

**Comparative genomics to investigate
genome function and adaptations in the
newly sequenced *Brachyspira*
hyodysenteriae and *Brachyspira pilosicoli***

Phatthanaphong Wanchanthuek

A Thesis presented for the degree of
Doctor of Philosophy



Murdoch University

Australia

March 2009

Dedication

I would like to dedicate this dissertation to my family
and my fiancée Ratchaneekorn.

Comparative genomics to investigate
genome function and adaptations in the
newly sequenced *Brachyspira hyodysenteriae*
and *Brachyspira pilosicoli*

Phatthanaphong Wanchanthuek

Submitted for the degree of Doctor of Philosophy

March 2009

ABSTRACT

Brachyspira hyodysenteriae and *Brachyspira pilosicoli* are anaerobic intestinal spirochaetes that are the aetiological agents of swine dysentery and intestinal spirochaetosis, respectively. As part of this PhD study the genome sequence of *B. hyodysenteriae* strain WA1 and a near complete sequence of *B. pilosicoli* strain 95/1000 were obtained, and subjected to comparative genomic analysis. The *B. hyodysenteriae* genome consisted of a circular 3.0 Mb chromosome, and a 35,940 bp circular plasmid that has not previously been described. The incomplete genome of *B. pilosicoli* contained 4 scaffolds. There were 2,652 and 2,297 predicted ORFs in the *B. hyodysenteriae* and *B. pilosicoli* strains, respectively. Of the predicted ORFs, more had similarities to proteins of the enteric *Clostridium* species than they did to proteins of other spirochaetes. Many of these genes were associated with transport and metabolism, and they may have been gradually acquired through horizontal gene transfer in the environment of the large intestine.

A construction of central metabolic pathways of the *Brachyspira* species identified a complete set of coding sequences for glycolysis, gluconeogenesis, a non-oxidative pentose phosphate pathway, nucleotide metabolism and a respiratory electron transport chain. A notable finding was the presence of *rfb* genes on the *B. hyodysenteriae* plasmid, and their apparent absence from *B. pilosicoli*. As these genes are involved in rhamnose biosynthesis it is likely that the composition of the *B. hyodysenteriae* lipooligosaccharide O-sugars is different from that of *B. pilosicoli*. O-antigen differences in these related species could be associated with differences in their specific niches, and/or with their disease specificity. Overall, comparison of *B. hyodysenteriae* and *B. pilosicoli* protein content and analysis of their central metabolic pathways showed that they have

diverged markedly from other spirochaetes in the process of adapting to their habitat in the large intestine.

The presence of overlapping genes in the two *Brachyspira* species and in other spirochaete species also was investigated to determine their functional role, if any. The number of overlapping genes in the 12 spirochaete genomes examined ranged from 11-45%. Of these, 80% were unidirectional. Overlapping genes were found non-uniform distributed within the *Brachyspira* genomes such that 70-80% of them occurred on the same strand (unidirectional, $\rightarrow\rightarrow/\leftarrow\leftarrow$), with 16-28% occurring on opposite DNA strands (divergent, $\leftarrow\rightarrow$). The remaining 4-6% of overlapping genes were convergent ($\rightarrow\leftarrow$). The majority of the unidirectional overlap regions were relatively short, with >50% of the total observations overlapping by >4 bp. A small number of overlapping gene pairs was duplicated within each genome and there were some triplet overlapping gene pairs. Unique orthologous overlapping gene pairs were identified within the various spirochaete genera. Over 75% of the overlapping genes in the *Brachyspira* species were in the same or related metabolic pathway. This finding suggests that overlapping genes are not only likely to be the result of functional constraints but also are constrained from a metabolomic context. Of the remaining 25% overlapping genes, 50% contained one hypothetical gene with unknown function. In addition, in one of the orthologous overlapping gene pairs in the *Brachyspira* species, a promoter was shared, indicating the presence of a novel class of overlapping gene operon in these intestinal spirochaetes.

Declaration

The work in this thesis is based on research carried out at the Centre for Comparative Genomics (CCG), Murdoch University, Australia. I declare that this thesis is my own account of my research and contains as its main content work which has not previously been submitted for a degree at any tertiary institution.

.....

