This may be the author's version of a work that was submitted/accepted for publication in the following source:

Ahmed, Niyaz, Wanchanthuek, Phatthanaphong, Bellgard, Matthew, La, Tom, Ryan, Karon, Moolhuijzen, Paula, Chapman, Brett, Black, Michael, Schibeci, David, Hunter, Adam, Barrero, Roberto, Phillips, Nyree, & Hampson, David (2010) The complete genome sequence of the pathogenic intestinal spirochete Brachyspira pilosicoli and comparison with other Brachyspira genomes. *PLoS One, 5*(7), Article number: e11455.

This file was downloaded from: https://eprints.qut.edu.au/125426/

© The Author(s)

This work is covered by copyright. Unless the document is being made available under a Creative Commons Licence, you must assume that re-use is limited to personal use and that permission from the copyright owner must be obtained for all other uses. If the document is available under a Creative Commons License (or other specified license) then refer to the Licence for details of permitted re-use. It is a condition of access that users recognise and abide by the legal requirements associated with these rights. If you believe that this work infringes copyright please provide details by email to qut.copyright@qut.edu.au

License: Creative Commons: Attribution 4.0

Notice: Please note that this document may not be the Version of Record (i.e. published version) of the work. Author manuscript versions (as Submitted for peer review or as Accepted for publication after peer review) can be identified by an absence of publisher branding and/or typeset appearance. If there is any doubt, please refer to the published source.

https://doi.org/10.1371/journal.pone.0011455
The Complete Genome Sequence of the Pathogenic Intestinal Spirochete \textit{Brachyspira pilosicoli} and Comparison with Other \textit{Brachyspira} Genomes

Phatthanaphong Wanchanthuek$^{1,2}$, Matthew I. Bellgard$^1$, Tom La$^3$, Karon Ryan$^1$, Paula Moolhuijzen$^1$, Brett Chapman$^1$, Michael Black$^1$, David Schibeci$^1$, Adam Hunter$^1$, Roberto Barrero$^1$, Nyree D. Phillips$^3$, David J. Hampson$^3$*

$^1$Centre for Comparative Genomics, Murdoch University, Perth, Western Australia, Australia, $^2$Faculty of Informatics, Mahasarakham University, Mahasarakham, Thailand, $^3$Animal Research Institute, School of Veterinary and Biomedical Science, Murdoch University, Perth, Western Australia, Australia

Abstract

\textbf{Background:} The anaerobic spirochete \textit{Brachyspira pilosicoli} colonizes the large intestine of various species of birds and mammals, including humans. It causes “intestinal spirchotosis”, a condition characterized by mild colitis, diarrhea and reduced growth. This study aimed to sequence and analyze the bacterial genome to investigate the genetic basis of its specialized ecology and virulence.

\textbf{Methodology/Principal Findings:} The genome of \textit{B. pilosicoli} 95/1000 was sequenced, assembled and compared with that of the pathogenic \textit{Brachyspira hyodysenteriae} and a near-complete sequence of \textit{Brachyspira murochii}. The \textit{B. pilosicoli} genome was circular, composed of 2,586,443 bp with a 27.9 mol% G+C content, and encoded 2,338 genes. The three \textit{Brachyspira} species shared 1,087 genes and showed evidence of extensive genome rearrangements. Despite minor differences in predicted protein functional groups, the species had many similar features including core metabolic pathways. Genes distinguishing \textit{B. pilosicoli} from \textit{B. hyodysenteriae} included those for a previously undescribed bacteriophage that may be useful for genetic manipulation, for a glycine reductase complex allowing use of glycine whilst protecting from oxidative stress, and for aconitase and related enzymes in the incomplete TCA cycle, allowing glutamate synthesis and function of the cycle during oxidative stress. \textit{B. pilosicoli} had substantially fewer methyl-accepting chemotaxis genes than \textit{B. hyodysenteriae} and hence these species are likely to have different chemotactic responses that may help to explain their different host range and colonization sites. \textit{B. pilosicoli} lacked the gene for a new putative hemolysin identified in \textit{B. hyodysenteriae} WA1. Both \textit{B. pilosicoli} and \textit{B. murochii} lacked the \textit{rfdBADC} gene cluster found on the \textit{B. hyodysenteriae} plasmid, and hence were predicted to have different lipooligosaccharide structures. Overall, \textit{B. pilosicoli} 95/1000 had a variety of genes potentially contributing to virulence.

\textbf{Conclusions/Significance:} The availability of the complete genome sequence of \textit{B. pilosicoli} 95/1000 will facilitate functional genomics studies aimed at elucidating host-pathogen interactions and virulence.

Citation: Wanchanthuek P, Bellgard MI, La T, Ryan K, Moolhuijzen P, et al. (2010) The Complete Genome Sequence of the Pathogenic Intestinal Spirochete \textit{Brachyspira pilosicoli} and Comparison with Other \textit{Brachyspira} Genomes. PLoS ONE 5(7): e11455. doi:10.1371/journal.pone.0011455

Editor: Niyaz Ahmed, University of Hyderabad, India

Received May 3, 2010; Accepted June 13, 2010; Published July 6, 2010

Copyright: © 2010 Wanchanthuek et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was supported by the Australian Research Council and Novartis Animal Vaccines through Linkage grant number LP0348441 for genomic sequencing and comparative genomic analysis for animal bacterial vaccine discovery, the Western Australian Government in establishment of the WA Centre of Excellence for Comparative Genomics and for support for this project, and the Mahasarakham University Research Affairs Division. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: This study included identifying bacterial proteins that could be useful for a vaccine. Part of this project was funded by Novartis Animal Vaccines. Several of the proteins that were identified are being patented by Spirogene Pty, a Murdoch University spin-out company. This does not alter the authors’ adherence to all the PLoS ONE policies on sharing data and materials.

* E-mail: d.hampson@murdoch.edu.au

Introduction

Spirochetes of the genus \textit{Brachyspira} are anaerobic bacteria that colonize the large intestine of animals and birds [1]. There are currently seven officially named species in the genus, including both pathogenic and commensal representatives. The names of the species and their phylogenetic relationships based on their 16S rRNA gene sequences are shown in Figure 1. The two main pathogenic species are \textit{Brachyspira hyodysenteriae}, the agent of a major pig disease called swine dysentery, and \textit{Brachyspira pilosicoli}, the cause of a condition known as intestinal (or colonic) spirchotosis. \textit{B. pilosicoli} has a wider host range than \textit{B. hyodysenteriae}, and colonizes a variety of species, including human beings [2–6]. Infections with \textit{B. pilosicoli} are particularly common in intensively housed pigs and chickens: colonized individuals may develop focal areas of inflammation in the large intestine, with diarrhea and reduced rates of growth and production. Colonization with \textit{B. pilosicoli} also occurs at a high prevalence rate in human beings living in crowded and unhygienic conditions, particularly in developing countries [7–11], as well as amongst homosexual males.
Despite the potential importance of *B. pilosicoli* as a pathogen, it has not been extensively studied. Progress has been hampered by the spirochete’s specialised growth requirements and slow growth rate, as well as by a lack of genomic information and an absence of means for genetic manipulation.

A characteristic feature of infection with *B. pilosicoli* is the attachment of the spirochete by one cell end to the luminal surface of enterocytes in the large intestine; with time, dense matts of these attached bacteria may form a “false brush-border” covering the whole luminal surface of the enterocytes [12,13]. Experimentally, following attachment to cultured Caco-2 cells *B. pilosicoli* has been shown to induce apoptosis, actin rearrangement and increased expression of interleukin-1ß (IL-1ß) and IL-8 [14]. Such pathological changes may contribute to the local colitis and diarrhea that is observed *in vivo*.

The recent publication of the genome sequence of *B. hyodysenteriae* strain WA1 [15] and the availability of a near-complete genome sequence of *Brachyspira murdochii* type strain 56-150T (DSM 12563) in GenBank represent important opportunities to advance research on these *Brachyspira* species. *B. murdochii* is generally considered to be a harmless commensal in pigs, chickens and rats [16,17]; however, there have been recent reports that it may have some potential to cause colitis in pigs [18,19]. In the current study the genome of *B. pilosicoli* strain 95/1000 was sequenced and subjected to comparative genomic analysis, particularly in relation to the sequences of WA1 and 56-150T. As the sequence of the latter was incomplete some comparisons were not appropriate. The overall aim of this work was to enhance knowledge about *B. pilosicoli* and its relationships to the other *Brachyspira* species, particularly in regard to identifying the genetic basis of their different ecologies and pathogenic potentials.

**Materials and Methods**

**Spirochete strain and growth conditions**

*Brachyspira pilosicoli* strain 95/1000 (ATCC BAA-1826) was originally isolated in Western Australia from the diarrheic feces of a commercial pig with porcine intestinal spirochetosis [13]. The spirochete was purified by repeated subculture, grown to mid-log phase in pre-reduced anaerobic broth [20], and a cell pellet prepared.

