dr Standfast, Lister Institute, Elstree	1956	
CN	4161	GL.353ESCI(7-9)		Dr Standfast, Lister Institute, Elstree	1956	
CN	4329	25593		Institute for the production of sera and vaccines, Zagreb	1957	Croatian strain
CN	4330	26426		Institute for the production of sera and vaccines, Zagreb	1957	Croatian strain
CN	4331	79792		Institute for the production of sera and vaccines, Zagreb	1957	Croatian strain
CN	4951	134		Dr Cohen, Utrecht via Mr G Turner, WRL	1960	Dutch strain
CN	5420	32316		Hither Green Hospital via Mr G Turner, WRL	1963	
CN	5432	G.1.332		A.F.B. Standfast, Lister Institute, Elstree, Glaxo	1963	
CN	5433	G.1.260		A.F.B. Standfast, Lister Institute, Elstree, Glaxo	1963	
CN	5434	G.1.219		A.F.B. Standfast, Lister Institute, Elstree, Glaxo	1963	
CN	5436	G.1.18398	Michigan	A.F.B. Standfast, Lister Institute, Elstree, Glaxo	1963	USA strain
CN	5438	34901		Hither Green Hospital via Mr G Turner, WRL	1963	
CN	5445	35846		Hither Green Hospital via Mr G Turner, WRL	1963	
CN	5447	37046		Hither Green Hospital via Mr G Turner, WRL	1963	
CN	5478	39156		Hither Green Hospital via Mr G Turner, WRL	1963	
CN	5618	NCTC9797; CN1262		NCTC	1964	Duplicate strain

Source culture collection[1]	Strain number	Other strain designations	Further details	Original source of isolate	Earliest year known	Reasons for exclusion
CN	5631	M.9		Dr NW Preston	1964	
NCTC	10907	B16 (ATCC12742)	JM Dolby, LIPM in 1973; Serotype 1.0.3.0.0.0.	M Haire, Belfast	1964	
CN	5750	CN4132	non-haemolytic variant	Dr J Cameron, WRL	1965	Laboratory strain
CN	5756	CN2991	Hameolytic variant	Dr J Cameron, WRL	1965	Laboratory strain
NCTC	10909	D35726	JM Dolby, LIPM in 1973; Serotype 1.0.3.0.0.6.	PHLS Manchester	1966	
NCTC	10911	D30042	JM Dolby, LIPM in 1973; Serotype 1.2.3.4.0.0.	PHLS Manchester	1966	
NCTC	10910	D41633	JM Dolby, LIPM in 1973; Serotype 1.(2).3.4.0.0.	PHLS Coventry	1967	
Glasgow	77/18319			Prof. GT Stewart, Ruchill Hospital	1977	
Glasgow	77/19110			Prof. GT Stewart, Ruchill Hospital	1977	
Glasgow	77/24171			Prof. GT Stewart, Ruchill Hospital	1977	
Glasgow	77/24833			Prof. GT Stewart, Ruchill Hospital	1977	
Glasgow	77/25171			Prof. GT Stewart, Ruchill Hospital	1977	
Glasgow	77/26621			Prof. GT Stewart, Ruchill Hospital	1977	
Glasgow	77/31726			Prof. GT Stewart, Ruchill Hospital	1977	
Glasgow	77/35296			Prof. GT Stewart, Ruchill Hospital	1977	
Glasgow	77/37203			Prof. GT Stewart, Ruchill Hospital	1977	
Glasgow	77/6124			Prof. GT Stewart, Ruchill Hospital	1977	
MANCH	18335				1978	
MANCH	36379				1978	
MANCH	39424				1978	
MANCH	8002				1979	
MANCH	29836				1979	
CAMR	DCH10	Morse 114			1981	Laboratory strain
CAMR	DCH7	M2 (Small colony type)	small cell isolate from DCH2-6		1981	Laboratory strain
CAMR	DCH9	Tohama			1981	Laboratory strain
CAMR	DCH100	41833		Odstock 1982/83?	1982	
CAMR	DCH101	42944		Odstock 1982/83?	1982	
CAMR	DCH102	15624			1982	
CAMR	DCH103	15671			1982	

Source culture collection[1]	Strain number	Other strain designations	Further details	Original source of isolate	Earliest year known	Reasons for exclusion
CAMR	DCH104	15839			1982	
CAMR	DCH105	16324			1982	
CAMR	DCH138	3577		Odstock 1982/83?	1982	
CAMR	DCH145	9304		Odstock 1982/83?	1982	
CAMR	DCH152	7845		Odstock 1982/83?	1982	
CAMR	DCH18	2402			1982	
CAMR	DCH19	2888			1982	
CAMR	DCH20	2571			1982	
CAMR	DCH25	8061			1982	
CAMR	DCH26	8214			1982	
CAMR	DCH27	8295			1982	
CAMR	DCH28	8398			1982	
CAMR	DCH29	8399			1982	
CAMR	DCH30	9508			1982	
CAMR	DCH31	9564			1982	
CAMR	DCH32	9601			1982	
CAMR	DCH33	9641			1982	
CAMR	DCH34	9642			1982	
CAMR	DCH35	9780			1982	
CAMR	DCH36	9963			1982	
CAMR	DCH37	10019			1982	
CAMR	DCH38	10117			1982	
CAMR	DCH39	10118			1982	
CAMR	DCH40	10266			1982	
CAMR	DCH41	10267			1982	
CAMR	DCH42	10338			1982	
CAMR	DCH43	10127			1982	
CAMR	DCH44	10287			1982	
CAMR	DCH45	10379			1982	
CAMR	DCH46	10428			1982	
CAMR	DCH47	10520			1982	
CAMR	DCH48	10723			1982	

