The Journal of Veterinary Medical Science

FULL PAPER

Bacteriology

Phylogeny and putative virulence gene analysis of *Bartonella bovis*

Sun Tee TAY[1)]*, Kai Ling KHO[1)], Siew Fen LYE[2)] and Yun Fong NGEOW[3)]

[1)]Department of Medical Microbiology, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia

[2)]BioEasy Sdn Bhd. Setia Avenue, 33A-3, Jalan Setia Prima S, U13/S, Setia Alam, Seksyen U13, 40170 Shah Alam, Selangor, Malaysia

[3)]Department of Pre-Clinical Sciences, Faculty of Medicine and Health Sciences, Universiti Tunku Abdul Rahman, 43000 Kajang, Selangor DE, Malaysia

ABSTRACT. *Bartonella bovis* is a small Gram-negative bacterium recognized as an etiological agent for bacteremia and endocarditis in cattle. As few reports are available on the taxonomic position of *B. bovis* and its mechanism of virulence, this study aims to resolve the phylogeny of *B. bovis* and investigate putative virulence genes based on whole genome sequence analysis. Genome-wide comparisons based on single nucleotide polymorphisms (SNP) and orthologous genes were performed in this study for phylogenetic inference of 27 *Bartonella* species. Rapid Annotation using Subsystem Technology (RAST) analysis was used for annotation of putative virulence genes. The phylogenetic tree generated from the genome-wide comparison of orthologous genes exhibited a topology almost similar to that of the tree generated from SNP-based comparison, indicating a high concordance in the nucleotide and amino acid sequences of *Bartonella* spp. The analyses show consistent grouping of *B. bovis* in a cluster related to ruminant-associated species, including *Bartonella australis*, *Bartonella melophagi* and *Bartonella schoenbuchensis*. RAST analysis revealed genes encoding flagellar components, in corroboration with the observation of flagella-like structure of BbUM strain under negative straining. Genes associated with virulence, disease and defence, prophages, membrane transport, iron acquisition, motility and chemotaxis are annotated in *B. bovis* genome. The flagellin (*flaA*) gene of *B. bovis* is closely related to *Bartonella bacilliformis* and *Bartonella clarridgeiae* but distinct from other Gram-negative bacteria. The absence of type IV secretion systems, the *bona fide* pathogenicity factors of bartonellae, in *B. bovis* suggests that it may have a different mechanism of pathogenicity.

KEY WORDS: *Bartonella bovis,* genome-based phylogeny, putative virulence genes

J. Vet. Med. Sci.
80(4): 653–661, 2018
doi: 10.1292/jvms.17-0448

Received: 19 August 2017
Accepted: 14 December 2017
Published online in J-STAGE: 29 December 2017

Bartonella species are small, fastidious Gram-negative bacteria which are responsible for a wide variety of clinical syndromes in humans and animals [9–11]. The bacteria infect mammalian erythrocytes and endothelial cells and are usually transmitted through the bites of hematophagous arthropods, such as fleas, lice, flies, and ticks [11]. Humans acquire *Bartonella henselae* or *Bartonella clarridgeiae* infections following scratches or bites by cats infested with flea (*Ctenocephalides felis*). *Bartonella bacilliformis,* the etiologic agent of Carrion's disease or Oroya fever, is transmitted by the sandflies of the genus *Lutzomyia*. The transmission of trench fever (caused by *Bartonella quintana*) occurs through the bites of infected human louse (*Pediculus humanus*) [2]. Rodent-borne *Bartonella* spp., such as *Bartonella elizabethae, Bartonella tamiae, Bartonella rattimassiliensis*, and *Bartonella tribocorum* have been associated with zoonotic diseases [31, 32].

Several *Bartonella* spp. including *Bartonella bovis, Bartonella melophagi, Bartonella schoenbuchensis,* and *Bartonella australis* have been isolated from ruminants. Cattle have been reported as the main reservoir for *B. bovis*. The prevalence of *B. bovis* in cattle is generally high but varies widely across studies from different continents [5]. Long-term persistence of *B. bovis* and adaptation to cattle has been postulated to mimic *Bartonella* infection in cats and mice [37]. As the potential role of *B. bovis* as a zoonotic pathogen has been suggested [13, 43], further studies on the genomic and virulence of this bacterium is required. Several *B. bovis* strains have been identified using multilocus sequence typing (MLST) approach in a Malaysian study [27]. One of the strains (herein referred to as BbUM) isolated from a dairy cow, has been identified as MLST ST27 and proposed under a new lineage (IIa) of *B. bovis* [27].

In this study, the draft genome sequence of the BbUM strain was determined to enable further study on the phylogeny and prediction of virulence genes. To verify the species status of BbUM strain, average nucleotide identity (ANI) and tetranucleotide

*Correspondence to: Tay, S. T.: tayst@um.edu.my

usage patterns (TETRA) analyses were performed [20, 28, 54]. ANI determination is based on pair-wise comparisons of sequences shared between two strains. The approach has been used widely for classification of bacterial species [24]. For species definition, ANI values >94% correspond to 70% DNA–DNA re-association standard [20, 29, 30]. TETRA analyzes tetranucleotide usage patterns in genomic fragments, and a correlation co-efficient value of above 0.99 suggests the probability of two strains as the same species [51]. Previously, the genetic classification of *Bartonella* spp. has been based on sequence analysis of 16S rDNA and several protein-encoding genes [25, 39, 42]. However, the high degree of 16S rDNA sequence similarity provides very little statistical support for the determination of phylogenetic relationship within the genus *Bartonella* [42]. Additionally, the prediction of phylogenetic relationship between species can be confusing when two or more protein genes are analysed [25].

With the availability of whole genome sequences for most *Bartonella* spp., comparative genome-wide analyses based on single nucleotide polymorphisms (SNP) and orthologous genes are now possible. In a previous study [23], the phylogenetic relationship of 16 *Bartonella* species has been analyzed based on orthologous core genes. This study reports the phylogenetic analyses of 27 *Bartonell*a species based on SNP and orthologous genes. Additionally, putative virulence genes are identified by using Rapid Annotation using Subsystem Technology (RAST) analysis.