**Genomic DNA preparation, library construction and sequencing**

Preparation of genomic DNA, library construction and sequencing was as previously described for *B. hyodysenteriae* [15]. Sequencing was undertaken at the Australian Genome Research Facility. The first round of sequencing was performed via Sanger sequencing, with a total of 42,565 reads generated. The second round was performed using a pyrosequencing approach on a Roche-454 GS20 instrument, generating more than 25 times coverage of the genome. The quality filtered reads were assembled into contiguous sequences using the Newbler Assembler software (http://www.454.com/). To finish the genome sequence, remaining gaps were closed by PCR walking between un-linked contiguous sequences [21].

**Sequence analysis and annotation**

For both the newly sequenced *B. pilosicoli* 95/1000 and for the incomplete genome sequence of *B. murdochii* 56-150T obtained from GenBank (ABTG00000000), sequence analysis and annotation were as previously described for *B. hyodysenteriae* WA1 [15]. Dot matrix plots comparing the genomes of 95/1000, 56-150T and WA1 were generated using Freckle, an in-house development of the Dotter tool [22]. The minimum size of matched sequences was set to 20 base pairs (bp).

The complete nucleotide sequence and annotation of *B. pilosicoli* 95/1000 has been deposited in GenBank (accession number CP002025; Project ID: 48097). Annotations and functional assignments for *B. pilosicoli* 95/1000 and for *B. murdochii* 56-150T also can be accessed at the CCG website (http://ccg.murdoch.edu.au/spirochaetales/).

**Results and Discussion**

**General genome features**

The general features of the genome of *B. pilosicoli* 95/1000, together with those of *B. hyodysenteriae* WA1 and the near-complete genome of *B. murdochii* 56-150T are summarised in Table 1. The complete genome of *B. pilosicoli* 95/1000 consisted of a single circular chromosome of 2,586,443 bp (Figure 2), making it the...
the three Brachyspira genomes analysed to date (with B. murdochii being the largest). The overall G+C content of the B. pilosicoli genome was 27.9 mol%. Unlike B. hyodysenteriae WA1, B. pilosicoli 95/1000 did not contain a ~36 Kb plasmid. Based on the available sequence it also appeared unlikely that B. murdochii contained a similar plasmid. The B. pilosicoli genome encoded 2,338 genes, with an overall 85% coding region. In comparison, B. murdochii 56-150T had 3,055 predicted genes in the seven contigs that were available. The percent coding region was similar across the three genomes. The average size of the predicted genes in B. pilosicoli 95/1000 was 997 bp, and tentative function was assigned to 1,645 of these. The B. pilosicoli strain had 569 more functionally assigned genes than did B. murdochii 56-150T and 282 less than B. hyodysenteriae. B. pilosicoli 95/1000 had 655 genes with unknown function and this was somewhat fewer than for B. hyodysenteriae (704), but much fewer than the extremely high number of unmatched genes for B. murdochii (1,925). A total of 1,201 (51.4%) genes in B. pilosicoli had matches in the COG database and 1,410 (60.3%) had matches in the KEGG database (Table 1). Of these, only 426 (18.2%) were assigned to enzymes (EC number) and 635 (27.2%) were involved in KEGG pathways. Of the B. pilosicoli open reading frames (ORFs) screened, 244 were predicted to encode proteins with signal peptides and 48 were predicted to have transmembrane helices. Eleven genes encoded proteins predicted to have seven transmembrane helices.

The three Brachyspira species had similar numbers of tRNA genes, representing all 20 amino acids (Table 1). As with B. hyodysenteriae [15], only single copies of the 16S, 23S, and 5S ribosomal RNA genes were found in B. pilosicoli and B. murdochii. In B. pilosicoli 95/1000, the rrS and rrL genes were closely linked, with the rrL (16S) gene being about 645 Kb from the other two genes. This rRNA gene organisation has been noted earlier; it distinguishes the Brachyspira species from other spirochetes, and presumably pre-dates speciation in the Brachyspira genus [23]. Although spirochetes have a monophyletic origin, the copy number and organisation of rRNA genes differ in the different genera. For instance, the Treponema pallidum rRNA genes appear to be arranged in two typical rrn operons [24,25]. A single rRNA locus is found in most Borrelia species, with rrS separated from rrL and rrF by a small segment of DNA (~4 Kb). In Borrelia burgdorferi the rrF-rrL cluster is duplicated and is found immediately adjacent to the rrS-rrL-rrF cluster [26]. Pathogenic Leptospira species possess two copies each of rrS and rrL and one copy of rrF [27]. The non-pathogenic Leptospira biflexa contains two copies of each rRNA gene [24], and these are dispersed around the genome. Differences in the sequences of the 16S rRNA genes have been used to examine phylogenetic relationships between spirochete genera and species [28]. On this basis, in the case of the Brachyspira species, B. pilosicoli is more distantly related to B. hyodysenteriae than to B. murdochii (Figure 1).

The origin and terminus of replication in the genome of B. pilosicoli 95/1000 were predicted based on the position of dnaA (BP951000_0595), as bacterial chromosome replication origins are typically located near this gene. The putative oriC origin of replication was identified in an AT-rich intergenic region upstream of dnaA, in the vicinity of a cluster of hypothetical DnaA boxes. Similar putative oriC with comparable DnaA boxes have been identified in other spirochete genomes (Figure 3). The original of replication in B. murdochii was not found due to the genome sequence being incomplete. The putative origins of replication are located centrally within the most highly conserved and syntenic regions of the various spirochete genomes (Figure 3). B. pilosicoli contained the genes gagF, dnaK, hyp, ark, hyp, arg and gys1 downstream of dnaA, an arrangement similar to B. hyodysenteriae WA1 [15]. B. pilosicoli had a unique hypothetical coding region immediately upstream of the dnaA gene, whereas there were three hypothetical protein encoding genes upstream of dnaA in B. hyodysenteriae. This finding suggests that in B. pilosicoli 95/1000 and B. hyodysenteriae WA1 oriC relocated during evolution, presumably as the result of a DNA rearrangement. The differences in the origins of replication compared to other spirochetes suggest that there could be different mechanisms for replication of the spirochete chromosomes. Experimental studies will be required to verify the origin of replication in B. pilosicoli, as has been accomplished for Borrelia burgdorferi [29].

---

Table 1. General genomic features predicted for B. pilosicoli 95/1000 and comparison with those of B. hyodysenteriae WA1 and B. murdochii 56-150T.

| Feature                        | B. pilosicoli 95/1000 | B. hyodysenteriae WA1 | B. murdochii 56-150T* |
|-------------------------------|----------------------|-----------------------|------------------------|
| Genome size (bp)              | 2,586,443            | 3,036,634             | 3,189,383              |
| Coding region (%)             | 85%                  | 86%                   | 86%                    |
| G+C content (%)               | 27.90%               | 27.06%                | 26.91%                 |
| Total genes                   | 2,338                | 2,669                 | 3,055                  |
| tRNA genes                    | 3                    | 3                     | 3                      |
| Hypothetical protein/Conserved hypothetical protein | 655 | 704 | 1,925 |
| Genes with function prediction| 1,645                | 1,927                 | 1,076                  |
| Protein–coding genes          | 1,987                | 2,153                 | 2,997                  |
| Number of multiple copies of gene (%) | 99 (4.3)     | 210 (7.9)             | 204 (6.7)              |
| Genes assigned to COG         | 1,201                | 1,217                 | 1,217                  |
| Genes assigned to KEGG        | 1,410                | 1,492                 | 1,509                  |
| Genes assigned E.C. numbers   | 426                  | 509                   | 436                    |
| Genes with signal peptide     | 244                  | 360                   | 220                    |
| Genes with transmembrane helices | 48                 | 52                    | 34                     |

*Incomplete genome containing 7 contigs (GenBank accession number ABTG00000000).

doi:10.1371/journal.pone.0011455.t001
Whole-genome alignment

A dot plot comparison of the genomes of *B. pilosicoli*, *B. hyodysenteriae* and *B. murdochii* showed no large scale conservation of gene order, but only conservation of some genes in clusters (Figure 4). Gene cluster conservation appeared to be greatest between *B. hyodysenteriae* and *B. murdochii*, consistent with these two species being phylogenetically more closely related to each other than to *B. pilosicoli*. A 26 Kb region of unknown significance that is partially conserved in *B. hyodysenteriae* WA1, *B. pilosicoli* 95/1000, *Enterococcus faecalis* and *Escherichia coli* [30], also was identified in *B. murdochii*. Alignment of the two complete *Brachyspira* genomes using the Artemis Comparison Tool [31] identified extensive gene

Figure 2. Genome of *B. pilosicoli* 95/1000. Circles outer to inner: Genes, forward strand, COG coded; Genes, reverse strand, COG coded; tRNA; rRNA; GC skew; AT skew. All genes are color-coded according to Cluster of Orthologous Gene Categories (COG) functions: violet for translation, ribosomal structure and biogenesis; plum for RNA processing and modification; pink for transcription; deep pink for DNA replication, recombination and repair; hot pink for chromatin structure and dynamics; wheat for cell division and chromosome partitioning; light salmon for nuclear structure; yellow for defence mechanisms; gold for signal transduction mechanisms; pale green for cell envelope biogenesis, outer membrane; spring green for cell motility and secretion; lawn green for cytoskeleton; yellow green for extracellular structures; aquamarine for intracellular trafficking, secretion, and vesicular transport; medium aquamarine for post translational modification, protein turnover, chaperones; cyan for energy production and conversion; deep sky blue for carbohydrate transport and metabolism; sky blue for amino acid transport and metabolism; light slate blue for nucleotide transport and metabolism; orchid for coenzyme metabolism; medium orchid for lipid metabolism; dark orchid for inorganic ion transport and metabolism; blue violet for secondary metabolites biosynthesis, transport and catabolism; slate grey for general function prediction only; grey for function unknown; grey for not in COGs; black for tRNA.
doi:10.1371/journal.pone.0011455.g002
rearrangements between them (Figure S1 in supporting information), consistent with the findings of an earlier comparison of partial physical maps of the two species [32].