Source culture collection[1]	Strain number	Other strain designations	Further details	Original source of isolate	Earliest year known	Reasons for exclusion
CAMR	DCH49	10755			1982	
CAMR	DCH50	10753			1982	
CAMR	DCH51	10755			1982	
CAMR	DCH52	10772			1982	
CAMR	DCH53	11537			1982	
CAMR	DCH54	11581			1982	
CAMR	DCH55	11583			1982	
CAMR	DCH56	11620			1982	
CAMR	DCH57	11621			1982	
CAMR	DCH58	11693			1982	
CAMR	DCH59	11755			1982	
CAMR	DCH60	11774			1982	
CAMR	DCH61	11830			1982	
CAMR	DCH62	11856			1982	
CAMR	DCH63	11922			1982	
CAMR	DCH64	11927			1982	
CAMR	DCH65	11929			1982	
CAMR	DCH66	11995			1982	
CAMR	DCH67	11996			1982	
CAMR	DCH68	12072			1982	
CAMR	DCH69	12083			1982	
CAMR	DCH70	12475			1982	
CAMR	DCH71	12692			1982	
CAMR	DCH72	12744			1982	
CAMR	DCH73	12746			1982	
CAMR	DCH74	12751			1982	
CAMR	DCH75	12754			1982	
CAMR	DCH76	12793			1982	
CAMR	DCH77	12821			1982	
CAMR	DCH78	13028			1982	
CAMR	DCH79	13091			1982	
CAMR	DCH80	13199			1982	

Source culture collection[1]	Strain number	Other strain designations	Further details	Original source of isolate	Earliest year known	Reasons for exclusion
CAMR	DCH81	13201			1982	
CAMR	DCH82	13255			1982	
CAMR	DCH83	13450			1982	
CAMR	DCH84	13570			1982	
CAMR	DCH85	13630			1982	
CAMR	DCH86	13657			1982	
CAMR	DCH87	13736			1982	
CAMR	DCH88	13787			1982	
CAMR	DCH89	14067			1982	
CAMR	DCH90	13930			1982	
CAMR	DCH91	13960			1982	
CAMR	DCH92	14041			1982	
CAMR	DCH93	14059			1982	
CAMR	DCH94	14061			1982	
CAMR	DCH95	14062			1982	
CAMR	DCH96	14446			1982	
CAMR	DCH97	14565			1982	
CAMR	DCH98	14671			1982	
CAMR	DCH99	14908			1982	
CAMR	DCH110	K3747	ExAmp 22		1982	Laboratory strain
CAMR	DCH127	260			1983	
CAMR	DCH128	451			1983	
CAMR	DCH129	424			1983	
CAMR	DCH130	461			1983	
CAMR	DCH131	1035			1983	
CAMR	DCH132	1359			1983	
CAMR	DCH133	1340			1983	
CAMR	DCH134	1434			1983	
CAMR	DCH135	1675			1983	
CAMR	DCH136	1676			1983	
CAMR	DCH137	1868			1983	
CAMR	DCH139	2105			1983	

Source culture collection[1]	Strain number	Other strain designations	Further details	Original source of isolate	Earliest year known	Reasons for exclusion
CAMR	DCH140	2492			1983	
CAMR	DCH141	2977			1983	
CAMR	DCH142	3209			1983	
CAMR	DCH143	3341			1983	
CAMR	DCH144	3438			1983	
CAMR	DCH147	6153			1983	
CAMR	DCH148	6837			1983	
CAMR	DCH149	7153			1983	
CAMR	DCH150	7454			1983	
CAMR	DCH151	7726			1983	
CAMR	DCH153	5448			1983	
CAMR	DCH154	8519			1983	
CAMR	DCH155	8837			1983	
CAMR	DCH156	8838			1983	
CAMR	DCH157	9634			1983	
CAMR	DCH158	9804			1983	
CAMR	DCH159	10038			1983	
CAMR	DCH160	10802			1983	
CAMR	DCH161	12894			1983	
CAMR	DCH162	12895			1983	
CAMR	DCH163	13109			1983	
CAMR	DCH164	13967			1983	
CAMR	DCH165	14086			1983	
CAMR	DCH166	14234			1983	
CAMR	DCH167	14746			1983	
CAMR	DCH168	15108			1983	
CAMR	DCH169	15120			1983	
CAMR	DCH170	15533			1983	
CAMR	DCH171	16182			1983	
CAMR	DCH172	16326			1983	
CAMR	DCH173	16743			1983	
CAMR	DCH174	16744			1983	

Source culture collection[1]	Strain number	Other strain designations	Further details	Original source of isolate	Earliest year known	Reasons for exclusion
CAMR	DCH175	16873			1983	
CAMR	DCH176	16874			1983	
CAMR	DCH177	26251			1983	
CAMR	DCH180	29551			1983	
CAMR	DCH181	20582			1983	
CN	8474	348		Dr Alison A Weiss, University of Virginia, School of Medicine, Charlottesville, USA via MR K Cownley, WRL, Beckenham	1984	USA strain
CAMR	DCH182	84/88			1984	
CAMR	DCH183	84/2271			1984	
CAMR	DCH184	2761			1984	
CAMR	DCH185	4789			1984	
CAMR	DCH186	5967			1984	
CAMR	DCH1 Passage	A3, Novotny	3 passage		1984	Laboratory strain
CAMR	DCH8	Tohama (ex Manclark)	M/C		1984	Laboratory strain
CN	8486				1985	
CN	8530				1985	
CN	8533				1985	
CN	8535				1985	
CN	8537				1985	
CN	8539				1985	
CN	8540				1985	
CN	8541				1985	
CN	8549				1985	
CN	8552				1985	
CN	8554				1985	
CN	8555				1985	
CN	8636				1985	
CAMR	Wellcome 28	BP/MS	Jan-85		1985	Laboratory strain
Glasgow	Taberman				1986	

[1]NCTC=National Collection of Type Cultures; CN=Wellcome Bacterial Collection; Glasgow=University of Glasgow; MANCH=Manchester University; CAMR=Centre for Applied Microbiology and Research.