MATERIALS AND METHODS

Growth conditions and whole genome sequencing of BbUM strain

BbUM strain was cultured on fresh Columbia agar supplemented with 5% sheep blood (Isolab, Shah Alam, Malaysia) at 37°C in 5% CO_2 atmosphere for four days. After Gram staining, the strain was observed under light microscopy. Bacterial morphology was observed using transmission electron microscopy (EM JEM1200EX; JEOL, Tokyo, Japan) at 100 kV with negative staining. The colonies were suspended in lysis buffer for DNA extraction using a QIAmp DNA mini kit (Qiagen, Hilden, Germany). The concentration and purity of the eluted DNA were measured using a NanoDrop™ 2000 Spectrophotometer (Thermo Scientific, Wilmington, DE, U.S.A.) and a Qubit fluorometer (Invitrogen Corp., Carlsbad, CA, U.S.A.). The DNA sample was sequenced using Illumina platform with 100 bp paired-end reads. The reads were then subjected to adapters and low quality reads trimming based on Q25 using Trimmomatics [8] and Sycthe (https://github.com/vsbuffalo/scythe). String Graph Assembler (SGA) [50] was used for error correction. After pre-analysis, the reads were subjected to *de novo* genome assembly, using IDBA-UD, a de novo assembler for sequencing data with highly uneven depth [41]. Only contigs larger than 200bp were used for downstream analysis. The assembled contigs were subjected to left over adapters and contaminant filter against NCBI univec database via seq clean (default parameter) (http://hpc.ilri.cgiar.org/seqclean/seqclean). The whole genome sequencing project has been submitted to the GenBank database (accession no: MWVG00000000).

Inter-strain comparisons with ANI and TETRA analyses

Analyses to determine average nucleotide identity (ANI) [20] and tetra-nucleotide usage patterns TETRA (http://www.megx.net/tetra) [54] were performed by using JSpecies v1.2.1 [44], using 27 *Bartonella* genome sequences available in the GenBank database. Table 1 is the list of *Bartonella* species included in this analysis.

Investigation of phylogenetic relationships amongst Bartonella species

In this study, genome-wide SNP and orthologous gene analyses were performed for inference of phylogenetic relationship amongst *Bartonella* species. i) Genome-wide SNP-based comparison. Completed and draft sequences downloaded from NCBI (Table 1) were subjected to a series of pre-processing to generate a FASTA input file compatible with the analytic software. The analysis was carried out using the software package kSNP3.0, in which SNP discovery was based on k-mer analysis, using single variant position within sequences of nucleotide length k [19]. The kchooser script was used to select an optimum value of k-mer size for the data set. Jellyfish [38] in the kSNP3.0 package was used to enumerate a list of k-mers at optimum k-value for each genome in the data set. The sequence variation of the data set was evaluated and the fraction of core-k-mers at optimum k-value was identified. kSNP was used to search for putative SNP loci in the data set and the identified SNP alleles were concatenated into a string. Multiple sequence alignment was produced for each nucleotide in the SNP matrix using MUSCLE (version 3.8.31) [18]. Best fit nucleotides substitution model was selected using jModelTest [17, 21] and trees of different models were assessed using Akaike Information Criterion framework (AIC) [1]. PhyML 3.0 [22] was used to generate maximum likelihood tree with 100 bootstrap replicates using the jModelTest suggested model. ii) Genome-wide comparison of orthologous genes. The predicted proteins from BbUM genome were compared to proteins from other *Bartonella* species and *Brucella abortus* (Table 1) using OrthoMCL (version 2.0.9) [14, 35]. Subsequently, shared proteins (core proteins) of three *B. bovis* strains (*B. bovis* m02, 91-4 and BbUM) were analysed. A Venn diagram was drawn to show the distribution of genes amongst three *B. bovis* strains. Orthologue groups with single-copy genes in all species in the study were identified. For each orthologue group, the protein sequences for all species were extracted and aligned using MUSCLE (version 3.8.31) [18]. Each multiple sequence alignment was trimmed using trimAl (version 1.4) [12] before merging into one long sequence alignment of orthologous proteins using a custom script. The alignment was then subjected to substitution model tests using ProtTest (version 3.4 release 20140123) [16] and phylogeny estimation was performed using PhyML (release 20111216) [21]. Owing to the large alignment size, branch support for each node in the protein evolution estimation was tested with the approximate likelihood ratio test (aLRT) [3] instead of bootstrap. Trees were generated separately using three supported aLRTs in PhyML, i.e., aLRT, Chi2-based aLRT and SH-like aLRT.

Table 1. List of *Bartonella* species and outgroup selected for SNP and orthologous gene analysis

Genome	Accession number	Source	Cluster (SNP)[a)]	Cluster (Orthologous gene)[b)]
B. quintana RM-11	NC_018533	Monkey (*Rhesus macaques)*	1	I
B. henselae MVT02	NZ_LN879429	Human (*Homo sapiens*)	1	I
B. henselae str. Houston-1	NC_005956	Human (*Homo sapiens*)	1	I
B. koehlerae C-29	NZ_AHPL00000000	Cat (*Felis domesticus*)	1	I
B. senegalensis OS02	NZ_CALV00000000	Soft tick (*Ornithodoros sonrai*)	1	I
B. elizabethae (BeUM)	LFMF00000000	Wild rat (*Rattus diardii*)	2	II
B. elizabethae F9251	NZ_AIMF00000000	Human (*Homo sapiens*)	2	II
B. elizabethae Re6043vi	NZ_AILW00000000	Polynesian rat (*Rattus exulans*)	2	II
B. queenslandensis AUST/NH15	NZ_CALX00000000	Wild rat (*Rattus leucopus*)	2	II
B. rattimassiliensis 15908	NZ_AILY00000000	Wild rat (*Rattus norvegicus*)	2	II
B. tribocorum BM1374166	NZ_HG969192	Wild rat (*Rattus norvegicus*)	2	II
B. tribocorum CIP 105476	NC_010161	Wild rat (*Rattus norvegicus*)	2	II
B. alsatica IBS 382	NZ_AIME00000000	Wild rabbit (*Oryctolagus cuniculus*)	2	IV
B. vinsonii spp. *berkhoffii*	NC_020301	Dog (*Canis lupus*)	2	IV
B. birtlesii IBS 325	NZ_CM001557	Mouse (*Apodemus* spp.)	3	IV
B. taylorii 8TBB	NZ_AIMD00000000	Vole (*Microtus agrestis*)	3	IV
B. washoensis Sb944nv	NZ_AILU00000000	California ground squirrel (*Spermophilus beecheyi*)	3	I
B. doshiae NCTC 12862	NZ_AILV00000000	Vole (*Microtus agrestis*)	3	-
B. australis Aust/NH1	NC_020300	Kangaroo (*Macropus giganteus*)	3	III
B. bacilliformis KC583	NC_008783	Human (*Homo sapiens*)	3	III
B. bovis (BbUM)	MWVG00000000	Crossbred cattle (Mafriwal)	3	III
B. bovis 91–4	NZ_CM001844	Cattle (*Bos taurus*)	3	III
B. bovis m02	NZ_AGWB00000000	Moose (*Alces alces*)	3	III
B. clarridgeiae 73	NC_014932	Cat (*Cattus felis*)	3	III
B. melophagi K-2C	NZ_AIMA00000000	Sheep ked (*Melophagus ovinus*)	3	III
B. rochalimae ATCC BAA-1498	NZ_AHPK00000000	Human (*Homo sapiens*)	3	III
B. schoenbuchensis m07a	NZ_AGWC00000000	Moose (*Alces alces*)	3	III
Brucella abortus 2308	NC_007618 (Chr 1), NC_007624 (Chr 2)	Aborted fetus of a cow	N/A	N/A