An exception to the apparent general lack of conservation of gene order between the three Brachyspira genomes was a high conservation of the operons encoding ribosomal proteins. B. pilosicoli 95/1000 contained genes for 33 r-proteins organised in a 18 Kb region, whereas the ribosome clusters in the B. hyodysenteriae and B. murdochii strains included a total of 32 and 36 genes encoding r-proteins organised in 15 Kb and 18 Kb regions, respectively (Figure 5). The sequences of the B. pilosicoli r-proteins in this cluster were highly similar to the B. hyodysenteriae and B. murdochii homologs. Ribosomal protein genes are interesting because of their ubiquity and similar conservation rate such that horizontal transfer between lineages is unlikely. Apparently there has been a positive selection for clustering of these physically interacting proteins in the sequenced Brachyspira species, whilst there has been no absolute requirement for juxtaposition of other genes in the genomes. Synteny therefore has been lost at a much faster rate than is useful for prediction of gene function. This is consistent with the situation in other bacterial species, where orthologous genes are not necessarily located at the same relative position in the genomes, and only certain gene clusters are syntenic [33]. Genomes of closely related species typically maintain a high degree of synteny [34], whereas genomes of moderately distant species, such as these Brachyspira species, usually have no significant overall synteny [35].

Functional predictions amongst the Brachyspira species

The COG categories of the protein-coding genes are shown in Table 2. Generally there were few significant differences amongst the three genomes, although the smaller genome of B. pilosicoli 95/1000 contained more genes than the other two species in five categories: (N) Cell motility, (U) Intracellular trafficking, secretion, and vesicular transport, (C) Energy production and conversion, (H) Co-enzyme transport and metabolism, and (I) Lipid transport and metabolism. On the other hand, B. hyodysenteriae had more genes functioning in (K) Transcription, (V) Defense mechanisms, (T) Signal transduction mechanisms, (O) Posttranslational modification, protein turnover, chaperones, (G) Carbohydrate transport and metabolism, (E) Amino acid transport and metabolism, (P) Inorganic ion transport and metabolism, and (Q) Secondary metabolites biosynthesis, transport and catabolism. B. murdochii had more genes involved with (L) Replication, recombination and repair, and (M) Cell wall/membrane/envelope biogenesis. It was noteworthy that B. pilosicoli had the largest number of genes involved in energy production and conversion, and this capacity may enhance its potential to colonize various hosts.

Previously it has been shown that B. hyodysenteriae lacks many genes for the biosynthesis of small molecules, and therefore must
Figure 4. Dot matrix plots comparing the genomes of 95/1000, 56-150T and WA1 generated using Freckle. The output displays a two dimensional plot with dots representing matched regions between the three genomes. The minimum size of matched sequences was set to 20 bp. doi:10.1371/journal.pone.0011455.g004

Figure 5. Conservation of the gene cluster for ribosomal proteins in B. pilosicoli 95/1000, B. hyodysenteriae WA1 and B. murdochii 56-150T. doi:10.1371/journal.pone.0011455.g005
acquire these from the environment [15]. Interestingly, *B. pilosicoli* was predicted to contain substantially fewer genes involved in inorganic ion transport and metabolism (P) than both *B. hyodysenteriae* and *B. murdochii* (53 versus 74 and 73, respectively; Table 2). Furthermore, it is not known which *B. pilosicoli* genes may act to compensate for the shortage of such biosynthetic pathways, since the substrates of many of the genes regarded as encoding transporters due to their possession of the motif sequences were unknown. These differences were unexpected, given that *B. pilosicoli* has a wider host range than the other two *Brachyspira* species, and hence might be predicted to have a greater metabolic capacity and flexibility in order to survive in these more varied nutritional environments.

Global gene comparisons between the *Brachyspira* species

As expected, comparative analysis of the *B. pilosicoli* genome across available microbial genomes in the non-redundant (nr) database at NCBI identified greatest similarities with *B. hyodysenteriae* and *B. murdochii*. However, as with *B. hyodysenteriae* [15], the next highest levels of similarities at the protein level were with *Clostridium* species (~10%) and *E. coli* (~5%). The three sequenced *Brachyspira* species contained a total of 1,087 conserved or "core" genes (Figure 6). It would be instructive to determine whether these same genes are conserved in other *Brachyspira* species. The majority of genes in *B. pilosicoli* (1,769, 77%) were conserved in *B. hyodysenteriae*; therefore, most of the genome-inferred metabolic potential of *B. hyodysenteriae* described previously can be extrapolated to *B. pilosicoli*. Only an additional 99 genes were conserved between *B. pilosicoli* and *B. murdochii*, whereas an additional 311 were conserved between *B. hyodysenteriae* and *B. murdochii*. *B. hyodysenteriae* had 1,014 genes that were not shared with *B. pilosicoli*, whereas *B. murdochii* contained a remarkable 1,900 genes not found in *B. pilosicoli*. On this basis, *B. pilosicoli* seemed to be more similar to *B. hyodysenteriae* than to *B. murdochii*. A large proportion of the unique genes in both *B. hyodysenteriae* and *B. murdochii* were of unknown function (Table 1).

Although the *B. hyodysenteriae* genome was ~450 Kb larger than that of *B. pilosicoli*, both species contained roughly the same

### Table 2. Distribution of Cluster of Orthologous Genes (COGs) in *B. pilosicoli* 95/1000, *B. hyodysenteriae* WA1 and *B. murdochii* 56-150T, showing the number and percentage of proteins within a genome assigned to each functional group (the *B. murdochii* genome is incomplete).

| Functions                                      | *B. pil* | %    | *B. hyo* | %    | *B. mur* | %    |
|------------------------------------------------|---------|------|----------|------|----------|------|
| **Cellular Processes**                         |         |      |          |      |          |      |
| J Translation, ribosomal structure and biogenesis | 122     | 5.30 | 122      | 4.57 | 120      | 4.00 |
| K Transcription                                | 51      | 2.21 | 63       | 2.36 | 46       | 1.53 |
| L Replication, recombination and repair        | 51      | 2.21 | 51       | 1.91 | 63       | 2.10 |
| **Cellular Processes and Signalling**          |         |      |          |      |          |      |
| D Cell cycle control, cell division, chromosome partitioning | 10      | 0.43 | 10       | 0.37 | 9        | 0.30 |
| V Defence mechanisms                           | 35      | 1.52 | 40       | 1.50 | 32       | 1.07 |
| T Signal transduction mechanisms               | 16      | 0.69 | 21       | 0.79 | 18       | 0.60 |
| M Cell wall/membrane/envelope biogenesis       | 74      | 3.21 | 75       | 2.81 | 81       | 2.70 |
| N Cell motility                                | 40      | 1.74 | 38       | 1.42 | 38       | 1.27 |
| U Intracellular trafficking, secretion, and vesicular transport | 11      | 0.48 | 10       | 0.37 | 8        | 0.27 |
| O Posttranslational modification, protein turnover, chaperones | 40      | 1.74 | 41       | 1.54 | 37       | 1.23 |
| **Metabolism**                                 |         |      |          |      |          |      |
| C Energy production and conversion             | 89      | 3.86 | 83       | 3.11 | 82       | 2.74 |
| G Carbohydrate transport and metabolism        | 101     | 4.39 | 108      | 4.05 | 104      | 3.47 |
| E Amino acid transport and metabolism          | 141     | 6.12 | 148      | 5.55 | 115      | 3.84 |
| F Nucleotide transport and metabolism          | 49      | 2.13 | 48       | 1.80 | 49       | 1.63 |
| H Coenzyme transport and metabolism            | 47      | 2.04 | 42       | 1.57 | 46       | 1.53 |
| I Lipid transport and metabolism               | 41      | 1.78 | 37       | 1.39 | 40       | 1.33 |
| P Inorganic ion transport and metabolism       | 53      | 2.30 | 74       | 2.77 | 73       | 2.44 |
| Q Secondary metabolites biosynthesis, transport and catabolism | 9      | 0.39 | 16       | 0.60 | 10       | 0.33 |
| **Poorly characterised**                      |         |      |          |      |          |      |
| R General function prediction only             | 149     | 6.47 | 171      | 6.41 | 175      | 5.84 |
| S Function unknown                             | 72      | 3.13 | 70       | 2.62 | 71       | 2.37 |
| **Unassigned**                                |         |      |          |      |          |      |
| X not in COG                                   | 1,201   | 52.15| 1,401    | 52.49| 1,780    | 59.39|
| **Total**                                      | 2,303   | 100  | 2,669    | 100  | 2,997    | 100  |

doi:10.1371/journal.pone.0011455.t002
and B. murdochii associated genes in the region were absent from components of an integrated bacteriophage (Figure 7). The 29 a 25 Kb region (location 1,595,515–1,615,675 bp) encoding that has been associated with a gene transfer agent in B. hyodysenteriae, whilst 13 were bacteriophage-associated: hypothetical proteins, whilst 13 were bacteriophage-associated:

Figure 7. Organisation of the bacteriophage genome in B. pilosicoli 95/1000. doi:10.1371/journal.pone.0011455.g007

number of gene across the identified functional gene categories, and only a relatively small number of genes in B. pilosicoli were obviously different from those in B. hyodysenteriae. Of the 525 potential genes that were specific for B. pilosicoli, 319 were of unknown function. The remaining 206 functionally annotated B. pilosicoli-specific genes included those predicted to be involved in energy production, carbohydrate metabolism, amino acid metabolism, capsule biosynthesis (although B. pilosicoli is not known to have a capsule), or encoding transcriptional regulators, transporters and predicted surface antigens, as well as forming a bacteriophage region. Selected differences found in B. pilosicoli that appear to be potentially significant are described below.

Bacteriophage genes. Brachyspira pilosicoli 95/1000 contained a 25 Kb region (location 1,595,515–1,615,675 bp) encoding components of an integrated bacteriophage (Figure 7). The 29 associated genes in the region were absent from B. hyodysenteriae and B. murdochii, apart from the OrfG VSH-1 protein homolog that has been associated with a gene transfer agent in B. hyodysenteriae. Sixteen genes were identified as encoding hypothetical proteins, whilst 13 were bacteriophage-associated:

putative integrase XerDC family protein (BP951000_1455); OrfG VSH-1 phage protein homolog (BP951000_1459); putative Lys - endolysin; glycoside hydrolase (BP951000_1461); putative RNA polymerase domain-containing protein (BP951000_1463); phage tail tape measure protein homolog, TP901 family (BP951000_1469); bacteriophage TP901-like family protein homolog (BP951000_1474); putative DNA packaging, Phage QLRG family protein (BP951000_1476); phage major capsid protein homolog, HK97 family (BP951000_1477); phage prohead protease (BP951000_1478); putative phage portal protein, HK97 family (BP951000_1479); DNA methylase N-4/N-6 domain protein homolog (BP951000_1480); putative phage terminase, small subunit, P27 family protein (BP951000_1481); and putative phage terminase, large subunit (BP951000_1482). The GC content in this bacteriophage region was 29.65%, which was slightly higher than the average for the whole genome (27.90%). It is not yet clear whether this integrated bacteriophage is functional, or capable of transferring genetic material between B. pilosicoli strains. If it is, it could be developed into a tool for genetic manipulation of the spirochete. By analogy with other prophages, it potentially also could be involved in lysogenic conversion, for example modifying the spirochete phenotype to enhance its virulence [36]. It will be important to determine how widespread and conserved this bacteriophage region is amongst B. pilosicoli strains. B. hyodysenteriae WA1 also contained two genes predicted to be bacteriophage-associated: phage terminase large subunit (strB, BHWA1_01969) and integrase (BHWA1_02688), but these were distinct from the bacteriophage genes found in B. pilosicoli. A predicted integrase-recombinase protein (4083292.C14.orf00918) was found in B. murdochii, but it only had 25.61% identity at the protein level.

The novel bacteriophage region in B. pilosicoli 95/1000 was distinct from the smaller prophage-like gene transfer agent (GTA) that previously was described in the genomes of B. pilosicoli 95/1000 and in B. hyodysenteriae WA-1 [37], and which was originally described in B. hyodysenteriae B204 as VSH-1 [38,39]. The genes of the GTAs in 95/1000 and WA1 were contiguous, but showed gene rearrangements. Elements of the GTA also have been detected in different contigs of a partial genome sequence of B. intermedia HB60 [37]. In the current analysis OrfE was found immediately in front of the other 17 GTA genes that were previously described in B. pilosicoli 95/1000 [37]. B. murdochii 56-150T was found to have a cluster of 31 genes on contig 4083292.C21 (orf00108–orf00140) that included many of the described components of the GTA, as well as 17 hypothetical proteins. Genes for Hvp19/Hvp22, Hvp13, Hvp38, and Hvp28 were not identified in the partial genome of B. murdochii. Again, the known genes showed rearrangements from the order found in B. hyodysenteriae and B. pilosicoli, and it is unclear whether the B. murdochii GTA is functional.
Genes of the glycine reductase complex system. *B. pilosicoli* 95/1000 and *B. murchoi* 56-150\(^1\) differed from *B. hyodysenteriae* WA1 in that they possessed nine genes within a glycine reductase complex (gd cluster); this arrangement is similar to that in the anaerobic bacterium *Clostridium sticklandii* [40] (Figure 8). The locus in *B. pilosicoli* included genes encoding the GridX protein (gdE, BP951000_1852), thioredoxin reductase (trxB, BP951000_1853), glycine reductase (gdE, BP951000_1854), sarcosine reductase (gdD, BP951000_1855), glycine reductase selenoprotein B (gdB, BP951000_1856), two copies for the glycine/sarcosine/betaine reductase complex, component C, alpha subunit (gdC, BP951000_1857 and BP951000_1858), and two copies for a sodium:alanine symporter family protein (BP951000_1859 and BP951000_1860). In *B. murchoi*, the locus consisted of genes encoding putative glycine reductase complex component (gdX, 4083292.C42.orf00552), thioredoxin reductase (trxB, 4083292.C42.orf00553), glycine/sarcosine/betaine reductase component B alpha/beta subunit (gdE, 4083292.C42.orf00554), two copies of glycine reductase complex selenoprotein A (gdA, 4083292.C42.orf00556 and 4083292.C42.orf00557), two copies of selenoprotein B, BP951000_0858, BP951000_0859 and BP951000_0861, were identified in *B. pilosicoli* 95/1000, whilst one putative copy was found in *B. murchoi* (4083292.C42.orf00884), and none in *B. hyodysenteriae* BHWA1. The predicted ability of *B. pilosicoli* to rapidly respond to an oxidative stress or a redox insult was consistent with its reported ability to survive outside the host for a longer period than *B. hyodysenteriae* [42,43], and also may explain why it is able to survive in more oxygenated host tissues than *B. hyodysenteriae*, such as in the bloodstream of immunocompromised patients [44]. *B. murchoi* also has been isolated from extra-intestinal sites [45]. This predicted ability to withstand oxidative stress also may make *B. pilosicoli* more adaptable in terms of its abilities to colonize the large intestines of a wide variety of host species, where environmental conditions are likely to vary.

**Sulfatase genes.** Three copies of genes encoding sulfatases (BP951000_0838, BP951000_0859 and BP951000_0861) were identified in *B. pilosicoli* 95/1000, whilst one putative copy was found in *B. murchoi* (4083292.C42.orf00884), and none in *B. hyodysenteriae*. Sulfatases play important roles in the cycling of sulfur in the environment, in the degradation of sulfated glycosaminoglycans and glycolipids in the lysosome, and in remodeling sulfated glycosaminoglycans in the extracellular space. The sulfatase genes potentially could encode for enzymes that modify glycosaminoglycans to generate binding sites required for attachment of *B. pilosicoli* [46].

**Fucosyltransferase genes.** A novel set of two copies of genes encoding alpha-1,2-fucosyltransferase (fucT, BP951000_1232 and BP951000_1235) and beta-1,3-galactosyltransferase (BP951000_1768) were identified in *B. pilosicoli* 95/1000; these are key enzymes in the biosynthesis of Lewis antigens, structures found on the surface of human erythrocytes and epithelial cells. The gastric pathogen *Helicobacter pylori* can express Lewis and related antigens in the O-chains of its surface lipopolysaccharide [47], and this activity is believed to be important for bacterial persistence, immune evasion, and possibly generation of inflammation [48]. Further work is required to determine whether *B. pilosicoli* expresses Lewis antigens, and to investigate their potential involvement in the pathogenesis of infection by this spirochete species.