Part B. Contemporary UK *Bordetella pertussis* isolates (see 2.1.2, page 57)

Isolate	Sex	Date of isolation	Date of birth	Age group	Laboratory	Region	Other notes	Reasons for exclusion of thesis
98K290	M	21/09/98	20/01/92	5+ years	Frimley Park Hospital	South Thames		
98K297	M	12/09/98	23/08/97	1-4 years	Alder Hey County Hospital			
98K299	F	09/09/98	07/05/98	2-5 months	Good Hope General Hospital	West Midlands		
98K300	M	27/09/98	26/06/98	2-5 months	Wrexham Maelor General Hospital	Wales		
98K301	F	30/11/98	15/07/98	2-5 months	Alder Hey County Hospital	North West		
98K302	F	29/09/98	23/06/98	2-5 months	Manchester Booth Hall Hospital	North West	Typical asthma bronchiolitis	
98K303		12/10/98	19/08/98	< 2 months	Lincoln Public Health Laboratory			
98K304	M	30/09/98	31/10/94	1-4 years	Conquest Hospital	South Thames		
98K305	F	25/09/98	03/03/92	5+ years	Bristol Royal Hospital	South and West	Vomiting	
98K306	F	08/10/98	25/04/98	2-5 months	Gloucester Public Health Hospital	South and West		
98K310	F	28/09/98	24/03/97	1-4 years	Banbury Horton General Hospital	Anglia and Oxford	Paroxysmal cough	
98K311	M	19/10/98	14/01/98	6-11 months	Alder Hey County Hospital	North West		
98K313	M	14/10/98	27/06/91	5+ years	Nottingham Public Health Laboratory	Trent	Worsening of asthma	
98K314		07/10/98	27/03/98	6-11 months	Londonderry Altnaglevin Hospital			
98K315	M	14/09/98	09/12/97	6-11 months	Ashford Public Health Laboratory	South Thames		
98K316	F	15/10/98	12/08/98	2-5 months	Truro Public Health Laboratory	South and West	Baby, 9 weeks, presented with a cough; 1 week respiratory distress	
98K317	F	19/10/98	08/09/98	< 2 months	Sheffield Public Health Laboratory	Trent		
98K318	M	19/10/98	23/03/94	1-4 years	Manchester Trafford General Hospital	North West		
98K319	F	16/10/98	27/11/97	6-11 months	Leeds Public Health Laboratory	Northern and Yorkshire		
98K320	M	13/10/98	12/01/95	1-4 years	Alder Hey County Hospital	North West		
98K321	F	13/10/98	09/02/93	5+ years	South Tyneside District Hospital	Northern and Yorkshire		
98K322	M	31/10/98	07/02/69	5+ years	Aberystwyth Bronglais Hospital	Wales		
98K323	F	31/10/98	25/02/97	1-4 years	Aberystwyth Bronglais Hospital	Wales		
98K325	F	26/10/98	26/03/93	5+ years	Oxford Public Health Laboratory	Anglia and Oxford		
98K327	M	26/10/98	22/05/98	2-5 months	Portsmouth Public Health Laboratory	South and West		
98K328	F	02/11/98	06/08/98	2-5 months	Salisbury Public Health Laboratory	South and West		
98K329	M	02/11/98	29/08/98	2-5 months	Chertsey ST. Peters Hospital	South Thames		
98K331	F	30/10/98	31/07/98	2-5 months	Poole Public Health Laboratory	South and West		
98K332	F	06/11/98	03/10/98	< 2 months	Bristol Public Health Laboratory	South and West	Possibly deaf, ventilated	
98K333	M	03/11/98	25/09/98	< 2 months	Keighly Airdale Hospital	Northern and Yorkshire		
98K335	F	16/11/98	24/07/98	2-5 months	Poole Public Health Laboratory	South and West		

Isolate	Sex	Date of isolation	Date of birth	Age group	Laboratory	Region	Other notes	Reasons for exclusion of thesis
98K336	F	12/11/98	26/06/87	5+ years	Manchester Booth Hall Hospital	North West	Had cough for three weks	
98K337	M	19/11/98	08/09/98	2-5 months	Manchester Booth Hall Hospital	North West		
98K338	M	16/11/98	11/03/92	5+ years	Manchester Booth Hall Hospital	North West	Persistent cough admitted diagnosis	
98K340	F	18/11/98	23/10/98	< 2 months	Whiston Hospital	West Midlands		
98K351	F	10/12/98	27/10/98	< 2 months	Bristol Southmead Hospital	South and West		
98K352	F	16/12/98	25/10/98	< 2 months	Bristol Public Health Laboratory	South and West	Baby admitted with pneumonia; died after two days	
99K001	M	02/12/98	16/10/98	< 2 months	Cardiff Public Health Laboratory	Wales		
99K002	F	17/12/98	10/11/98	< 2 months	Truro Public Health Laboratory	South and West	Persistent cough admitted diagnosis in hospital	
99K003	F	03/12/98	28/11/92	5+ years	Leeds Public Health Laboratory	Northern and Yorkshire		
99K004	M	17/12/98	11/10/98	2-5 months	Nottingham Public Health Laboratory	Trent		
99K005	F	02/12/98	16/08/98	2-5 months	Margate Queen Elizabeth the Queen Mother Hospital	South and West		
99K006	F	15/12/98	09/10/98	2-5 months	Leeds General Infirmary	Northern and Yorkshire		
99K008	F	14/12/98	06/07/87	5+ years	Merthyr Tydfil Prince Charles Hospital	Wales		
99K009	F		25/10/98		Bristol Public Health Laboratory			contaminated culture
99K010	F	08/12/98	27/10/98	< 2 months	Southampton Public Health Laboratory	South and West	Tachypnea with occasional apnoea; progressed ventilatory failure and death	
99K011		06/12/98			Dumfries Royal Infirmary			
99K012		15/12/98	17/10/98	< 2 months	Altnagelvin Area			
99K013	F	18/12/98	30/10/98	< 2 months	Chester Public Health Laboratory	Trent		
99K014	F	11/12/98	25/09/98	2-5 months	Stoke on Trent Public Health Laboratory	Trent		
99K015		03/12/98	01/11/98	< 2 months	Altnagelvin Area			
99K016	F	31/12/98	09/07/98	2-5 months	Altnagelvin Area			
99K017	F	07/01/99	02/09/98	2-5 months	East Glamorgan Hospital	Wales		
99K018	F	23/12/98	22/01/95	1-4 years	Leeds Public Health Laboratory	Northern and Yorkshire		
99K020	M	07/01/99	08/12/93	5+ years	Manchester Booth Hall Hospital	North West		
99K021	M	12/01/99	03/10/98	2-5 months	Middlesbrough Public Health Laboratory	Northern and Yorkshire		
99K022	F	19/01/99	16/11/98	2-5 months	Crawley Hospital	South Thames		
99K023	F	31/01/99	25/12/98	< 2 months	Poole Public Health Laboratory	South and West		
99K024	F	21/01/99	25/12/98	< 2 months	Poole Public Health Laboratory	South and West		
99K025	M	18/01/99	11/02/94	1-4 years	Gloucester Public Health Hospital	South and West		