a) refer to Fig. 2 (i); b) refer to Fig. 2 (ii); N/A, not available.

Annotation and gene prediction

The contigs were subjected to gene prediction using Prokaryotic Dynamic Programming Genefinding Algorithm (PRODIGAL) Version 2.60 [26]. Annotation of predicted coding sequences was performed by homology search against the nr database. tRNA and rRNA were predicted using ARAGORN [34] and RNAmmer [33], respectively. The assembled contigs were annotated using RAST (Rapid Annotation using Subsystem Technology) pipeline [4].

Analysis of flagella (flaA) gene

Two *flaA* sequences (type 1 and 2) of BbUM strain were retrieved based on the data generated from RAST analysis and compared for similarity with other flagella sequences in the GenBank database using BioEdit Sequence Alignment Editor Software (Version 7.0.5.3). A dendrogram was constructed based on the sequences using neighbour-joining method of the MEGA software and bootstrap analysis with 1,000 re-samplings [52]. *E. coli* flagellin (*fli*C, NC000913) was used as an outgroup for comparison.

RESULTS

B. bovis BbUM strain demonstrated small (1–2 mm in diameter), smooth and greyish colonies upon culturing on blood agar after four days of incubation. Electron microscopy of negative staining reveals flagella-like structure at one pole and a thin slime layer surrounding the cell body of the BbUM strain (measuring 1.45 μm in length and 0.59 μm in width) (Fig. 1).

The assembly from the sequencing reads of BbUM strain generated 331 contigs with an average length of 4,977 bp (200–406,386 bp). N50, the length of a collection of contigs that covers at least 50% of the genome assembly, was 158,089 bp. The draft genome of BbUM strain has a total length of 1,647,450 bp and a GC content of 37.7%. BbUM strain shares the highest values of ANI (98.12%) and TETRA (0.99438) with *B. bovis* 91–4 type strain, followed by *B. bovis* m02 (Suppl Table 1), thus placing the strains under the same species. The next genetically closely related species with *B. bovis* were *B. melophagi* K-2C, and *B. schoenbuchensis* m07a (Suppl. Table 1), two *Bartonella* species originated from sheep ked and moose, respectively.

Figure 2 (i) shows the SNP-based phylogenetic tree of *Bartonella* spp. using *B. abortus* 2308 as an outgroup. The SNP-based

phylogenetic tree shows the differentiation of *Bartonella* spp. into three distinct clusters (1–3) (Table 1). Cluster 1 consists of two cat-associated species, i.e., *B. henselae* and *B. koehlerae*, *B. quintana* and a tick-derived *B. senegalensis*. Cluster 2 is mainly comprised of rodent-associated *Bartonella* spp. i.e., *B. elizabethae*, *B. tribocorum*, *B. queenslandensis*, and *B. rattimassiliensis*, a dog-derived (*B. vinsonii*) and a rabbit-derived *Bartonella* spp. (*B. alsatica)*. Cluster 3 has the largest number of *Bartonella* spp., with four ruminant-associated species (*B. australis*, *B. bovis*, *B. melophagi* and *B. schoenbuchensis*), four rodent-associated species (*B. washoensis*, *B. birtlesii*, *B. doshiae* and *B. taylorii*) and three human pathogens (*B. rochalimae*, *B. clarridgeiae*, and *B. bacilliformis*).

In the analysis of orthologous genes, a total of 2,690 orthologue groups were identified, of which, 1,299 groups contained at least one *B. bovis* (*B. bovis* m02, *B. bovis* 91–4 and BbUM) protein. There was a high number of shared orthologue groups (1,092), suggesting the high degree of similarity amongst *B. bovis* strains. The protein evolution estimation using three different aLRTs produced the same tree with different branch-support values. Figure 2 (ii) shows the dendrogram constructed based on SH-like aLRT branch support. Four clusters (I-IV) were obtained. The cat-associated species (*B. henselae* and *B. koehlerae*), *B. quintana* and a tick-derived *B. senegalensis* are grouped in the same phylogenetic group (cluster I) with *B. washoensis*. Cluster II is mainly comprised of rodent-associated *Bartonella* spp. (*B. elizabethae*, *B. tribocorum*, *B. queenslandensis*, and *B. rattimassiliensis*). BbUM strain is placed in the same phylogenetic group (cluster III) with *B. bovis* 91–4 strain and other ruminant-associated *Bartonella* spp. (*B. australis*, *B. bovis*, *B. melophagi*, and *B. schoenbuchensis*) and three human pathogens (*B. rochalimae*, *B. clarridgeiae*, *B. bacilliformis*), similar to the results derived from SNP-based phylogenetic analysis.

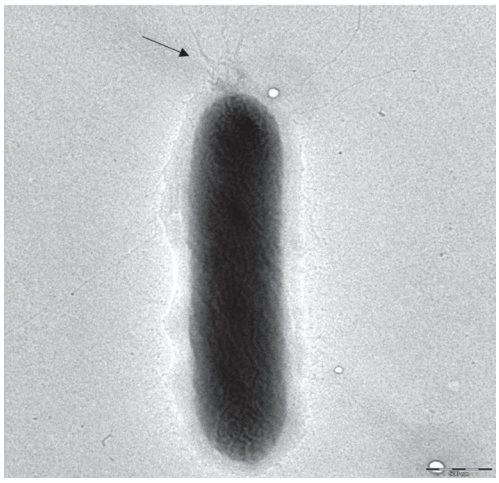

Fig. 1. Transmission electron micrograph showing the presence of a flagella-like structure (arrow) at the polar region and slime layer surrounding the cell body of *B. bovis* BbUM strain.