(Phatthanaphong Wanchanthuek)

Acknowledgements

The present dissertation was conducted from 2006 to 2009 at the Centre for Comparative Genomics, Murdoch University. Many people helped tremendously in the completion of this thesis. First, I wish to express my deepest thanks and enormous debt of gratitude, to my supervisors Professor Matthew Bellgard and Professor David Hampson, who have been a wealth of knowledge and support throughout this study. Their understanding, enthusiasm and generosity were endless. As supervisors of a student originally from non-English country, they spent considerable time and effort in improving my English communication skills, which will certainly have great impact on my future research career.

My deepest thanks also go to Dr. Roberto Barrero, Dr. Karon Ryan, Paula Moolhuijzen and Dr. Tom La, whose excellent expertise, patience, kindness and friendship helped me to get through all the difficult stages of the project. Also thanks to the other CCG staff for giving me advice, and for sharing their experience and knowledge.

The project was financed by a grant from the Australian Research Council and Novartis Animal Vaccines (NAV) as the industry partner. Dr. Ian Thompson from NAV was particularly supportive throughout the project. I wish to acknowledge and thank the Royal Thai Government for providing me with a PhD scholarship to undertake this work.

Finally, I owe particular thanks to my parents, my family and my fiancée Ratchaneekorn. They always support me and are a source of pride for me over many years. Their faith, sustenance, understanding and companionship were the sources of my strength to pursue this dream.

List of publications:

1) **Wanchanthuek, P.**, Hampson, D.J., and Bellgard, M. 2009. Analysis of overlapping genes in the newly sequenced of *Brachyspira* and other spirochaetes indicates their likely role in streamlining metabolic pathways (*in preparation*).

2) Bellgard, M*, **Wanchanthuek, P***, La, T., Ryan, K., Moolhuijzen, P., Zlbertyn, A., Shaban, B., Motro, Y., Dunn, D., Schibeci, D., Hunter, A., Barrero, R., Phillips, N., and Hampson, D. 2009. Genome sequence of the pathogenic intestinal spirochete *Brachyspira hyodysenteriae* reveals adaptations to its lifestyle in the porcine large intestine. *PLoS ONE* 4, e4641.

* These authors contributed equally to this work.

3) **Wanchanthuek, P.**, Hampson, D.J. and Bellgard, M. 2008. Conservation and metabolic functional significance of overlapping gene in the bacterial genomes. The 20th Annual Meeting and International Conference of the Thai Society for Biotechnology. October 14-17, 2008 Maha Sarakham, Thailand.

4) **Wanchanthuek, P.**, Ryan, K., Moolhuijzen, P., Albertyn, Z., Shaban, B., La, T., Hampson, D.J. and Bellgard, M. 2008. Comparison of *Brachyspira* central metabolism pathway using the genomic DNA sequence. The International Conference on Genome Informatics. Gold Coast, Australia, 1-3 December 2008.

5) **Wanchanthuek, P.**, Ryan, K., Moolhuijzen, P., Albertyn, Z., Shaban, B., La, T., Hampson, D.J. and Bellgard, M. 2008. A plasmid-borne O-antigen in *Brachyspira hyodysenteriae*. The Seventh Asia Pacific Bioinformatics Conference, Beijing, China, 13-16 January 2009.

Content

Dedication	i
Title page	ii
Abstract	iii
Declaration	v
Acknowledgements	vi
List of publications	vii
Introduction	1
General introduction	2
Thesis outline.....	3
1 Literature review.....	4
1.1 General information on the spirochaetes.....	5
1.2 <i>Brachyspira</i> phylogeny and taxonomy.....	6
1.3 The genus <i>Brachyspira</i>	8
1.3.1 <i>Brachyspira hyodysenteriae</i>	9
1.3.2 <i>Brachyspira pilosicoli</i>	10
1.4 Genome of <i>Brachyspira</i> species.....	11
1.5 Genome sequencing and analysis.....	13
1.5.1 Genome sequencing.....	13
1.5.1.1 Library construction and template preparation	15
1.5.1.2 Automatic high-throughput sequencing.....	15
1.5.1.3 Genome sequence assembly.....	17
1.5.1.3.1 Construction of contigs and scaffolds.....	18
1.5.1.3.2 Generation of consensus sequences.....	19
1.5.1.4 Genome finishing.....	20
1.5.1.5 Scaffold and gap closure.....	21
1.5.2 Genome annotation and analysis.....	21
1.5.2.1 Genome annotation pipeline.....	22
1.5.2.1.1 Structural annotation.....	22
1.5.2.1.2 Functional annotation.....	23
1.5.2.1.2.1 Homology-based prediction of function.....	23
1.5.2.1.2.2 Protein domain and protein localisation....	25