Isolate	Sex	Date of isolation	Date of birth	Age group	Laboratory	Region	Other notes	Reasons for exclusion of thesis
99K027	M	26/01/99	10/09/98	2-5 months	Alder Hey County Hospital	North West		
99K028	F	01/03/99	14/05/98	6-11 months	Gateshead Queen Elizabeth Hospital	Northern and Yorkshire		
99K030	M	29/01/99	22/10/88	5+ years	Lancaster Royal Infirmary	North West		
99K031	F	01/02/99	12/08/98	2-5 months	Carlisle Public Health Laboratory	Northern and Yorkshire		
99K032	F	15/01/99	10/11/98	2-5 months	Peterborough Public Health Laboratory	Anglia and Oxford		
99K033	M	04/01/99	31/03/96	1-4 years	Swansea Public Health Laboratory	Wales	Croupy cough with some features of paroxysms	
99K034	F	19/01/99	24/11/98	< 2 months	Bristol Public Health Laboratory	South and West		
99K035	M	02/02/99	21/12/98	< 2 months	Portsmouth Public Health Laboratory	South and West		
99K036	M	08/02/99	28/12/98	< 2 months	Haverfordwest Withybush General Hospital	Wales		
99K037	M	02/02/99	03/09/92	5+ years	Swansea Singleton Hospital	Wales		
99K038	F	04/02/99	24/12/98	< 2 months	Cambridge Public Health Laboratory	Anglia and Oxford		
99K039	F	18/02/99	26/10/90	5+ years	Furness General Hospital	Northern and Yorkshire		
99K040	M	10/02/99	06/12/98	2-5 months	Boston Pilgrim Hospital	Trent		
99K041	F	19/02/99	21/12/98	< 2 months	Epsom Public Health Laboratory	South Thames		
99K042	M	18/02/99	14/11/98	2-5 months	Exeter Public Health Laboratory	South and West	Chest infection with cough	
99K043					Dumfries Royal Infirmary			
99K044	F	13/02/99	27/09/98	2-5 months	Portsmouth Public Health Laboratory	South and West		
99K045	M	24/02/99	08/01/99	< 2 months	Alder Hey County Hospital	North West		
99K046	M	01/03/99	28/01/99	< 2 months	Manchester Booth Hall Hospital	North West		
99K047	F	20/02/99	17/03/90	5+ years	Ipswich Public Health Laboratory	Anglia and Oxford		
99K048	M	02/03/99	12/02/68	5+ years	Epsom Public Health Laboratory	South Thames	Aplastic anaemia and bone marrow transplant	
99K049	F	25/02/99	28/01/99	< 2 months	Salisbury Public Health Laboratory	South and West		
99K050	M	05/02/99	19/12/98	< 2 months	Nottingham Public Health Laboratory	Trent	Admit hospital with Hirschsprung; bronchiolitis with secondary consolidation	
99K051	F	21/02/99	30/09/98	2-5 months	Alder Hey County Hospital	North West	Admitted with croup, cough and stridor	
99K052	F				Swansea Public Health Laboratory			contaminated culture
99K053	F	26/02/99	15/12/98	2-5 months	Swansea Singleton Hospital	Wales		
99K055	M	22/02/99	29/01/99	< 2 months	Bishops Stortford Herts and Essex Hospital	North Thames	Two weeks admitted to hospital, persistent cough and respiratory cough	
99K057	F	09/03/99	19/02/98	1-4 years	Westmorland General Hospital	Northern and Yorkshire		
99K058	M	08/03/99	29/11/98	2-5 months	Norwich Public Health Laboratory	Anglia and Oxford	29Dec98 upper respiratory tract infection, snuffles 3Feb99, snuff 5Mar 99 diarrhoea, 6Mar99 admitted	

Isolate	Sex	Date of isolation	Date of birth	Age group	Laboratory	Region	Other notes	Reasons for exclusion of thesis
99K059	M	04/03/99	30/05/92	5+ years	Preston Hall Hospital			*B.parapertussis*
99K060		03/03/99	18/12/98	2-5 months	Southmead Hospital			*B.parapertussis*
99K061	M	03/03/99	23/12/92	5+ years	Taunton Public Health Laboratory	South and West		
99K062	M	05/03/99	28/02/91	5+ years	Gloucester Public Health Hospital	South and West		
99K063	F	05/03/99	27/05/86	5+ years	Gloucester Public Health Hospital			
99K064	M	23/02/99	27/09/82	5+ years	Bodelwyddan Rhyl Public Health Laboratory	Wales	Continuing cough phlegm apyrexial	
99K065	M	01/03/99	15/02/94	5+ years	Conquest Hospital	South Thames		
99K070	F	17/03/99	26/03/88	5+ years	Cardiff Public Health Laboratory	Wales		
99K071	F	07/03/99	21/02/96	1-4 years	Swindon Princess Margaret Hospital	South and West		
99K072	F	16/03/99	10/01/99	2-5 months	Swindon Princess Margaret Hospital	South and West		
99K073	M	17/03/99	15/02/60	5+ years	Nottingham Public Health Laboratory	Trent		
99K077	F	15/03/99	21/04/95	1-4 years	Bristol Royal Hospital	South and West		
99K078	F	24/03/99	05/02/99	< 2 months	Furness General Hospital	Trent		
99K079	M	04/06/99	23/12/53	5+ years	Nottingham Public Health Laboratory	Trent		
W9025245	F	>2/9/99	27/08/99		Whittington Hospital			
PICU278	M	16/08/99	22/04/99	2-5 months	RSIL		Baby	
PICU475	F	04/10/99	30/04/99	6-11 months	RSIL		Baby	
PICU519	M	03/09/99	17/05/64	5+ years	RSIL		Father to PICU522	
PICU522	F	03/09/99	01/07/99	2-5 months	RSIL		Baby of PICU519	
PICU523	F	03/09/99	01/07/88	5+ years	RSIL		Sister of PICU522	
PICU524	M	03/09/99	07/02/97	1-4 years	RSIL		Brother of PICU522	
PICU643	M	2000					Father of PICU 644	
PICU644	F	2000					Baby	
PICU666	F	2000					Baby	
PICU705	F	2000					Baby	
PICU708	F	2000					Sister of PICU 705	
Bp 245	F	2001	28/07/26		North Tees General Hospital		Isolated from blood culture from immunocompromised F with non-Hodgkins lymphoma	
Bp 436		2002			Glamorgan Royal Hospital			
Bp 442	M	2002		2-5 months	Farnborough Public Health Laboratory			
Bp 463	F	2002	15/02/95		Nobles Isle of Man Hospital			
Bp 568	F	2002	13/08/02		Whittington Hospital			