**Sialidase genes.** Three copies of a gene encoding a sialidase (neuraminidase) family protein homolog (nana, BP951000_2021, BP951000_2022 and BP951000_2023) were identified in *B. pilosicoli* 95/1000. **Figure 8. Physical map of the glycine reductase genes and adjacent regions in *B. pilosicoli* 95/1000, *B. murchoi* 56-150\(^1\) and *Clostridium sticklandii*.**

doi:10.1371/journal.pone.0011455.g008
pilosici 95/1000, but not in B. hydysenteriae or B. murdochii. NanA proteins are produced by a wide variety of mucosal pathogens, and are a potential virulence factor in bacteria [49,50]. As they are widespread and conserved among a very broad range of important human pathogens, this implies that they have a critical role in microbial ecology [51]. These enzymes may enhance B. pilosici colonization or induce tissue damage.

Aconitase gene. B. pilosici and B. murdochii were both found to have a gene encoding aconitase (BP951000_0370 and 4083292.C42.orf00530, respectively), and this was not present in B. hydysenteriae. Aconitase is a tricarboxylic acid (TCA) cycle protein that catalyses the conversion of citrate to isocitrate, and, amongst other things, has been implicated in the virulence of Staphylococcus aureus [52]. It is possible that the aconitase gene may contribute to the control of virulence factor synthesis [53]. As none of the three anaerobic Brachyspira species contained genes for a complete TCA cycle, this suggests that retention of the aconitase gene may have an important functional significance. Two other genes were located adjacent to the aconitase gene in both B. pilosici and B. murdochii, these being putative citrate synthase (BP951000_0368 and 4083292.C42.orf00523) and putative isocitrate dehydrogenase (BP951000_0369 and 4083292.C42.orf00525). They were not identified in B. hydysenteriae WA1. The presence of aconitase implies that B. pilosici and B. murdochii can utilise the TCA to produce glutamate, and this capacity is discussed later in the section on central metabolic pathways.

Potential virulence factor screening

Independent of the previous analysis, screening of all coding sequences (CDS) of the three species using Blast, SignalP, PSORT and TMHMM to look for similarities to genes reported to be associated with virulence in other bacteria identified a total of 235 genes with putative roles in virulence in B. pilosici 95/1000, compared to 303 in B. hydysenteriae WA1 and 142 in B. murdochii 56-150T. The number of genes in the different categories for all three Brachyspira species is shown in Table 3. Interpretation of the results for B. murdochii was complicated by the fact that the genome was incomplete, and additional genes are likely to be identified once a complete genome sequence becomes available. Overall, apart from what is shown, the predicted gene products did not have significant similarities with those of other well-characterized toxins or adhesins described in the major species of enteropathogenic bacteria - such as in E. coli or Clostridium species. Nevertheless, it is important to remember that other virulence-associated genes are likely to exist amongst those that are currently recorded as encoding proteins that are hypothetical or of unknown function.

It was of interest that B. hydysenteriae, the most pathogenic of the three species, contained more genes involved with lipopolysaccharide biosynthesis, motility and chemotaxis, and adhesion and/or surface proteins than did the other two species. The latter two did not have more of any category of these potential virulence-associated genes, apart from the bacteriophage genes in B. pilosici.

Lipooligosaccharides. The outer envelope of the Brachyspira species is considered to contain lipooligosaccharide (LOS) rather than lipopolysaccharide, based on the lack of a typical Gram-negative ladder formation of repeated O-sugar components on silver-stained SDS-PAGE gels [54,55]. LOS is considered to be involved in the induction of colonic lesions associated with B. hydysenteriae [56,57], and hence should be considered as potentially being involved in virulence in other Brachyspira species. LOS is also antigenic, and protective immunity to B. hydysenteriae is specific to the LOS serogroup [58]. The importance of LOS in relation to protective immunity to B. pilosici is less clear, as pigs experimentally infected with B. pilosici 95/1000 did not develop a systemic antibody response against the spirochete [59].

B. pilosici contained a set of 25 core genes involved in the biosynthesis of LOS (Table 3). This was fewer than the 30 genes identified for B. hydysenteriae, but more than the 10 so far identified in B. murdochii. The predicted pathways for biosynthesis of LOS and peptidoglycan for B. hydysenteriae and B. pilosici are shown in Figure 9. The main difference between the species related to the presence of an rfbBADC cluster in the virulence of ~36 kB plasmid in B. hydysenteriae WA1 that encode proteins for nucleotide sugar biosynthesis (dTDP-rhamnose) that are likely to be involved in O-antigen assimilation. A similar rfbBADC cluster was not found in B. pilosici or in B. murdochii, implying that they were unable to produce these antigens. Similar rfb gene clusters have been linked to virulence in numerous Gram negative bacteria [60,61], and, for example, are plasmid encoded in the case of some Salmonella enterica serovars [62]. Further work is required to examine the potential role of the rfb cluster in the virulence of B. hydysenteriae. It was interesting that despite the presence of genes encoding glycosyltransferases in this pathway, B. hydysenteriae does not appear to produce ladder-like O-antigens [54]. This discrepancy requires further investigation. The most precise way to describe the diversity of LPS/LOS in the Brachyspira species will be to make a detailed comparison of polysaccharide content and structures. It is unfortunate that no such structural data currently are available.

| Putative gene | B. pil | B. hyo | B. murd |
|---------------|-------|-------|--------|
| Chemotaxis    |       |       |        |
| Putative methyl-accepting chemotaxis protein | 6     | 10    | 7      |
| methyl-accepting chemotaxis protein A (mcpA) | 2     | 8     | 8      |
| methyl-accepting chemotaxis protein B (mcpB) | 6     | 19    | 16     |
| methyl-accepting chemotaxis protein C (mcpC) | -     | 3     | 1      |
| chemotaxis protein | 13 | 12 | 7 |
| Flagella      | 32    | 33    | 33     |
| Adhesion and/or surface protein |       |       |        |
| Lipoprotein   | 13    | 34    | 6      |
| Variable surface protein | 3   | 9    | 6      |
| Outer membrane protein | 6   | 4    | 2      |
| Integral membrane protein | 1   | 6    | -      |
| Inner membrane protein | 2   | 7    | -      |
| Host cell membrane degradation |       |       |        |
| Hemolysis     | 8     | 8     | 4      |
| Phospholipase | 2     | 4     | 3      |
| Peptidase     | 39    | 40    | 21     |
| Protease      | 15    | 15    | 14     |
| Oxidative stress (nox) | 2   | 2    | 2      |
| Phage         | 29    | 2     | -      |
| Ankyrin-like protein | 31 | 51   | 72     |
| Total         | 235   | 303   | 142    |

*Incomplete genome.

**core LOS biosynthesis genes. doi:10.1371/journal.pone.0011455.t003
Figure 9. Different biosynthetic pathways for lipooligosaccharide in *B. pilosicoli* 95/1000 and *B. hyodysenteriae* WA1. The gene name is shown in pink and corresponding predicted ORFs present in the *B. hyodysenteriae* genome are indicated in red, and those for *B. pilosicoli* in blue. doi:10.1371/journal.pone.0011455.g009
**Chemotaxis.** The capacity for chemotactic responses and motility are extremely important for *Brachyspira* species, allowing them to colonize specialized niches in the complex nutritional, physical and microbiological environment of the large intestine. Interestingly, a striking difference was found in the numbers of chemotaxis genes. *B. hyodysenteriae* contained 52 genes associated with chemotaxis, compared to 27 in *B. pilosicoli* and 39 in the partial genome of *B. murdochii* [Table 3]. The distribution of methyl-accepting chemotaxis genes also was different between the species, with *B. pilosicoli* having no mcpC genes. These differences are likely to be highly significant in relation to the chemotactic signals that the different species can respond to; they may explain the differences that have been observed in the attraction of *B. pilosicoli* and *B. hyodysenteriae* to mucin, and their tendency to occupy different local niches within the large intestine [63]. *B. hyodysenteriae* colonizes the lumen of the colon, but also penetrates deep into the colonic crypts, where it enters goblet cells. Although *B. pilosicoli* also can enter the crypts, unlike *B. hyodysenteriae* it penetrates the dense mucus layer at the crypt shoulders, and attaches to the underlying colonic enterocytes. As *B. hyodysenteriae* has a much more limited host range than *B. pilosicoli*, it seems paradoxical that it has a greater number and variety of methyl-accepting chemotaxis genes.

*B. pilosicoli* contained 13 chemosensory transducer genes, and it was interesting that seven of these, and genes for five hypothetical proteins, were located within a cluster at a ~5.7 Kb locus. This locus comprised cheB (BP951000_0460), cheD (BP951000_0461), cheR (BP951000_0462), cheW (BP951000_0463), cheF (BP951000_0464), cheT (BP951000_0465), five copies of hypothetical proteins (BP951000_0466 to BP951000_0470), and cheX (BP951000_0471). Similarly, six of the chemosensory transducer genes of *B. hyodysenteriae* were within one cluster at a locus of similar size. These genes encoded chemotaxis response regulator (cheR, BHWA1_00488), chemotaxis histidine kinase (cheA, BHWA1_00489), chemotaxis protein (cheW, BHWA1_00490), chemotaxis protein methyltransferase (cheR, BHWA1_00491), chemotaxis protein (cheD, BHWA1_00492), and response regulator receiver modulated CheB methylesterase (cheB, BHWA1_00493). The other chemosensory transduction genes were located elsewhere on the two genomes, and a similar clustering of che genes was not found in *B. murdochii*.