The taxonomic positions of *B. washoensis*, *B. doshiae*, *B. birtlesii*, *B. taylorii*, *B. vinsonii* and *B. alsatica* were inconsistent by SNP and orthologous gene analyses. Three rodent-associated *Bartonella* spp. i.e., *B. birtlesi*, *B. taylorii*, *B. alsatica* and a dog-derived *Bartonella* spp. (*B. vinsonii*) previously grouped in cluster 3 in the SNP-based phylogeny tree formed a new cluster (IV) in the orthologous gene analysis (Fig. 2 i and ii).

Figure 3 shows the distribution of BbUM genes in different functional categories assigned by the RAST pipeline. These genes include those involved in the basic biological functions such as protein metabolism (199 genes), amino acid and derivatives (128 genes), cofactors, vitamins, prosthetic group pigments (99 genes), RNA metabolism (97 genes), DNA metabolism (71 genes), carbohydrates (58 genes), and fatty acids, lipids and isoprenoids (50 genes) and motility and chemotaxis (50 genes). Based on RAST analysis, 10 genes associated with phages (encoding phage major tail tube, sheath proteins, and the large subunit of phage terminase in the subsystem of phage packaging machinery) were annotated in the BbUM genome (Table 2).

The RAST prediction of genes associated with virulence, disease and defence, phages, prophages, transposable elements, plasmids; membrane transport, and motility and chemotaxis is shown in Table 2. The genes associated with the virulence, defence and intracellular resistance of BbUM strains including colicin V and bacteriocin production cluster, lysozyme inhibitors, cobalt-zinc-cadmium resistance, resistance to fluoroquinolones, copper homeostasis: copper tolerance, and beta-lactamase, and *Mycobacterium* virulence operons involved in protein synthesis (SSU and LSU ribosomal proteins) and DNA transcription are annotated in the BbUM genome (Table 2). In the membrane transport category, type IV secretion systems (T4SSs), VirB/VirD4, Vbh/TraG and Trw, and T4SS-translocated effector proteins were not annotated in the BbUM strain. Additionally, the adhesins in bartonellae, the orthologous trimeric autotransporter adhesins, BadA in *B. henselae* and the Vomp family in *B. quintana* [56] are also not annotated.

Several genes encoding protein translocation across the cytoplasmic membrane, ABC transporters, cation transporters and Uni-Sym- and Antiporters are annotated in BbUM genome (Table 2). Various components of iron acquisition metabolism including iron compound ABC uptake transporter substrate-binding proteins PiuA and PiuC, hemin transport protein HmuS, periplasmic hemin-binding protein, ABC-type hemin transport system, ATPase component, electron transfer flavoprotein, beta subunit, ferric siderophore transport system, periplasmic binding protein TonB, TonB-dependent hemin, and ferrichrome receptor are also annotated in the BbUM genome.

Since flagella-like structure was observed under electron microscopy (Fig. 1), the presence of flagellin gene in BbUM genome was investigated. Based on RAST analysis, 50 genes associated with flagellar motility were annotated under the motility and chemotaxis category (Table 2). The genes included *flaA*, *flgB*, *flgC*, *flgD*, *flgE*, *flgF*, *flgG*, *flgH*, *flgI*, *flgK*, *flgL*, *flhA*, *flhB*, *fliE*, *fliF*, *fliG*, *fliI*, *fliL*, *fliM*, *fliP*, *fliQ*, *fliR*, *motA* and *motB*. Additionally, the genes for flagellar protein FlgJ and flagellar hook-associated switch protein (FliN) were also annotated. Two flagellin genes (labelled as *flaA* sequence types 1 and 2) sharing the highest sequence similarities (60.6 and 73.0% nucleotide similarity, respectively) with that of *B. bacilliformis,* and 59.8 and 73.6% nucleotide similarity, respectively, with that of *B. clarridgeiae,* were retrieved for further analysis. The *flaA* type 1 gene of BbUM strain was also detected in *B. bovis* 91-4 and m02 strains with 99.6 and 94.2% sequence similarities, while the type 2 gene was detected in *B. bovis* 91-4 and m02 strains with 99.2 and 90.7% sequence similarities, with the BbUM strain. The similarity between