1.5.2.2 Comparative Genomics.....	28
1.5.2.2.1 General genome features of spirochaetes.....	28
1.5.2.2.2 Analysis of genome context.....	31
1.5.2.2.3 Comparative analysis of overlapping gene.....	31
1.5.2.2.4 Evolution of metabolic networks.....	34
1.5.2.3 Vaccine candidates and drug targets.....	37
1.6 Thesis aims and objectives.....	38
2 Assembly and annotation of a complete genome of <i>Brachyspira</i>	40
 <i>hyodysenteriae</i> and an incomplete genome of <i>Brachyspira pilosicoli</i>.	
2.1 Introduction.....	41
2.2 Materials and Methods.....	42
2.2.1 Spirochaete strains.....	42
2.2.2 Genome sequencing.....	42
2.2.3 Genome assembly.....	43
2.2.4 Structural and functional annotation.....	43
2.2.5 Genome map.....	45
2.2.6 Genomic nucleotide skew.....	46
2.3.7 Codon usage.....	46
2.3.8 Ortholog analysis.....	46
2.3.9 Bioinformatics analysis.....	47
2.3 Results.....	48
2.3.1 Genome sequencing progress.....	48
2.3.2 Genomic features.....	50
2.3.2.1 Complete genome sequence of <i>B. hyodysenteriae</i> WA1.....	50
2.3.2.2 Draft genome sequence of <i>B. pilosicoli</i> 95/1000.....	55
2.3.3 Annotation of the <i>Brachyspira</i> genome sequences.....	56
2.3.4 Genome analysis.....	58
2.3.4.1 General genome features of both <i>Brachyspira</i> species.....	58
2.3.4.1.1 Complete genome sequence of <i>B. hyodysenteriae</i>	58
2.3.4.1.2 A draft <i>B. pilosicoli</i> genome sequence.....	62
2.3.4.2 Origin of replication.....	63
2.3.4.3 Genome synteny.....	65
2.3.4.4 Codon usage and amino acid composition of ORFs.....	65
2.3.4.5 Frequency of overlapping genes.....	67

2.3.4.6	Taxonomic distribution of gene homology.....	69
2.3.4.7	The core <i>Brachyspira</i> genes.....	71
2.3.4.8	COG analysis among bacterial genomes.....	72
2.4	Discussion.....	76
2.4.1	Genome assembly.....	76
2.4.2	Genome analysis.....	77
2.4.2.1	Genome size and number of genes.....	77
2.4.2.2	Ribosomal genes.....	79
2.4.2.3	G+C content.....	80
2.4.2.4	Origin of replication.....	81
2.4.2.5	Genome synteny.....	82
2.4.2.6	Gene associated with gene transfer.....	84
2.4.2.7	The core <i>Brachyspira</i> genes.....	85
2.4.2.8	Taxonomy distribution.....	86
2.5	Summary.....	87
3	Genome-based construction of the metabolic pathways of	89
	<i>Brachyspira hyodysenteriae</i> and <i>Brachyspira pilosicoli</i> and	
	comparative analysis with those of other spirochaetes.	
3.1	Introduction.....	90
3.2	Materials and Methods.....	92
3.2.1	Genome sequences.....	92
3.2.2	Prediction and annotation of protein coding sequences.....	92
3.2.3	Metabolic network reconstruction.....	93
3.3	Results.....	94
3.3.1	<i>Brachyspira</i> central metabolic pathways.....	94
3.3.2	Carbohydrate and energy metabolism.....	97
3.3.2.1	Glycolytic pathway.....	97
3.3.2.2	Pentose phosphate pathway.....	101
3.3.3	Nucleotide metabolism.....	102
3.3.3.1	Purine biosynthesis.....	102
3.3.3.2	Pyrimidine biosynthesis.....	106
3.3.3.3	Folate cycle.....	106
3.3.4	Amino acid biosynthesis.....	107
3.3.4.1	Aspartate family.....	110