**Motility.** Only minor difference were found in the number and types of flagellar-associated genes between the three *Brachyspira* species (Table 3). *B. hyodysenteriae* WA1 and *B. murdochii* 56-150T contained four genes for FlaB core proteins while *B. pilosicoli* 95/1000 contained three. WA1 and 95/1000 contained three FlaA sheath protein genes while 50-150T contained two (Table 3). 50-150T also contained two copies of f1n, encoding flagellar synthesis regulator FlfN, while the other two strains had single copies. As expected, the three strains had many minor differences in the predicted sequences of the flagellar proteins. When considering the predicted flagellar structure, it should be remembered that *B. pilosicoli* only has 4–6 periplasmic flagella at each cell end compared to 7–14 for *B. hyodysenteriae* [1], and this difference could involve the need for some modifications. Furthermore, differences in flagellin protein ratios can affect the stiffness of the periplasmic flagella in spirochetes, and this stiffness directly correlates with their motility [64]. Hence, depending on expression of the flagellin proteins, the species are predicted to have different capacities for motility.

**Lipoproteins and surface proteins.** As with other bacteria, predicted surface associated proteins and lipoproteins of the *Brachyspira* species are likely to be important in interactions with the host; for example, cell surface proteins are thought to be involved with the polar attachment of *B. pilosicoli* to colonic enterocytes [65]. *B. pilosicoli* contained substantially fewer genes encoding lipoproteins than did *B. hyodysenteriae* (13 versus 34). Such lipoproteins, where they are surface associated, are likely to be important as potentially targets for immune recognition and could be used as subunit vaccine candidates [66]. An example of the difference between the two fully sequenced *Brachyspira* species was the presence of a gene encoding a peptidoglycan-associated outer membrane lipoprotein (PAL, BHWA1_02111) in *B. hyodysenteriae* that was absent in *B. pilosicoli*. PAL is involved in maintenance of the integrity of the cell envelope [67], and contributes to bacterial virulence and inflammation in Gram-negative sepsis [68]. Examples of lipoprotein-encoding genes that were shared between the species include those for basic membrane lipoprotein, outer membrane lipoprotein, apolipoprotein N acyltransferase, lipoprotein releasing system, transmembrane protein, LolC/E family, and lipoprotein releasing system ATP binding protein. The species had other predicted lipoproteins, but these were of unknown function.

*B. hyodysenteriae* has been reported to have a set of variable surface proteins of about 39 KDa [69], and strain WA1 contained nine genes encoding these proteins (VspA to VspF, and VspH), with two copies of VspA and VspD. Genes for VspA to VspF were located adjacent to each other (BHWA1_01456 to BHWA1_01601), whilst those for one copy of VspA and for VspH were located elsewhere on the genome (BHWA1_00889 and BHWA1_02382, respectively). *B. pilosicoli* contained single copies of the genes for VspD, VspE, and VspH located in different regions, while *B. murdochii* 56-150T had three copies of vspG, two of vspF and one of vspG. Differential expression of these genes may be involved in immune evasion, allowing persistence of the spirochetes [69].

Amongst the other outer membrane proteins, *B. pilosicoli* 95/1000 contained three copies of bspA, predicted to encode a BspA-like surface protein. BspA-like proteins are expressed on the surface of many pathogenic bacteria: they may bind to fibrinectin, stimulate a significant serological response [70], and induce secretion of IL-8 [71].

**Hemolysis-associated genes.** *B. hyodysenteriae* is strongly hemolytic, whereas *B. pilosicoli* and *B. murdochii* are weakly hemolytic [1]. The strong hemolytic activity of *B. hyodysenteriae* is thought to be an important virulence trait [72,73]. In addition to the seven genes associated with hemolysis previously identified in the genome of *B. hyodysenteriae* WA1 [15], an additional putative hemolysin gene was identified (BHWA1_00962).

Only single copies of thyA, thyB, thyC, and acpP (also known as hylA [73]) were identified in the partial genome of *B. murdochii*. On the other hand, *B. pilosicoli* 95/1000 contained eight genes potentially involved with hemolysis, of which seven were similar to those in *B. hyodysenteriae*. These encoded hemolysin A (thyA, BP951000_0123), ATP-dependent Clp protease proteolytic subunit (clpP, also known as thyD, BP951000_1802), hemolysin C (thyC, BP951000_1288), hemolysin related protein (hly, BP951000_0473), acyl carrier protein: contains beta-hemolysin (acpP or hylA, BP951000_0533), putative hemolysin III (BP951000_0424), and putative channel protein hemolysin III family protein (BP951000_0888). The eighth was a putative hemolysin (BP951000_2207) that was similar to a protein of *Parabacteroides distasonis* ATCC 8503, with ~42.9% protein identity. *B. pilosicoli* lacked a gene with similarity to that encoding the new putative hemolysin of *B. hyodysenteriae* (BHWA1_00962). From this, it is possible that the new putative hemolysin identified in *B. hyodysenteriae* is responsible for its strongly hemolytic phenotype. Alternatively, it has been suggested that differences between *B. hyodysenteriae* and the weakly hemolytic...
Brachyspira species such as B. pilosicoli in the FabF ACP synthase and FabG ACP reductase proteins that flank the acyl carrier protein (HlyA) may result in different lipid moieties being attached to the HlyA proteins [32]. Acyl carrier protein-dependant fatty acylation is known to be important in activation of prohemolysin to active hemolysin in E. coli [74], and such differences therefore may be important for expression of the strongly hemolytic phenotype in B. hyodysenteriae.

**Proteases.** B. hyodysenteriae WA1, B. pilosicoli 95/1000 and B. murdochii 56-150T contained similar numbers of genes predicted to encode proteases (15, 15 and 14, respectively). The proteolytic capacity in all three species was linked with the large number of ORFs encoding enzymes involved in uptake and metabolism of amino acids. These proteases are likely to be required for survival in the intestinal environment, but also may be involved in induction of local destruction of host tissues, and hence contribute to virulence. A serine protease previously described in B. pilosicoli (BP951000_0826, BP951000_1141 and BP951000_2083) [75] was found to have >92% sequence identity to proteases of B. murdochii 56-150T (ZP_04048074) and B. hyodysenteriae WA1 (YP_002721633).

**Secretion systems.** As with B. hyodysenteriae [15], B. pilosicoli lacked genes for the specialized secretion systems found in pathogenic Gram-negative bacteria, although it had genes for the Sec pathway and ABC-type transporters. Six sec genes were identified: secA (BP951000_0230), secY (BP951000_1006), secF (BP951000_1532), secD (BP951000_1533), secE (BP951000_1559) and secG (BP951000_1896), and there were 36 genes encoding ABC transporters (compared to 35 and 29 for B. hyodysenteriae and B. murdochii, respectively). Ten flagella-associated genes that can form part of a type III secretory system were found in B. pilosicoli 95/1000, but no needle associated genes that encode proteins required for the injection of toxins into the host cell were not identified.

**Central metabolic pathways**

A reconstruction of the central metabolic pathways of B. pilosicoli showed that it shared many metabolic capabilities with B. hyodysenteriae. The pathways for B. pilosicoli are shown in Figure 10, and can be compared with those for B. hyodysenteriae by examining Figure 2 in reference 15. In both species the glycolysis-gluconeogenesis metabolic axis constituted the backbone of energy production and the starting point of many biosynthetic pathways. The biosynthesis of peptidoglycan, phospholipids, aromatic amino acids, fatty acids and cofactors commences from pyruvate, or from intermediates in the glycolytic pathway. As with B. hyodysenteriae, a complete set of genes for the non-oxidative pentose phosphate pathway, nucleotide metabolism and a respiratory electron transport chain were identified in B. pilosicoli and B. murdochii. It can be seen from the figures that some of the obvious ways in which B. pilosicoli differed from B. hyodysenteriae related to the presence of pathways associated with citrate and 2-oxo-glutarate, a pathway from acetyl-CoA to ethanol, and a lack of the rfbBADC genes located on the B. hyodysenteriae plasmid (predicted to be involved in O-antigen biosynthesis).