Isolate	Sex	Date of isolation	Date of birth	Age group	Laboratory	Region	Other notes	Reasons for exclusion of thesis
Bp 584	F	2002	13/05/64		Guernsey Princess Elizabeth Hospital			
Bp 602	F	2002	15/03/02		Whittington Hospital			
Bp 609	M	2002	05/03/02		Stoke on Trent Public Health Laboratory			
Bp 610	M	2002	02/03/02		Grantham Hospital			
Bp 618		2002					Internal Quality specimen	
Bp 632	F	2002	23/01/02		Exeter Public Health Laboratory			
Bp 637	F	2002	14/03/02		Birmingham Public Health Laboratory			
Bp 663	M	2002	03/08/95		Leicester Public Health Laboratory		Leicester outbreak	
Bp 704	F	2002			Nottingham Public Health Laboratory		Keyworth medical practice, Nottingham	
Bp 705	F	2002			Nottingham Public Health Laboratory		Keyworth medical practice, Nottingham	
Bp 706	M	2002	10/07/89		Nottingham Public Health Laboratory		Keyworth medical practice, Nottingham	
Bp 707	F	2002	10/12/93		Nottingham Public Health Laboratory		Keyworth medical practice, Nottingham	
Bp 708	F	2002			Nottingham Public Health Laboratory		Keyworth medical practice, Nottingham	
Bp 709	M	2002			Nottingham Public Health Laboratory		Keyworth medical practice, Nottingham	
Bp 710	M	2002	01/02/59		Nottingham Public Health Laboratory		Keyworth medical practice, Nottingham	
Bp 711	M	2002	06/07/95		Nottingham Public Health Laboratory		Keyworth medical practice, Nottingham	
Bp 712	M	2002			Nottingham Public Health Laboratory		Keyworth medical practice, Nottingham	
Bp 713	F	2002			Nottingham Public Health Laboratory		Keyworth medical practice, Nottingham	

Additional enhanced information on selected isolates from September 1998 to April 1999

Isolate	Typically clinical symptoms	Complications encountered	Conjunctivitis	Pneumonia	Apnoea attacks	Convulsions	Admitted to hospital	Vaccinated?	Number of doses	Died	Serotype
98K290	YES	YES	NO	YES	NO	NO	NO	YES	3	U	1,3
98K299	YES	NO	NO	NO	NO	NO	YES	YES	1	U	1,3
98K300	NO	NO	NO	NO	NO	NO	YES	NO	0	U	1,2
98K301	NK	NK	NK	NK	NK	NK	YES	YES	1	U	1,2
98K302	NO	NO	NO	NO	NO	NO	NO	YES	1	U	1,2
98K304	YES	NO	NO	NO	NO	NO	NO	YES	3	U	1,2
98K305	YES	YES	NO	NO	YES	NO	NO	YES	3	U	1,3
98K306	YES	NO	NO	NO	NO	NO	YES	YES	1	U	1,2
98K310	YES	NO	NO	NO	NO	NO	NO	NO	0	U	1,2
98K311	YES	NO	NO	NO	NO	NO	NO	NO	0	U	1,3
98K313	YES	YES	NO	NO	NO	NO	NK	YES	3	U	1,2
98K315	YES	NK	NK	NK	NK	NK	NO	NO	0	U	1,3
98K316	NO	YES	NO	NO	YES	NO	YES	NO	0	U	1,2
98K317	YES	YES	NO	NO	YES	NO	NK	NO	0	U	1,2
98K318	YES	NO	NO	NO	NO	NO	NO	YES	3	U	1,3
98K319	YES	NO	NO	NO	NO	NO	NO	NO	0	U	1,3
98K320	NK	NK	NK	NK	NK	NK	NK	NK	NK		1,3
98K321	YES	NK	NK	NK	NK	NK	YES	YES	3	U	1,2
98K322	YES	NK	NK	NK	NK	NK	NO	YES	NK	U	1,2
98K323	YES	NO	NO	NO	NO	NO	YES	YES	3	U	1,2
98K325	NO	YES	NO	YES	NO	NO	NO	YES	3	U	1,3
98K327	NK	NO	NO	NO	NO	NO	YES	YES	1	U	1,2
98K328	YES	YES	NO	YES	YES	NO	YES	YES	1	U	1,3
98K329	YES	YES	NO	NO	YES	NO	YES	NO	0	U	1,2
98K331	YES	NO	NO	NO	NO	NO	YES	YES	1	U	1,3
98K332	YES	YES	NO	NO	YES	NO	YES	NO	0		1,2
98K333	YES	NO	NO	NO	NO	NO	YES	NO	0	U	1,2
98K335	YES	NK	NK	NK	NK	NK	YES	YES	1	U	1,3
98K336	NK	NO	NO	NO	NO	NO	NK	NK	NK	U	1,2
98K337	YES	YES	NO	NO	YES	NO	YES	NO	0	Y	1,3