3.3.4.2	Glutamate family.....	110
3.3.4.2	Other amino acid families.....	112
3.3.5	Lipid metabolism.....	114
3.3.6	Lipooligosaccharide (LOS) biosynthesis.....	117
3.3.6.1	Genes encoding O-antigen component.....	119
3.3.6.2	Genes encoding protein involved in lipid A biosynthesis....	120
3.3.6.3	Genes encoding core polysaccharide biosynthesis.....	121
3.3.6.4	Genes encoding peptidoglycan biosynthesis.....	121
3.4	Discussion.....	123
3.4.1	Catabolic energy producing pathways.....	124
3.4.2	Anabolic pathways.....	129
3.4.2.1	Nucleotide metabolism.....	129
3.4.2.2	Amino acid biosynthesis.....	133
3.4.2.3	Fatty acid and lipid biosynthesis.....	137
3.4.3	Cell wall structure biosynthesis.....	140
3.4.4	Host colonisation.....	145
3.5	Summary.....	148
4	Analysis of overlapping genes in the genomes of <i>Brachyspira</i>	149
	<i>hyodysenteriae</i> and <i>Brachyspira pilosicoli</i>, and in other spirochaetes.	
4.1	Introduction.....	150
4.2	Materials and Methods.....	153
4.2.1	Sequence Data.....	153
4.2.2	Gene prediction and genome annotation.....	153
4.2.3	Identification of the overlapping genes.....	154
4.2.4	Metabolic pathway data.....	154
4.2.5	Promoter prediction.....	155
4.2.6	Horizontal gene transfer (HGT) analysis	155
4.3	Results	156
4.3.1	Identification of overlapping genes within spirochaetes.....	156
4.3.2	Conservation of overlapping genes.....	159
4.3.2.1	Orthologous overlapping gene pairs.....	159
4.3.2.2	Duplicated- and tri-overlapping gene pairs.....	161
4.3.3	Coexpression of overlapping genes involved in the same biological process	164

4.4 Discussion.....	175
4.4.1 Relation between genome and overlapping genes.....	175
4.4.2 Conservation of overlapping genes.....	177
4.4.3 Coexpression of overlapping genes.....	181
4.4.4 Evolutionary of overlapping gene pairs in <i>Brachyspira</i>	183
4.5 Summary.....	185
5 Conclusions and outlook	186
5.1 Assembly and annotation of the newly sequenced <i>Brachyspira</i>	187
5.2 Comparative genome analysis of <i>Brachyspira</i>	189
5.3 Metabolism of <i>Brachyspira</i>	190
5.4 Analysis of overlapping genes in <i>B. hyodysenteriae</i> , <i>B. pilosicoli</i> and other spirochaetes	191
5.5 Outlook.....	192
References	196
Appendix	223

List of Figures

Figure

- 1.1 Phylogeny of the order Spirochaetales as inferred from 16S rRNA gene sequences (Paster & Dewhirst, 2000). The genus *Brachyspira* is highlighted in red. 7
- 1.2 Description of the main steps required for the whole-genome shotgun sequencing of a bacterium and generation of genome data. 14
- 2.1 Genome annotation pipeline. 44
- 2.2 Overall progress of the *Brachyspira* genome sequencing projects demonstrating the relationship between an increasing number of reads with each successive edit and the resultant decrease in the number of contigs in each assembly. The x-axis displays the number of reads and sequencing method employed with each edit. The left-hand y-axis corresponds to the percent genome fraction (●), whilst the right-hand y-axis shows the number of contigs (■) and the average contig size (◆). The arrow corresponds to the progress of the sequencing project in relation to time. 49
- 2.3 Stages in the alignment of a complete genome sequence of *B. hyodysenteriae* WA1 (top panels) and a draft genome sequence of *B. pilosicoli* 95/1000 (bottom panels). 52
- 2.4 A summary of the outcome of the initial annotation of *B. hyodysenteriae* WA1 and *B. pilosicoli* 95/1000 genomes. 57
- 2.5 Genome of *B. hyodysenteriae* strain WA1 (A): Chromosome (CI); (B): Plasmid (PI). Circles range from 1 (outer circle) to 6 (inner circle) for C1 and I (outer circle) to IV (inner circle) for PI. Circles 1/I and 2/II, genes and forward and reverse strand; circle 3, tRNA genes; circle 4, rRNA genes, circle 5/III, GC bias/skew ((G-C)/(G+C); red indicates values >0; green indicates values < 0); circles 6/IV, A+T percentage content. 60
- 2.6 The *B. hyodysenteriae* WA1 plasmid. 62