![Figure 10. Central metabolic pathway construction for B. pilosicoli 95/1000.](https://www.plosone.org/doi/10.1371/journal.pone.0011455.g010)
Unlike *B. hyodysenteriae*, *B. pilosicoli* and *B. murdochii* both appeared to have the capacity to use enzymes of the incomplete TCA cycle to synthesize glutamate. Relevant predicted enzymes in the two species were putative citrate synthase (gefA, BP51000_0367 and 4083292.C12orf00527), putative isocitrate dehydrogenase (idh, BP51000_0968 and 4083292.C12orf00530), aconitase (BP51000_0369 and 4083292.C12orf00525), malate dehydrogenase (malh, BP51000_1737 and 4083292.C15orf00095), and two copies of fumarate hydratase in *B. pilosicoli* (funJ, BP51000_1272 and BP51000_1273; funA was not identified in *B. murdochii*, possibly due to the genome being incomplete. This capacity emphasized the dual function of the TCA portion of the cycle: provision of NADH as part of the complete cycle in aerobes, and provision of alpha-oxoglutarate in aerobes and anaerobes. Aconitase allows the TCA cycle to function during periods of oxidative stress, and may play a role in mediating oxidative stress and regulating cell death [53,76]. The presence of malate dehydrogenase as an alternative enzyme for converting oxaloacetate into malate is likely to provide a source of malate for cell biosynthesis in *B. pilosicoli* and *B. murdochii*.

A key branch in the partial TCA cycle of *B. pilosicoli* and *B. murdochii* involved 2-oxoglutarate, succinyl-CoA and fumarate, precursors that could proceed in either the oxidative or reductive direction. Unlike *B. hyodysenteriae*, *B. pilosicoli* and *B. murdochii* also had a gene set for the production of ethanol from acetyl-CoA via a branched fermentation pathway. The significance of this difference between the species is unknown, but it is likely to influence overall metabolic flexibility.

*B. pilosicoli* can be distinguished from all the other *Brachyspira* species in that it lacks beta-glucosidase activity [77]. Consistent with this, no genes specifically encoding beta-glucosidase were identified in the *B. pilosicoli* genome; however, a novel system for metabolizing multiple sugars, including beta-glucosides was identified that involved alpha-galactosidase/6-phospho-beta-glucosidase (alphaGlc, BP51000_0253), PTS system, arbutin-like IIC component (gbG, BP51000_0254) and 6-phospho-alpha-glucosidase (gbA, BP51000_0255). These enzymes have specificity for 6-phospho-beta-D-glycoside substrates rather than the beta-D-glycoside substrates of beta-glucosidases [78,79]. Genes associated with beta-glucoside metabolism are likely to be important in *Brachyspira* species as they are known to regulate growth, adhesion and colonization in other bacterial species [80].

*B. hyodysenteriae* and *B. murdochii* had genes encoding phosphoenolpyruvate synthase, whereas *B. pilosicoli* did not appear to have the capacity to interconvert phosphoenolpyruvate (PEP) and pyruvate. Instead, it may utilize glyceral but not pyruvate for growth, and use pyruvate kinase (pyk, BP51000_0367 and 4083292.C12orf00530) to catalyse the reverse conversion of PEP to pyruvic acid [81]. This process occurs more significantly under microaerobic conditions, with a corresponding increase in the rate of constructive metabolism.

*B. hyodysenteriae* and *B. murdochii* possessed a complete set of genes for histidine metabolism, whereas *B. pilosicoli* did not. It is possible that *B. pilosicoli* has an alternative pathway for synthesizing histidine. It may however also obtain nutrients from the surrounding environment, particularly as the intestine is a rich source of amino acids, including histidine [82]. Consistent with this requirement, three copies of a polar amino acid uptake ABC transporter, PAAT family, amino acid-binding protein (BP51000_0968, BP51000_0969 and BP51000_0990) were identified in *B. pilosicoli*, and not in the other two species.

A complete fatty acid biosynthesis gene set previously has been found in *B. hyodysenteriae* [15], and was identified here in *B. murdochii*. In contrast, an incomplete gene set was identified in *B. pilosicoli*, with only the genes for (3R)-hydroxymyristoyl-acetyl carrier protein dehydratase (fabZ, BP51000_0773), malonyl CoA acyl carrier protein transacylase (fadB), BP51000_1348), and the previously mentioned loci comprising ACP synthase II (fabF, BP51000_0532), acyl carrier protein (acpP, BP51000_0533) and acyl reductase (fabG, BP51000_0534) being identified while FabZ provides the 3-hydroxyxymyristate moieties characteristic of lipid A. FabD is involved in elongation reactions in fatty acid synthesis. These findings suggested that *B. pilosicoli* does not incorporate carbon from simple carbon sources into fatty acids.

**Implications.** In this study the complete genome sequence of *B. pilosicoli*, a widespread but somewhat neglected enteric pathogen of humans and animals, was obtained and made available for analysis by the scientific community. The spirochete genome was shown to have basic similarities to the genomes of the related species *B. hyodysenteriae* and *B. murdochii*, but a number of distinct features were identified that may help to explain the different host ranges and colonization sites of the species. These included predicted differences in chemotactic responses, oxygen sensitivity, the use of some different substrates, the presence of different surface proteins, and predicted differences in LOS content and structure. The first description of genes for an apparently complete bacteriophage in *B. pilosicoli* also was made: this finding has important practical implications, since if the bacteriophage can be shown to be functional it could be used for genetic manipulation of *B. pilosicoli*. The current lack of such a genetic tool for *B. pilosicoli* is a constraint that limits functional studies of the genes that have been identified.

**Supporting Information**

**Figure S1** Whole-genome comparisons between *B. pilosicoli* 95/1000 and *B. hyodysenteriae* WAI displayed using the Artemis Comparison Tool. Regions of DNA:DNA similarity are joined by lines. Found at: doi:10.1371/journal.pone.0011455.s001 (2.24 MB TIF)

**Acknowledgments**

The authors thank Drs Peter J. Wilson and Annette McGrath from the Australian Genome Research Facility for undertaking the *B. pilosicoli* sequencing and for initial bioinformatics analysis in support of the sequencing effort. We also wish to acknowledge the generosity of Dr S. Lucas and colleagues for making the partial genome sequence of *B. murdochii* DSM 12563 freely available in GenBank.

**Author Contributions**

Conceived and designed the experiments: MIB TL DJH. Performed the experiments: PW TL. Analyzed the data: PW MIB TL KR PMM BC MB. Contributed reagents/materials/analysis tools: MIB TL DJH. Performed the experiments: PW TL. Contributed reagents/materials/analysis tools: DSM 12563 freely available in GenBank.

**References**

1. Stanton TB (2006) The genus *Brachyspira*. In: Falkow S, Rosenberg SE, Schleifer KH, Stackebrandt E, eds. The Prokaryotes (Volume 7). New York: Springer. pp 330–356.

2. Lee JI, McLaren AJ, Lymberry AJ, Hamson DJ (1993) Human intestinal spirochaete are distinct from *Serpulina hyodysenteriae*. J Clin Microbiol 31: 16–21.

3. Trotz DJ, Stanton TB, Jensen NS, Duhamel GE, Johnson JL, et al. (1996) *Serpulina pilosicoli* sp. nov., the agent of porcine intestinal spirochetosis. Int J Syst Bacteriol 46: 206–215.

4. Trotz DJ, Mikosza ASJ, Combs BG, Oxberry SL, Hamson DJ (1998) Population genetic analysis of *Serpulina pilosicoli* and its molecular epidemiology.

5. Trott DJ, Mikosza ASJ, Combs BG, Oxberry SL, Hamson DJ (1998) Population genetic analysis of *Serpulina pilosicoli* and its molecular epidemiology.
in villages of the Eastern Highlands of Papua New Guinea. Int J Syst Bacteriol 48: 659–668.

5. Ochiai S, Adachi Y, Mori K (1997) Unification of the genera Serpulina and Brachyspira, and proposals of Brachyspira hyodysenteriae comb. nov., Brachyspira aalborgi comb. nov., and Brachyspira pilosicoli comb. nov. Microbiol Immunol 41: 445–452.

6. Hampson DJ, Oxberry SL, La T (2006) Potential for zoonotic transmission of Brachyspira pilosicoli. Emerg Infect Dis 12: 899–907.

7. Lee J, Hampson DJ (1992) Intestinal spirochaetes colonising Aborigines from communities in the remote north of Western Australia. Epidemiol Infect 109: 131–144.

8. Trott DJ, Combos BG, Oxberry SL, Mikosz ASJ, Robertson ID, et al. (1997) The prevalence of Serpulina pilosicoli in humans and domestic animals in the Eastern Highlands of Papua New Guinea. Epidemiol Infect 119: 369–379.

9. Margawan KR, Robertson ID, Brooke CJ, Hampson DJ (2004) Prevalence, risk factors and molecular epidemiology of Brachyspira pilosicoli in humans on the island of Bali, Indonesia. J Med Microbiol 53: 325–332.

10. Munshi MA, Traub RJ, Robertson ID, Mikosz ASJ, Hampson DJ (2004) Colonization and risk factors for Brachyspira aalborgi and Brachyspira pilosicoli in humans and dogs on tea-estates in Assam, India. Epidemiol Infect 132: 137–144.

11. Nelson EG, Tanumihardjo A, Cheddar B, Kane AV, Quest M, et al. (2009) High prevalence of spirochaetosis in cholera patients, Bangladesh. Emerg Infect Dis 15: 571–573.