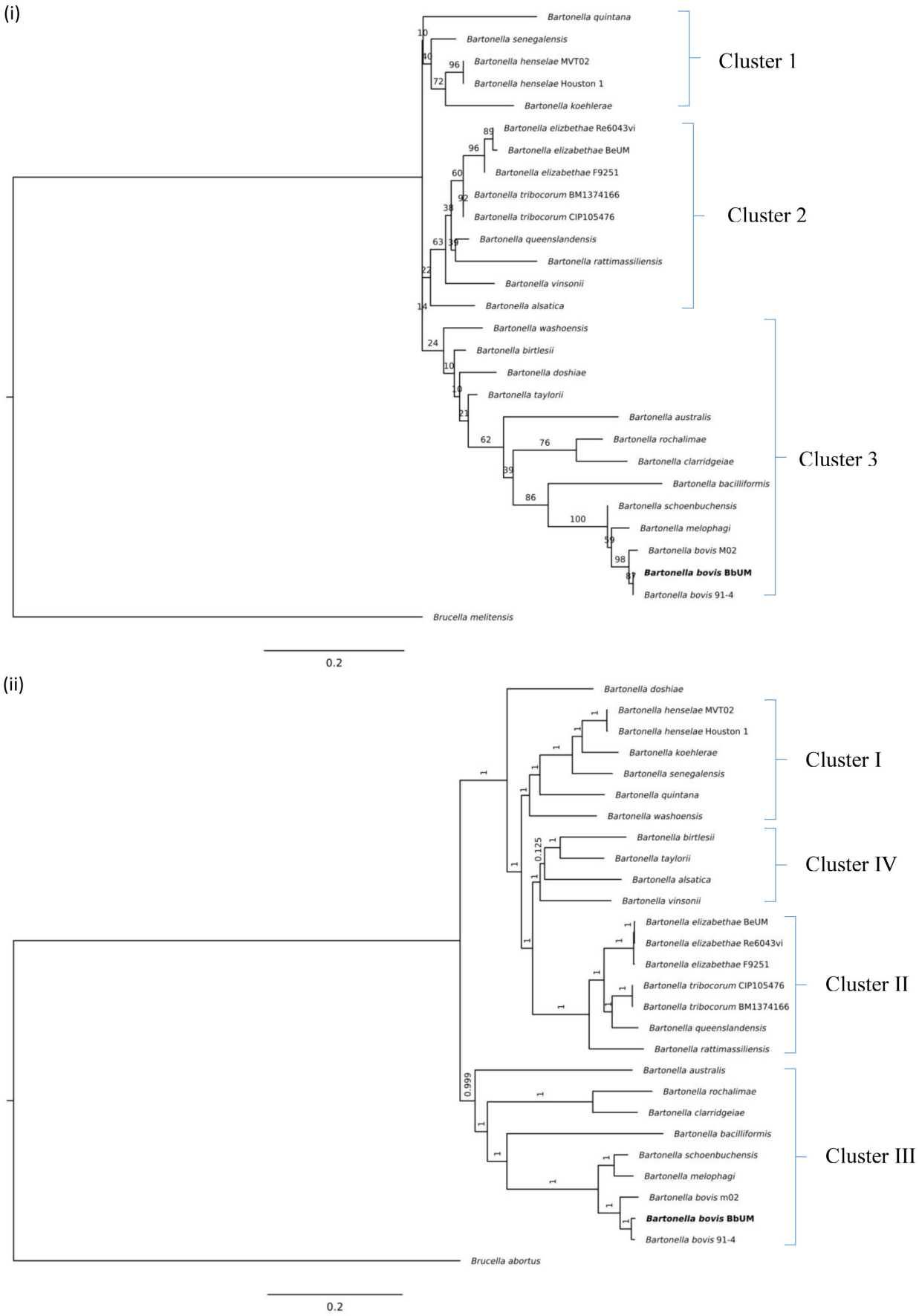

Fig. 2. (i) SNP-based phylogenetic tree with *Brucella abortus* 2308 as an outgroup. (ii) Construction of dendrogram based on protein evolution of *Bartonella* species. The analysis indicates that BbUM strain is originated from the same cluster with *B. bovis* type strain 91-4.

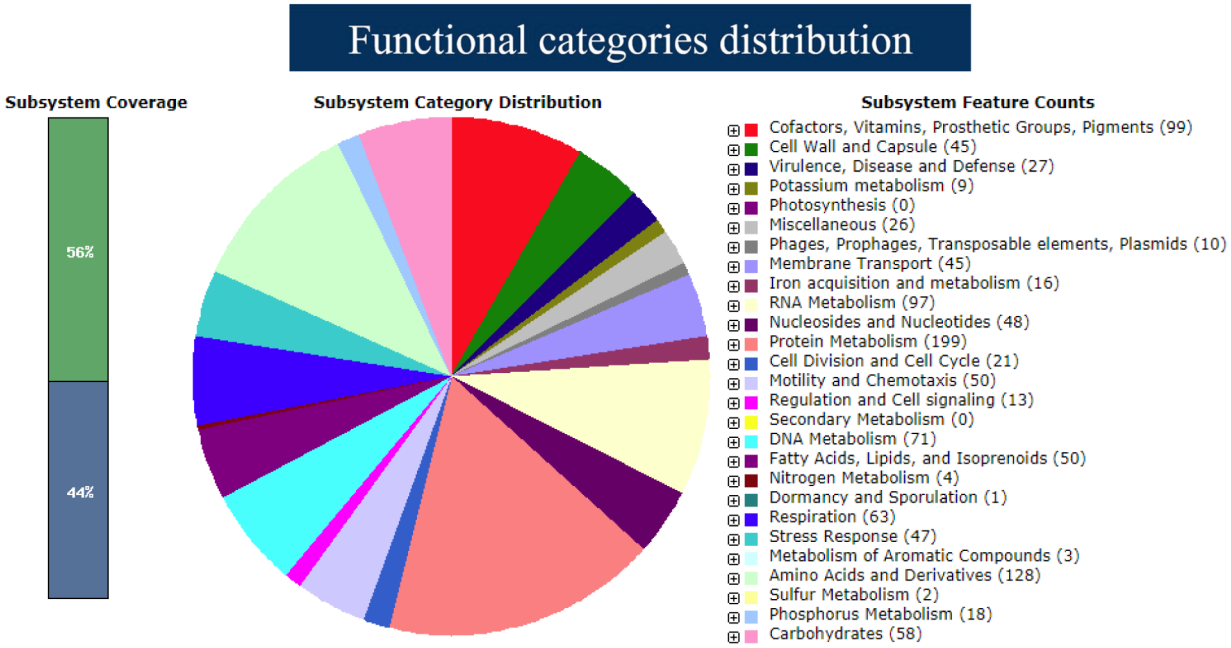

Fig. 3. Subsystem statistics (based on functional categories) of the RAST-predicted genes of BbUM strain.

Table 2. RAST prediction of genes in BbUM, classified under various functional categories

Virulence, disease and defense (n=27)	Phages, prophages, transposable elements, plasmids (n=10)	Membrane transport (n=45)	Motility and chemotaxis (n=50)
Bacteriocins, ribosomally synthesized antibacterial peptides (6) -Colicin V and bacteriocin adhesion (0) -Toxins and superantigens (0) -Production cluster (6) Resistance to antibiotics and toxic compounds (12) -Lysozyme inhibitors (1) -Cobalt-zinc-cadmium resistance (4) -Resistance to fluoroquinolones (4) -Copper homeostasis: copper tolerance (2) -Beta-lactamase (1) Invasion and intracellular resistance (9) -Mycobacterium virulence operon involved in protein synthesis (SSU ribosomal proteins) (4) -Mycobacterium virulence operon involved in DNA transcription (2) -Mycobacterium virulence operon involved in protein synthesis (LSU ribosomal proteins) (3)	Phages, prophages (10) Pathogenicity islands (0) Gene transfer agent (GTA) (0) Plasmid related functions (0) Phage family-specific subsystems (0) Transposable elements (0)	Protein translocation across cytoplasmic membrane (3) -ABC transporters (14) -Cation transporters (2) -Uni-sym and antiporters (8) Membrane transport-no subcategory (18) Protein secretion system, Type I, II, III, IV, V, VI (0) Protein secretion system, Type II (0) Protein secretion Protein secretion system, Type VII (chaperone/usher pathway, CU) (0) Protein secretion system, Type VIII (Extracellular nucleation/ precipitation pathway, ENP) (0) TRAP transporters (0) Sugar phosphotransferase systems, PTS (0)	Flagellar motility in Prokaryota (50) Magnetotaxis (0) Motility and chemotaxis-no subcategory (0) Social motility and nonflagellar swimming in bacteria (0)