- 2.7 (A) Predicted position of the origins of replication in *B. hyodysenteriae* and *B. pilosicoli*, (B) compared to the putative origin of replication (*oriC*) regions of other spirochaetes. Homologous genes commonly found at bacterial origins are indicated in similar colours. *dnaA* is indicated in red and hypothetical proteins are indicated in grey. Conservation of a cluster of genes is located around the origin of replication in several spirochaetes. The putative proteins and the origins of replication are indicated. 64
- 2.8 Conservation of the *r*-protein in *B. hyodysenteriae* WA1 (BH) and *B. pilosicoli* (BP) 95/1000. 65
- 2.9 Codon usage for all open reading frames in *B. hyodysenteriae* WA1 and *B. pilosicoli* 95/1000. (A): Relationship of ORFs and G+C content (B): Codon usage for all open reading frames within *Brachyspira* genomes. Note that several of the codons are rarely used, whilst others are quite common. 66
- 2.10 Distribution of taxonomic best matches for BLASTp homology of representative *B. hyodysenteriae* (BH) and *B. pilosicoli* (BP) predicted proteins against the representative proteomes of bacteria from the NCBI. (A) Best matched and (B) within best 5 matched were derived from Supplementary Table S2.4 in Appendix A. 70
- 2.11 Venn diagram showing the number of unique and shared genes amongst the *B. hyodysenteriae* and *B. pilosicoli* genomes. The two circles represent the total number of CDSs predicted in each genome whilst the area of overlap indicates the number of orthologs predicted by reciprocal BLASTp analysis (threshold e-value = $1e^{-05}$). The pie chart depicts the COG functional groups of the 737 orthologs. 72
- 2.12 Distribution of Cluster of Orthologous Genes (COGs) in *B. hyodysenteriae* WA1, *B. pilosicoli* 95/100 and some other bacterial genomes. The table shows the fraction of proteins within a genome assigned to each functional group. The pie charts show the proportion of functional groups within each species. The codes for the COG functional groups and species are detailed at the bottom of the figure. 74

3.1	Central metabolic pathway construction for <i>B. hyodysenteriae</i> strain WA1 and <i>B. pilosicoli</i> strain 95/1000. The main difference between these two species was the rhamnose biosynthesis pathway, shown in red, encoded by genes within the <i>B. hyodysenteriae</i> plasmid.	96
3.2	Glycolysis and the non-oxidative pentose phosphate pathway in <i>B. hyodysenteriae</i> WA1 and <i>B. pilosicoli</i> 95/1000. The gene name is shown in pink and corresponding predicted ORFs present in the <i>B. hyodysenteriae</i> genome are indicated in red, and those for in <i>B. pilosicoli</i> in blue.	98
3.3	Nucleotide biosynthesis pathway in <i>B. hyodysenteriae</i> WA1 and <i>B. pilosicoli</i> 95/1000. The gene name is shown in pink and corresponding predicted ORFs present in the <i>B. hyodysenteriae</i> genome are indicated in red, and and those for in <i>B. pilosicoli</i> in blue.	103
3.4	The metabolism of the aspartate family in <i>B. hyodysenteriae</i> WA1 and <i>B. pilosicoli</i> 95/1000. The gene name is shown in pink and corresponding predicted ORFs present in the <i>B. hyodysenteriae</i> genome are indicated in red, and those for in <i>B. pilosicoli</i> in blue.	108
3.5	The metabolism of the glutamate family in <i>B. hyodysenteriae</i> WA1 and <i>B. pilosicoli</i> 95/1000. The gene name is shown in pink and corresponding predicted ORFs present in the <i>B. hyodysenteriae</i> genome are indicated in red, and those for in <i>B. pilosicoli</i> in blue.	109
3.6	Schematic pathways for (A) Fatty acid biosynthesis (B) Glycerolipid metabolism in <i>B. hyodysenteriae</i> WA1 and <i>B. pilosicoli</i> 95/1000. The gene name is shown in pink and corresponding predicted ORFs present in the <i>B. hyodysenteriae</i> genome are indicated in red, and those for in <i>B. pilosicoli</i> in blue.	115
3.7	Different biosynthetic pathways for lipooligosaccharide in <i>B. hyodysenteriae</i> WA1 and <i>B. pilosicoli</i> 95/1000. The gene name is shown in pink and corresponding predicted ORFs present in the <i>B. hyodysenteriae</i> genome are indicated in red, and those for in <i>B. pilosicoli</i> in blue.	118