Isolate	Typically clinical symptoms	Complications encountered	Conjunctivitis	Pneumonia	Apnoea attacks	Convulsions	Admitted to hospital	Vaccinated?	Number of doses	Died	Serotype
98K338	NO	NO	NO	NO	NO	NO	NO	YES	3	U	1,3
98K340	YES	YES	NO	YES	YES	NO	YES	NO	0	U	1,2
98K351	YES	NO	NO	NO	NO	NO	YES	NO	0	U	1,2
98K352	NO	YES	NO	YES	YES	NO	YES	NO	0	Y	1,2
99K001	YES	NO	NO	NO	NO	NO	YES	NO	0	U	1,2
99K002	YES	YES	NO	NO	YES	NO	YES	NO	0	U	1,2
99K003	YES	NO	NO	NO	NO	NO	NO	YES	3	U	1,2
99K004	YES	NO	NO	NO	NO	NO	YES	YES	1	U	1,3
99K005	YES	NK	NK	NK	NK	NK	YES	YES	2	U	1,2
99K006	YES	NO	NO	NO	NO	NO	YES	NO	0	U	1,2
99K008	YES	NO	NO	NO	NO	NO	YES	NO	0	U	1,3
99K010	NO	YES	NO	YES	YES	NO	YES	NO	0	Y	1,2
99K013	YES	YES	NO	NO	YES	NO	YES	NO	0	U	1,3
99K014	YES	NO	NO	NO	NO	NO	YES	YES	1	U	1,3
99K017	YES	NO	NO	NO	NO	NO	YES	NO	0	U	1,2
99K018	YES	NO	NO	NO	NO	NO	NO	NO	0	U	1,3
99K020	YES	NO	NO	NO	NO	NO	YES	YES	3	U	1,2
99K021	YES	NO	NO	NO	NO	NO	YES	YES	1	U	1,2
99K022		NK	NK	NK	NK	NK	NK	NK	NK		1,3
99K023	YES	NO	NO	NO	NO	NO	YES	NO	0	U	1,3
99K024	YES	NO	NO	NO	NO	NO	YES	NO	0	U	1,3
99K025	YES	NO	NO	NO	NO	NO	NK	YES	3	U	1,3
99K027	YES	NO	NO	NO	NO	NO	NO	YES	2	U	1,3
99K028	YES	NO	NO	NO	NO	NO	YES	NO	0	U	1,2
99K030	YES	YES	YES	NO	NO	NO	YES	NO	0	U	1,2
99K031	YES	NO	NO	NO	NO	NO	YES	YES	3	U	1,2
99K032	YES	NO	NO	NO	NO	NO	NK	NO	0	U	1,2
99K033	NO	NO	NO	NO	NO	NO	NO	YES	3	U	1,3
99K034	YES	NO	NO	NO	NO	NO	NK	NO	0	U	1,2
99K035	YES	YES	NO	NO	YES	NO	YES	NO	0	U	1,2
99K036	YES	NO	NO	NO	NO	NO	YES	NO	0	U	1,2

Isolate	Typically clinical symptoms	Complications encountered	Conjunctivitis	Pneumonia	Apnoea attacks	Convulsions	Admitted to hospital	Vaccinated?	Number of doses	Died	Serotype
99K037	YES	NO	NO	NO	NO	NO	NO	NO	0	U	1,3
99K038	NK	YES	NO	YES	NO	NO	YES	NO	0	U	1,2
99K039	YES	NO	NO	NO	NO	NO	NO	YES	4	U	1,3
99K040	YES	NO	NO	NO	NO	NO	YES	NO	0	U	1,2
99K041	YES	NO	NO	NO	NO	NO	YES	NO	0	U	1,2
99K042	NO	NO	NO	NO	NO	NO	YES	YES	1	U	1,3
99K044	YES	NO	NO	NO	NO	NO	YES	YES	3	U	1,2
99K045	YES	YES	NO	NO	YES	NO	YES	NO	0	U	1,3
99K046	YES	NO	NO	NO	NO	NO	YES	NO	0	U	1,2
99K047	YES	NO	NO	NO	NO	NO	NO	NO	0	U	1,3
99K048	YES	NO	NO	NO	NO	NO	NO	YES	3	U	1,2
99K049	YES	NO	NO	NO	NO	NO	YES	NO	0	U	1,3
99K050	NO	YES	NO	YES	NO	NO	YES	NO	0	U	1,2
99K051	NO	NO	NO	NO	NO	NO	YES	YES	1	U	1,2
99K053	YES	NK	NK	NK	NK	NK	YES	NO	0	U	1,2
99K055	NO	YES	NO	YES	YES	NO	YES	NO	0	U	1,3
99K057	YES	NO	NO	NO	NO	NO	NO	YES	3		1,2
99K058	NO	NO	NO	NO	NO	NO	YES	YES	2	U	1,3
99K061	YES	NO	NO	NO	NO	NO	NO	NO	0	U	1,2
99K062	NK	NO	NO	NO	NO	NO	NO	NO	0	U	1,3
99K064	NO	NO	NO	NO	NO	NO	NO	NO	0	U	1,2
99K065	YES	NO	NO	NO	NO	NO	NK	NO	0	U	1,2
99K070	NO	NO	NO	NO	NO	NO	YES	YES	3	U	1,3
99K071	YES	NO	NO	NO	NO	NO	YES	NO	0	U	1,3
99K072	YES	YES	NO	NO	NO	NO	YES	YES	1	U	1,3
99K073	YES	NO	NO	NO	NO	NO	NO	NK	NK	U	1,3
99K077	YES	NO	NO	NO	NO	NO	NK	YES	3	U	1,3
99K078	YES	YES	NO	NO	YES	NO	YES	NO	0	U	1,3
99K079	YES	NO	NO	NO	NO	NO	NO	NK	NK	U	1,3

NK=Not known

Part C. *Bordetella pertussis* vaccine and reference strains (see 2.1.3, page 57)

	Source Culture collection	Strain	Other strain references	Further details	Earliest year known
Vaccine strains					
	CN	2992B	24503	Same strain as CN2992, Dr Holt, CPHL, Colindale	1949
	CN	3099	252E/7377	Dr Ungar, Glaxo Laboratories	1950
	NIBSC (originally CN)	5476	38098	Hither Green Hospital via Mr G Turner	1963
	NIBSC	Tohama-I		Lab Isolates	1981
	Aventis	10536			
B.pertussis type strain	NCTC	10739^T	18323^T	M Pittman, NIH Maryland/P Kendrick/Redeposition of now discarded NCTC9797/Potency test of pertussis vaccine	1970
Reference strains					
	RIVM	B410		IS1002 standard, as used in Van der Zee (1997)	
	RIVM	B665		IS1002 standard, type dt-29 (RFLP-1)	
	RIVM	B1000	APV110	PFGE standard	
	RIVM	B994	APV1561	PFGE standard	
	RIVM	B1073	Fr56	PFGE standard	
	RIVM	B5	134	PFGE standard, group III	
	RIVM	B902	RIVM	PFGE standard, group IV alpha	
	RIVM	B1121	18323	PFGE standard	
	RIVM	B1127	Fr287	PFGE standard, group V	
	Institut Pasteur	Fr743		PFGE standard	