The numbers in parenthesis=No. of genes.

the *flaA* type 1 and type 2 sequences for BbUM strain was 55.5%. Figure 4 shows the comparison of *flaA* sequence types 1 and 2 with those of *B. clarridgeiae* and *B. bacilliformis*, and several other published bacterial *flaA* genes. All *Bartonella flaA* genes formed a cluster (with 100% bootstrap value) separated from the plant-associated soil bacteria (*Agrobacteriun tumefaciens* and *Sinorhizobium meliloti*), *Campylobacter jejuni*, *Listeria monocytogenes*, *Vibrio cholerae*, *Aeromonas hydrophila* and *A. punctata* (Fig. 4).

DISCUSSION

The high values obtained from the interstrain comparison using ANI and TETRA analyses in this study (ANI=98.12%, TETRA=0.99438 with reference to *B. bovis* 91-4 type strain) confirms the high degree of genetic relatedness amongst *B. bovis*

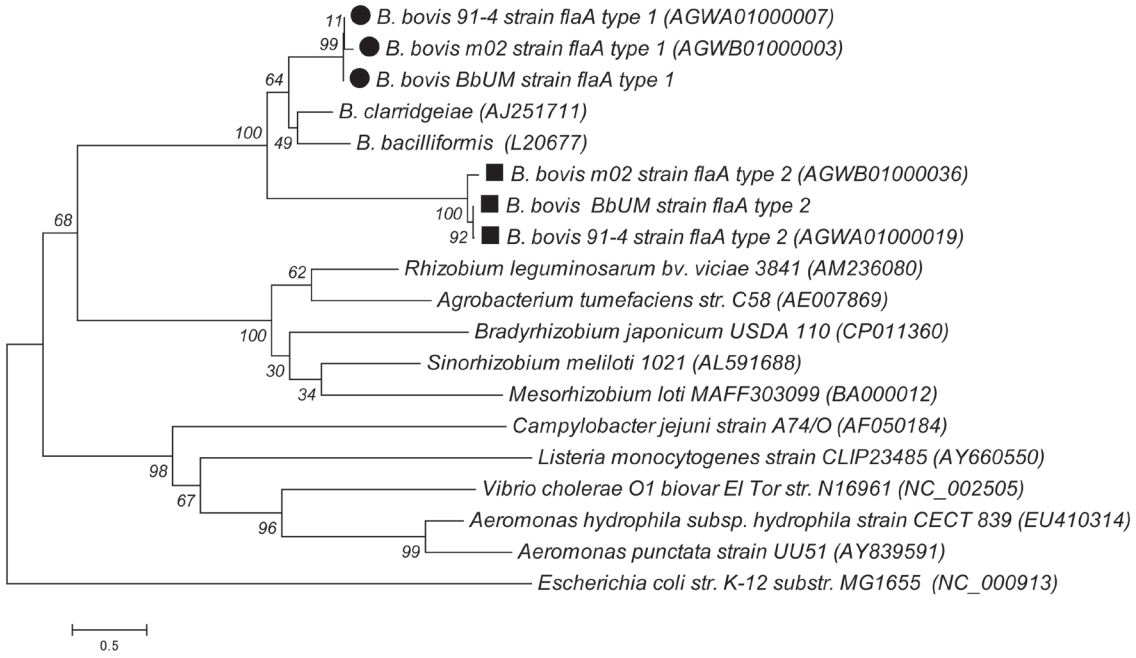

Fig. 4. Phylogenetic tree constructed based on *flaA* genes of BbUM strain, and others available in the GenBank database.

strains. Taken together, the phylogenetic tree generated from the genome-wide comparison of orthologous genes from 27 *Bartonella* species exhibited a topology almost similar to that of the phylogenetic tree generated from SNP-based comparison, indicating a high concordance in the nucleotide and amino acid sequences of *Bartonella* spp. Two cat-associated *Bartonella* spp. (*B. henselae* and *B. koehlerae*) and a tick-derived *Bartonella* sp. (*B. senegalensis*) are grouped in the same phylogenetic group (Cluster I) of the SNP and orthologous gene-based phylogenetic trees (Table 1). All members in Cluster 2 (*B. elizabethae*, *B. queenslandensis*, *B. rattimassiliensis* and *B. tribocorum*) are associated with the isolation from wild rodents. *B. bovis* is grouped in cluster 3 in the SNP phylogenetic tree (Fig. 2i), together with three ruminant-associated *Bartonella* spp., i.e., *B. australis* (isolated from kangaroo), *B. melophagi* (isolated from sheep ked) and *B. schoenbuchensis* (isolated from moose, *Alces alces*), and three human pathogens, *B. rochalimae*, *B. clarridgeiae* and *B. bacilliformis*. The clustering of some cat- or rodent-associated *Bartonella* spp. has led to the hypothesis that these bacteria and their reservoir hosts might have co-evolved [25]. The observation of the clustering of cat- (cluster 1/I), rodent- (cluster 2/II), and ruminant-associated *Bartonella* spp. (cluster 3/III) in the phylogenetic trees constructed from both SNP and orthologous genes in this study further strengthens the hypothesis. This study also confirms the finding of a previous study [23] that B. bacilliformis and *B. australis* are related to ruminant-associated *Bartonella* species.

The ability of *Bartonella* spp. to adapt to specific vectors and reservoirs has been recognized as a common strategy of bartonellae transmission [15]. Variable gene pool in bartonellae, as observed in *B. henselae*, has been reported to play a key role in the establishment of persistent infection in the natural host [36]. The host adaptation of Bartonella spp. has also been linked to the run-off replication with gene transfer agents [7].