3.8	Stepwise assembly of the peptidoglycan monomer in <i>B. hyodysenteriae</i> WA1 and <i>B. pilosicoli</i> 95/1000. The gene name is shown in pink and corresponding predicted ORFs present in the <i>B. hyodysenteriae</i> genome are indicated in red, and those for in <i>B. pilosicoli</i> in blue.	123
3.9	Schematic representation of the cell wall of Gram-negative bacteria. The three-part lipopolysaccharide complex (LPS) is anchored in the outer membrane by means of its lipid moiety. LPS is also known as endotoxin.	141
4.1	The three different types of orientation found in overlapping genes.	151
4.2	Correlation of overlapping genes and genome size in spirochaete genomes. Spirochaete names are the same as in Table 4.1.	158
4.3	Venn diagram showing the number of orthologous overlapping gene pairs in the newly sequenced <i>Brachyspira</i> species and other spirochaete genomes that were shared between the species in each genus. The number in the middle of the overlapped circle represented the number of orthologous overlapping gene pairs that were shared. Spirochaete names are the same as in Table 4.1.	160
4.4	Functional breakdown of overlapping genes among the <i>Brachyspira</i> species and other spirochaete core genes. Spirochaete names are the same as in Table 4.1. <i>E. coli</i> is included for comparison.	165
4.5	Variation in overlapping genes in the KEGG metabolic network. (A) Percentage of identified overlapping genes in the KEGG database, and (B) Distribution of overlapping genes in different metabolic and pathways. Spirochaete names are the same as in Table 4.1.	167
4.6	Overlapping genes involved in pyrimidine biosynthesis in both <i>Brachyspira</i> species.	169
4.7	Nucleotide sequence of <i>B. hyodysenteriae</i> strain WA1 <i>thyA</i> and <i>folA</i> .	172
4.8	Schema of the different reductive mechanisms of the folate cycle by thymidylate synthases. (A) <i>thyA</i> and (B) <i>thyX</i> . This schematic is modified from Myllykallio <i>et al.</i> , (2003).	173
4.9	Phylogenetic trees of <i>thyA</i> and <i>folA</i> in <i>Brachyspira</i> species and other bacterial species.	174

List of Tables

Table

1.1	<i>Brachyspira</i> species, pathogenicity and animal host.	8
2.1	Summary of the <i>B. hyodysenteriae</i> WA1 and <i>B. pilosicoli</i> 95/1000 assemblies, at various phases, and the main genomic features.	50
2.2	General genome features of <i>B. hyodysenteriae</i> WA1 and <i>B. pilosicoli</i> 95/1000.	59
2.3	Genome features among spirochaete genomes.	78
3.1	The key metabolic capabilities of <i>B. hyodysenteriae</i> WA1 and <i>B. pilosicoli</i> 95/1000 in comparison to other spirochaetes and <i>E. coli</i> .	94
3.2	The presence of important genes involved in folate biosynthesis in the spirochaetes species.	133
4.1	Number of overlapping genes and their patterns in spirochaete species and <i>Escherichia coli</i> .	157
4.2	The number of copies of duplicated- and tri-overlapping genes in spirochaete genomes. Spirochaete names are the same as in Table 4.1.	163
4.3	Distribution of overlapping genes of inferred genome-wide metabolic networks in spirochaetes and <i>E. coli</i> as found in the KEGG database. Spirochaete names are the same as in Table 4.1.	168
4.4	List of overlapping genes involved in pyrimidine metabolism in <i>B. hyodysenteriae</i> strain WA1 and <i>B. pilosicoli</i> strain 95/1000.	170
4.5	The number of overlapping <i>thyA</i> and <i>folA</i> genes found in bacterial genomes.	171
S2.1	A list of potential vaccine candidates predicted <i>in silico</i> from <i>B. hyodysenteriae</i> WA1 using SignalP and TMHMM.	224
S2.2	A list of potential vaccine candidates predicted <i>in silico</i> from <i>B. pilosicoli</i> 95/1000 using SignalP and TMHMM.	228
S2.3	Statistical analysis of codon usage in <i>B. hyodysenteriae</i> strain WA1 and <i>B. pilosicoli</i> strain 95/1000.	231
S2.4	Similarity of predicted <i>B. hyodysenteriae</i> strain WA1 (BH) and <i>B. pilosicoli</i> strain 95/1000 (BP) proteins to proteins from other taxa.	233
S2.5	Transporter genes identified in <i>B. hyodysenteriae</i> strain WA1.	235

S2.6	Transporter genes identified in <i>B. pilosicoli</i> strain 95/1000.	238
S3.1	Comparative analysis of metabolism among spirochaetes.	240