^{T}=Type strain

Part D. *Bordetella pertussis* isolates from Missouri, USA (see 2.1.4, page 58)

Strain	Sex	Age	Origin of strain	Pertussis vaccination history	Prophylactic history	Date received at MO Lab	Additional information
MO-482						25/01/1984	
MO-18						20/04/1984	
MO-814						14/05/1984	
MO-20						20/09/1984	
MO-1377						29/05/1986	
MO-34						23/07/1986	
MO-229						08/09/1986	
MO-919						02/02/1987	
MO-46						06/08/1987	
MO-704						04/02/1988	
MO-981						07/04/1988	
MO-620						02/09/1988	
MO-908						11/12/1988	
MO-121	M	6 months	West Plains, MO			20/07/1989	
MO-404	F	1 months	St. Louis, MO			05/09/1989	Unsure of details
MO-144	F	6 months	Smithville, MO			13/08/1996	
MO-87	M	2 months	Bonne Terre, MO			24/07/1997	
MO-1439	F	5 months	Kansas City, MO	up to date	unknown	03/06/1999	Kansas City "outbreak"
MO-1440	M	3 months	Kansas City, MO	1st dose DTP (?) 04/05/99	unknown	03/06/1999	Kansas City "outbreak"
MO-1459	M	6 months	Kansas City, MO	1 or 2 DTP doses	Albuterol	08/06/1999	Kansas City "outbreak"
MO-1460	F	3 months	Kansas City, MO	up to date	none	08/06/1999	Kansas City "outbreak"
MO-1482	M	2 months	Kansas City, MO	up to date	none	14/06/1999	Kansas City "outbreak"
MO-1484	M	10 months	Kansas City, MO	not up to date	Erythromycin	14/06/1999	Kansas City "outbreak"
MO-1503	F	3 months	Dixon, MO	1st dose DTP 28/05/99	none	21/06/1999	
MO-1532	M	4 months	Lohman, MO	up to date	Albuterol	29/06/1999	
MO-3	F	3 months	Kansas City, MO	1st dose DTP	none	01/07/1999	Kansas City "outbreak"
MO-9	F	3 months	Mexico, MO	1st dose DTaP 13/05/99	Amoxicillin	06/07/1999	
MO-14	F	3 months	Springfield, MO	unknown	none	07/07/1999	
MO-73	M	1 month	Kansas City, MO	up to date	Ampicillin, Gentamycin, Erythromicin	22/07/1999	Kansas City "outbreak"
MO-74	M	3 months	Kansas City, MO	up to date	none	22/07/1999	Kansas City "outbreak"

Strain	Sex	Age	Origin of strain	Pertussis vaccination history	Prophylactic history	Date received at MO Lab	Additional information
MO-80	F	1 month	Kansas City, MO		none	23/07/1999	Kansas City "outbreak"
MO-83	M	1 year	Jefferson City, MO	3 doses DTP; 18/05/98, 14/08/98, 06/01/99	none	23/07/1999	
MO-84	F	1 month	St. Louis, MO		unknown	23/07/1999	
MO-99	F	1 month	Alton, Illinois		none	27/07/1999	
MO-167	M	5 months	Moberley, MO	unknown	Amoxicillin for 1 week	11/08/1999	
MO-234	M	5 months	Kansas City, MO	2 doses DTaP; 17/05/99, 06/07/99	unknown	31/08/1999	Strain is from Kansas City and isolation date is within 5-21 days incubation period.
MO-21			Anderson, MO			05/01/2000	
MO-7			Blue Springs, MO			01/03/2000	
MO-301			Taylor, MO			29/03/2000	
MO-388			Joplin, MO			18/05/2000	
MO-452			Kansas City, MO			23/06/2000	
MO-495			St. Louis, MO			11/07/2000	
MO-592			Springfield, MO			16/08/2000	
MO-769			Jefferson City, MO			18/10/2000	Omitted from thesis as strain is *B.parapertussis*
MO-905			Rolla, MO			29/11/2000	
MO-96			Pevely, MO			25/01/2001	

Part E. *Bordetella parapertussis* strains (see 2.1.5, page 58)

Source Culture collection	Strain	Other strain references	Host and site of isolation	Depositor details	Earliest year known	References
NCTC	10853	G Eldering 508; Eldering 344; ATCC 15237	Man, cough plate	ATCC, USA in 1972	1938	{Eldering & Kendrick 1938 ID: 3791}
RIVM	B24			used for IS1002 RFLP probe		{van Loo, van der Heide, et al. 1999 ID: 94}

12.4 Appendix 4. Enhanced surveillance forms used by the Communicable Disease Surveillance Centre

SURVEILLANCE OF LABORATORY CONFIRMED PERTUSSIS INFECTION IN ENGLAND AND WALES

STRICTLY CONFIDENTIAL Form 1 (CDR) – last revision date Dec 1996

RESPONSIBLE CENTRES FOR LABORATORY CONFIRMED PERTUSSIS INFECTION IN ENGLAND AND WALES:
PHLS Communicable Disease Surveillance Centre, Immunisation Division
University of Manchester, Pertussis Reference Laboratory

OBJECTIVES:
- To obtain epidemiological and microbiological information, including serotyping, on all cases of laboratory confirmed pertussis in E&W
- To monitor the impact of pertussis vaccination in E&W, using laboratory, notification and vaccine coverage data

For CDSC/Ref. Lab. use only.
Study no.
Reporting Laboratory
Date of CDR report ____/____/19____
Date specimen taken ____/____/19____
Serotype

Please complete as far as possible, ticking appropriate boxes

REPORTING DETAILS

Was the isolate sent for serotyping? Yes ☐ No ☐

PERSONAL DETAILS

Name of patient: Surname ____________ First name ____________ **Sex:** M ☐ F ☐

Date of birth ____/____/____ **Age** ________ **If GP not known: name of clinician in charge of case**