Similar to *B. bacilliformis,* T4SSs which are the *bona fide* pathogenicity factors of bartonellae in mediating interbacterial DNA-transfer, and secretion of virulence factors into eukaryotic target cells [45, 48, 49, 53] are not annotated in BbUM genome (Table 2). The absence of *virB*, *vbh* or *trw* gene clusters has also been reported in *B. bovis* strain m02 [23]. Based on genomic analysis, it is speculated that *B. bovis* may have a different mechanism of pathogenicity which merits further investigation.

Flagellum-mediated motility has been postulated as a major virulence factor of *B. bacilliformis* for host invasion [6, 47]. Benson *et al.* [6] observed that *B. bacilliformis* was able to deform erythrocyte membrane in an irreversible fashion while nonmotile variants bound poorly and did not invade erythrocytes. The flagellin of *B. clarridgeiae* has been characterized and is closely related to the flagellin of *B. bacilliformis* [46]. In this study, two sequence types (types 1 and 2) of flagellin genes of BbUM, 91-4 and m02 strains were described (Fig. 4). It is yet to be investigated whether *flaA* gene is a major component of flagellar filament of *B. bovis,* as has been reported to be essential for motility of *Campylobacter jejuni* [40] and *H. pylori* [55].

In summary, comparative genome-wide SNP and orthologous gene analyses confirm the taxonomic placement of *B. bovis* with other ruminant-associated *Bartonella* spp. i.e., *B. australis*, *B. melophagi* and *B. schoenbuchensis*. The genome-based phylogeny study also strengthens the hypothesis on co-evolution of *Bartonella* species in specific animal hosts. The analysis of flagellin and other putative virulence genes provides the basis for further study of the pathogenicity of *B. bovis* as an etiological agent of bacteremia and endocarditis in cattle. With the availability of whole genome sequences, functional characterisation of certain genes is now possible to enhance research in the diagnosis and vaccine development for *B. bovis*.

ACKNOWLEDGMENTS. This project was funded by the University of Malaya Research Grant (UMRG no. RP013-2012A), High Impact Research Grant (UM.C/625/1/HIR-MOHE/CHAN/11) and the Ministry of Science, Technology and Innovation E-Science Fund (SF014-2015).

REFERENCES

1. Akaike, H. 1974. A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* **19**: 716–723. [CrossRef]
2. Alsmark, C. M., Frank, A. C., Karlberg, E. O., Legault, B. A., Ardell, D. H., Canbäck, B., Eriksson, A. S., Näslund, A. K., Handley, S. A., Huvet, M., La Scola, B., Holmberg, M. and Andersson, S. G. 2004. The louse-borne human pathogen *Bartonella quintana* is a genomic derivative of the zoonotic agent *Bartonella henselae*. *Proc. Natl. Acad. Sci. U.S.A.* **101**: 9716–9721. [Medline] [CrossRef]
3. Anisimova, M. and Gascuel, O. 2006. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Syst. Biol.* **55**: 539–552. [Medline] [CrossRef]
4. Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G. D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A. and Zagnitko, O. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **9**: 75. [Medline] [CrossRef]
5. Bai, Y., Malania, L., Alvarez Castillo, D., Moran, D., Boonmar, S., Chanlun, A., Suksawat, F., Maruyama, S., Knobel, D. and Kosoy, M. 2013. Global distribution of *Bartonella* infections in domestic bovine and characterization of *Bartonella bovis* strains using multi-locus sequence typing. *PLoS ONE* **8**: e80894. [Medline] [CrossRef]
6. Benson, L. A., Kar, S., McLaughlin, G. and Ihler, G. M. 1986. Entry of *Bartonella bacilliformis* into erythrocytes. *Infect. Immun.* **54**: 347–353. [Medline]
7. Berglund, E. C., Frank, A. C., Calteau, A., Vinnere Pettersson, O., Granberg, F., Eriksson, A. S., Näslund, K., Holmberg, M., Lindroos, H. and Andersson, S. G. 2009. Run-off replication of host-adaptability genes is associated with gene transfer agents in the genome of mouse-infecting *Bartonella grahamii*. *PLoS Genet.* **5**: e1000546. [Medline] [CrossRef]
8. Bolger, A. M., Lohse, M. and Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120. [Medline] [CrossRef]
9. Breitschwerdt, E. B., Maggi, R. G., Chomel, B. B. and Lappin, M. R. 2010. Bartonellosis: an emerging infectious disease of zoonotic importance to animals and human beings. *J. Vet. Emerg. Crit. Care (San Antonio)* **20**: 8–30. [Medline] [CrossRef]
10. Breitschwerdt, E. B., Sontakke, S., Cannedy, A., Hancock, S. I. and Bradley, J. M. 2001. Infection with *Bartonella weissii* and detection of *Nanobacterium* antigens in a North Carolina beef herd. *J. Clin. Microbiol.* **39**: 879–882. [Medline] [CrossRef]
11. Breitschwerdt, E. B., Linder, K. L., Day, M. J., Maggi, R. G., Chomel, B. B. and Kempf, V. A. 2013. Koch's postulates and the pathogenesis of comparative infectious disease causation associated with *Bartonella* species. *J. Comp. Pathol.* **148**: 115–125. [Medline] [CrossRef]
12. Capella-Gutiérrez, S., Silla-Martínez, J. M. and Gabaldón, T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**: 1972–1973. [Medline] [CrossRef]
13. Chang, C. C., Chomel, B. B., Kasten, R. W., Heller, R. M., Kocan, K. M., Ueno, H., Yamamoto, K., Bleich, V. C., Pierce, B. M., Gonzales, B. J., Swift, P. K., Boyce, W. M., Jang, S. S., Boulouis, H. J. and Piémont, Y. 2000. *Bartonella* spp. isolated from wild and domestic ruminants in North America. *Emerg. Infect. Dis.* **6**: 306–311. [Medline] [CrossRef]
14. Chen, F., Mackey, A. J., Stoeckert, C. J. Jr. and Roos, D. S. 2006. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Res.* **34**: D363–D368. [Medline] [CrossRef]
15. Chomel, B. B., Boulouis, H. J., Breitschwerdt, E. B., Kasten, R. W., Vayssier-Taussat, M., Birtles, R. J., Koehler, J. E. and Dehio, C. 2009. Ecological fitness and strategies of adaptation of *Bartonella* species to their hosts and vectors. *Vet. Res.* **40**: 29. [Medline] [CrossRef]
16. Darriba, D., Taboada, G. L., Doallo, R. and Posada, D. 2011. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* **27**: 1164–1165. [Medline] [CrossRef]
17. Darriba, D., Taboada, G. L., Doallo, R. and Posada, D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* **9**: 772. [Medline] [CrossRef]
18. Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**: 1792–1797. [Medline] [CrossRef]
19. Gardner, S. N., Slezak, T. and Hall, B. G. 2015. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics* **31**: 2877–2878. [Medline] [CrossRef]
20. Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P. and Tiedje, J. M. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* **57**: 81–91. [Medline] [CrossRef]
21. Guindon, S., Delsuc, F., Dufayard, J. F. and Gascuel, O. 2009. Estimating maximum likelihood phylogenies with PhyML. *Methods Mol. Biol.* **537**: 113–137. [Medline] [CrossRef]
22. Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W. and Gascuel, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**: 307–321. [Medline] [CrossRef]
23. Guy, L., Nystedt, B., Toft, C., Zaremba-Niedzwiedzka, K., Berglund, E. C., Granberg, F., Näslund, K., Eriksson, A. S. and Andersson, S. G. 2013. A gene transfer agent and a dynamic repertoire of secretion systems hold the keys to the explosive radiation of the emerging pathogen *Bartonella*. *PLoS Genet.* **9**: e1003393. [Medline] [CrossRef]
24. Han, N., Qiang, Y. and Zhang, W. 2016. ANItools web: a web tool for fast genome comparison within multiple bacterial strains. *Database (Oxford)* **2016**: 1. [Medline] [CrossRef]
25. Houpikian, P. and Raoult, D. 2001. Molecular phylogeny of the genus *Bartonella*: what is the current knowledge? *FEMS Microbiol. Lett.* **200**: 1–7. [Medline] [CrossRef]
26. Hyatt, D., Chen, G. L., Locascio, P. F., Land, M. L., Larimer, F. W. and Hauser, L. J. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**: 119. [Medline] [CrossRef]
27. Kho, K. L., Koh, F. X., Jaafar, T., Nizam, Q. N. H. and Tay, S. T. 2015. Prevalence and molecular heterogeneity of *Bartonella bovis* in cattle and *Haemaphysalis bispinosa* ticks in Peninsular Malaysia. *BMC Vet. Res.* **11**: 153. [Medline] [CrossRef]
28. Kim, M., Oh, H. S., Park, S. C. and Chun, J. 2014. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int. J. Syst. Evol. Microbiol.* **64**: 346–351. [Medline] [CrossRef]