12. Trivett-Moore NL, Gilbert GL, Law CLH, Trott DJ, Hampson DJ (1998) Recognition of two new species of intestinal spirochetes: Brachyspira aalborgi sp. nov. and Brachyspira pilosicoli sp. nov. Int J Syst Bacteriol 48: 1007–1012.

13. Hampson DJ, La T (2006) Reclassification of Serpulina intermedia and Serpulina murchardi as Brachyspira aalborgi intermedia comb. nov. and Brachyspira murchardi comb. nov. Int J Syst Evol Microbiol 56: 1009–1012.

14. Weisensee H, Maderner A, Herzog AM, Lassy H, Nowotny N (2005) Amplification and sequencing of Brachyspira spp. specific portions of nor using polymerase-embedded tissue samples from clinical colitis in Austrian pigs show frequent solitary presence of Brachyspira murchardi. Vet Microbiol 111: 67–75.

15. Jemsen TK, Christensen AS, Boye M (2010) Brachyspira murchardi colitis in pigs. Vet Pathol 47: 334–338.

16. Joens LA, Sahlberg E, Kiiskinen A, Cuervo J, Brismar K (2003) Effect of diet and vaccination on colonisation of pigs with the intestinal spirochaete Brachyspira pilosicoli. In: Current Protocols in Molecular Biology. New York: Wiley. pp 241–242.

17. Lichtensteiger CA, Vimr ER (1997) Neuraminidase (sialidase) activity of Helicobacter pylori. Arch Microbiol 175: 8–18.

18. Greer JM, Wannemuehler MJ (1989) Comparison of the biological responses to two strains and their serological cross-reactivities. J Med Microbiol 48: 411–415.

19. Finkmeier M, Okazako N, Mifuchi I, Yanagihara Y (1990) Inhibition of colonization by sulfated gastric mucin. Biochem Int 24: 324–328.

20. Whitfield C (1995) Biosynthesis of lipopolysaccharide O antigens. Trends Microbiol 3: 317–323.

21. Greer JM, Wannemuehler MJ (1991) Comparative analysis of the alphal,1,2-fucosyltransferase gene. Mol Microbiol 7: 773–778.

22. Nordgren R, Brismar K, Hulten H, Monin H, Bucht G (1994) Attraction of Brachyspira pilosicoli to mucin. Microbiology 140: 191–197.

23. Hulten H, Brismar K, Nordgren R (1995) Detection of Brachyspira pilosicoli by ELISA in the blood of critically ill patients. J Clin Microbiol 35: 482–485.

24. Hampson DJ, Robertson ID, Oxberry SL (1999) Isolation of Serpulina murchardi from the joint fluid of a lame pig. Aust Vet J 77: 48.

25. Price J, Stenlund A, Aufricht V, Fekete Z, Steenbock BL (1999) Inhibition of Helicobacter pylori colonization by sulfated gastric mucin. Biochem Int 41: 749–756.

26. Wang G, Rasko DA, Sherburne R, Taylor DE (1999) Molecular genetic basis for the variable expression of Lewis Y antigen in Helicobacter pylori of humans and dogs by electrophoretic analysis of the alpha(1,2)fucosyltransferase gene. Mol Microbiol 31: 1265–1274.

27. Wang G, Boulon PG, Chan NW, Palic MM, Taylor DE (1999) Novel Helicobacter pylori alpha 1,2-fucosyltransferase, a key enzyme in the synthesis of Lewis Y antigens. Microbiology 145: 3245–3251.

28. Whitfield C (1996) Biosynthesis of lipopolysaccharide O antigens. Trends Microbiol 4: 347–351.

29. Waibel K, Kehres B, Welti B, Kühn D, Porschen R, et al. (2004) Inhibition of mouse leukocyte adhesion by the Helicobacter pylori alpha 1,2-fucosyltransferase gene. Mol Microbiol 52: 1383–1395.

30. Wang G, Boulton PG, Chan NW, Palic MM, Taylor DE (1999) Novel Helicobacter pylori alpha 1,2-fucosyltransferase, a key enzyme in the synthesis of Lewis Y antigens. Microbiology 145: 3245–3251.

31. Lichtensteiger CA, Vimr ER (1997) Neuraminidase (sialidase) activity of Helicobacter pylori. Arch Microbiol 175: 8–18.

32. Novotny N, Arazm M, Wisk J, Reddy B, et al. (2006) Bacterial neuraminidase facilitates mucosal infection by participating in biofilm production. J Clin Invest 116: 2297–2305.

33. Doares SG (1988) Lipooligosaccharides from Brachyspira pilosicoli and Brachyspira aalborgi. In: Brent R, Kingston RE, Moore DD, Smith JA, Seidman JG, Struhl K, eds. The Bacteria and Archaea. Cold Spring Harbor: Cold Spring Harbor Press. pp 1–18.

34. Eckardt NA (2001) Everything in its place: Conservation of gene order among Listeria monocytogenes. Mol Microbiol 40: 2297–2305.
65. Trott DJ, Alt DP, Zaerner RL, Wannemuehler MJ, Stanton TB (2001) The search for Brachyspira outer membrane proteins that interact with the host. Anim Health Res Rev 2: 19–30.

66. Movahedi A, Hampson DJ (2008) New ways to identify novel bacterial antigens for vaccine development. Vet Microbiol 131: 1–13.

67. Rodriguez-Herva JJ, Ramos-Gonzalez MI, Ramos JL (1996) The Pseudomonas putida peptidoglycan-associated outer membrane lipoprotein is involved in maintenance of the integrity of the cell envelope. J Bacteriol 178: 1699–706.

68. Hellman J, Roberts JD, Jr., Trehar MM, Allaire JE, Warren HS (2002) Bacterial peptidoglycan-associated lipoprotein is released into the bloodstream in gram-negative sepsis and causes inflammation and death in mice. J Biol Chem 277: 14274–14280.

69. McCannan MT, Auer K, Foley W, Gabe JD (2003) Brachyspira hyodysenteriae contains eight linked gene copies related to an expressed 39-kDa surface protein. Microbes Infect 5: 1–6.

70. Sharma A, Sojar HT, Gmurich I, Hoorna K, Kuramitsu HK, et al. (1998) Cloning, expression, and sequencing of a cell surface antigen containing a leucine-rich repeat motif from Bacteroides forsythus ATCC 43037. Infect Immun 66: 5703–5710.

71. Onishi S, Honma K, Liang S, Stathopoulou P, Kirane D, et al. (2008) Toll-like receptor 2-mediated interleukin-8 expression in gingival epithelial cells by the Tannerella forsythia leucine-rich repeat protein BspA. Infect Immun 76: 198–205.

72. ter Huurne AA, Gaastra W (1995) Swine dysentery: More unknown than known. Vet Microbiol 46: 347–360.

73. Hsu T, Hutto D, Minion FC, Zaerner RL, Wannemuehler MJ (2001) Cloning of a beta-hemolysin gene of Brachyspira (Serpulina) hyodysenteriae and its expression in Escherichia coli. Infect Immun 69: 706–711.

74. Isartel JP, Koronakis V, Hughes C (1991) Activation of Escherichia coli prohaemolysin to the mature toxin by acyl carrier protein-dependent fatty acylation. Nature 351: 759–761.

75. Muniappa N, Dukhanel GE (1997) Outer membrane-associated serine protease of intestinal spirochetes. FEBS Microbiol Lett 154: 159–164.

76. Moeder W, Del Pazo O, Navarro DA, Martin GR, Klessig DF (2007) Aconitase plays a role in regulating resistance to oxidative stress and cell death in Arabidopsis and Nicotiana benthamiana. Plant Mol Biol 63: 273–287.

77. Fellstrom C, Granum A (1995) Phenotypical characterization of intestinal spirochetes isolated from pigs. Rev Sci Tech 14: 1–4.

78. Anderson RL (1972) Cellobiose metabolism in Aerobacter aerogenes. 3. Cleavage of cellobiose monophosphate by a phospho-beta-glucosidase. J Biol Chem 247: 3420–3423.

79. Dale M, Kopfler W, Chait I, Byers L (1986) Beta-glucosidase: substrate, solvent, and viscosity variation as probes of the rate-limiting steps. Biochemistry 25: 2522–2529.

80. Kilic AO, Tao L, Zhang Y, Lei Y, Khammanivong A, et al. (2006) Involvement of Streptococcus gordonii beta-glucoside metabolism systems in adhesion, biofilm formation, and in vivo gene expression. J Bacteriol 188: 4246–4253.

81. Podkopaeva DA, Grabovich M, Dubinina GA (2003) Oxidative stress and antioxidant cell protection systems in the microaerophilic bacterium Spirillum vineale. Mikrobiologia 72: 600–608.

82. Julio AO, Marris CF, Xie J, Gisedorf JR (2007) Histidine auxotrophy in communal and disease-causing nontypeable Haemophilus influenzae. J Bacteriol 189: 4994–5001.