Name of GP ____________ ____________

Address ____________ **Address** ____________

____________ Postcode ________ ____________ Postcode ________

CLINICAL HISTORY OF CASE

Did the patient have clinically typical pertussis? Yes ☐ No ☐ Not known ☐

Please add any relevant information (e.g. severity, treatment, complications) ____________

VACCINATION HISTORY OF CASE

At the time of diagnosis, was the patient vaccinated against pertussis? Yes ☐ No ☐ Not known ☐

If yes, number of doses ________

Completed by (please print) ____________

Position held ____________ Signature ____________ Date ____/____/____

PLEASE RETURN THIS FORM TO:
Dr Elizabeth Miller, Immunisation Division
PHLS Communicable Disease Surveillance Centre, 61 Colindale Avenue, London NW9 5EQ
Tel: 0181-200 6868 Fax: 0181-200 7868

PUBLIC HEALTH LABORATORY SERVICE
PROTECTING THE POPULATION FROM INFECTION

SURVEILLANCE OF LABORATORY CONFIRMED PERTUSSIS INFECTION IN ENGLAND AND WALES

STRICTLY CONFIDENTIAL Form 1 (Ref lab) – last revision date Dec 1996

RESPONSIBLE CENTRES FOR LABORATORY CONFIRMED PERTUSSIS INFECTION IN ENGLAND AND WALES:
PHLS Communicable Disease Surveillance Centre, Immunisation Division
University of Manchester, Pertussis Reference Laboratory

OBJECTIVES:
- To obtain epidemiological and microbiological information, including serotyping, on all cases of laboratory confirmed pertussis in E&W
- To monitor the impact of pertussis vaccination in E&W, using laboratory, notification and vaccine coverage data

For CDSC/Ref. Lab. use only.

Study no. ______

Reporting Laboratory ______

Date of CDR report ____/____/19____

Date specimen taken ____/____/19____

Serotype ______

Please complete as far as possible, ticking appropriate boxes

REPORTING DETAILS

Was the isolate reported to the CDR? Yes ☐ No ☐

PERSONAL DETAILS

Name of patient: Surname ______ First name ______ **Sex:** M ☐ F ☐

Date of birth ____/____/____ **Age** ______

Name of GP ______

Address ______

______

______ Postcode ______

If GP not known: name of clinician in charge of case

______

Address ______

______

______ Postcode ______

CLINICAL HISTORY OF CASE

Did the patient have clinically typical pertussis? Yes ☐ No ☐ Not known ☐

Please add any relevant information (e.g. severity, treatment, complications) ______

______

______

VACCINATION HISTORY OF CASE

At the time of diagnosis, was the patient vaccinated against pertussis? Yes ☐ No ☐ Not known ☐

If yes, number of doses ______

Completed by (please print) ______

Position held ______ Signature ______ Date ____/____/____

PLEASE RETURN THIS FORM TO:
Dr Elizabeth Miller, Immunisation Division
PHLS Communicable Disease Surveillance Centre, 61 Colindale Avenue, London NW9 5EQ
Tel: 0181 200 6868 Fax: 0181 200 7868

PUBLIC HEALTH LABORATORY SERVICE
PROTECTING THE POPULATION FROM INFECTION

SURVEILLANCE OF LABORATORY CONFIRMED PERTUSSIS INFECTION IN ENGLAND AND WALES

STRICTLY CONFIDENTIAL Form 2 (Clinical) - last revision date Dec 1996

RESPONSIBLE CENTRES FOR LABORATORY CONFIRMED PERTUSSIS INFECTION IN ENGLAND AND WALES:
PHLS Communicable Disease Surveillance Centre, Immunisation Division
University of Manchester, Pertussis Reference Laboratory

OBJECTIVES:
- To obtain epidemiological and microbiological information, including serotyping, on all cases of laboratory confirmed pertussis in E&W
- To monitor the impact of pertussis vaccination in E&W, using laboratory, notification and vaccine coverage data

For CDSC use only.
Study no. ______
Date B. pertussis isolated ____/____/19____

Please complete as far as possible, ticking appropriate boxes

PERSONAL DETAILS

Name of patient: Surname ______________ First name ______________ **Sex:** M ☐ F ☐

Date of birth ____/____/____ **Age** ______

CLINICAL HISTORY OF CASE

Did the patient have clinically typical pertussis? Yes ☐ No ☐

If no describe clinical presentation: ______________________________

____

Please indicate if there were any of the following complications

None ☐ Pneumonia ☐ Convulsions ☐ Conjunctival haemorrhage ☐ Apnoeic attacks ☐

Other (describe) ______________________________

Did the patient receive erythromycin? Yes ☐ No ☐

If yes, was this: for prevention ☐ Duration of course: ______ days Date started ____/____/____

for treatment ☐ Duration of course: ______ days Date started ____/____/____

Was the patient admitted to Hospital? Yes ☐ No ☐

VACCINATION HISTORY OF CASE

Has the patient been vaccinated against pertussis? Yes ☐ No ☐

If yes, how many doses were given **PRIOR** to onset of pertussis? ______

Dates of vaccination (if available) 1st dose ____/____/____ 2nd dose ____/____/____ 3rd dose ____/____/____

PLEASE GIVE NAME e.g. PENTAVAC OR INFANRIX OR BATCH NUMBER OF ANY DOSE GIVEN AFTER 1st Nov 1999

SOURCE OF INFECTION

In the month before onset, did this case have contact with a known or suspected case of pertussis? Yes ☐ No ☐

Not known ☐ If yes, please specify where and age of contact ______________________________

Please add any additional comments on the reverse of the form

Completed by (please print) ______________________________

Position held ______________ Signature ______________ Date ____/____/____

PLEASE RETURN THIS FORM TO:
Dr Elizabeth Miller, Immunisation Division
PHLS Communicable Disease Surveillance Centre, 61 Colindale Avenue, London NW9 5EQ
Tel: 0181-200 6868 Fax: 0181-200 7868

PUBLIC HEALTH LABORATORY SERVICE
PROTECTING THE POPULATION FROM INFECTION