29. Konstantinidis, K. T. and Tiedje, J. M. 2005. Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci. U.S.A.* **102**: 2567–2572. [Medline] [CrossRef]
30. Konstantinidis, K. T., Ramette, A. and Tiedje, J. M. 2006. The bacterial species definition in the genomic era. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **361**: 1929–1940. [Medline] [CrossRef]
31. Kosoy, M., Morway, C., Sheff, K. W., Bai, Y., Colborn, J., Chalcraft, L., Dowell, S. F., Peruski, L. F., Maloney, S. A., Baggett, H., Sutthirattana, S., Sidhirat, A., Maruyama, S., Kabeya, H., Chomel, B. B., Kasten, R., Popov, V., Robinson, J., Kruglov, A. and Petersen, L. R. 2008. *Bartonella tamiae* sp. nov., a newly recognized pathogen isolated from three human patients from Thailand. *J. Clin. Microbiol.* **46**: 772–775. [Medline] [CrossRef]
32. Kosoy, M., Bai, Y., Sheff, K., Morway, C., Baggett, H., Maloney, S. A., Boonmar, S., Bhengsri, S., Dowell, S. F., Sitdhirasdr, A., Lerdthusnee, K., Richardson, J. and Peruski, L. F. 2010. Identification of *Bartonella* infections in febrile human patients from Thailand and their potential animal reservoirs. *Am. J. Trop. Med. Hyg.* **82**: 1140–1145. [Medline] [CrossRef]
33. Lagesen, K., Hallin, P., Rødland, E. A., Staerfeldt, H. H., Rognes, T. and Ussery, D. W. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **35**: 3100–3108. [Medline] [CrossRef]
34. Laslett, D. and Canback, B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* **32**: 11–16. [Medline] [CrossRef]
35. Li, L., Stoeckert, C. J. Jr. and Roos, D. S. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13**: 2178–2189. [Medline] [CrossRef]
36. Lindroos, H., Vinnere, O., Mira, A., Repsilber, D., Näslund, K. and Andersson, S. G. 2006. Genome rearrangements, deletions, and amplifications in the natural population of *Bartonella henselae*. *J. Bacteriol.* **188**: 7426–7439. [Medline] [CrossRef]
37. Maillard, R., Grimard, B., Chastant-Maillard, S., Chomel, B., Delcroix, T., Gandoin, C., Bouillin, C., Halos, L., Vayssier-Taussat, M. and Boulouis, H. J. 2006. Effects of cow age and pregnancy on *Bartonella* infection in a herd of dairy cattle. *J. Clin. Microbiol.* **44**: 42–46. [Medline] [CrossRef]
38. Marçais, G. and Kingsford, C. 2011. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **27**: 764–770. [Medline] [CrossRef]
39. Marston, E. L., Sumner, J. W. and Regnery, R. L. 1999. Evaluation of intraspecies genetic variation within the 60 kDa heat-shock protein gene (*groEL*) of *Bartonella* species. *Int. J. Syst. Bacteriol.* **49**: 1015–1023. [Medline] [CrossRef]
40. Neal-McKinney, J. M., Samuelson, D. R., Eucker, T. P., Nissen, M. S., Crespo, R. and Konkel, M. E. 2014. Reducing *Campylobacter jejuni* colonization of poultry via vaccination. *PLOS ONE* **9**: e114254. [Medline] [CrossRef]
41. Peng, Y., Leung, H. C., Yiu, S. M. and Chin, F. Y. 2012. IDBA-UD: a *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* **28**: 1420–1428. [Medline] [CrossRef]
42. Pitulle, C., Strehse, C., Brown, J. W. and Breitschwerdt, E. B. 2002. Investigation of the phylogenetic relationships within the genus *Bartonella* based on comparative sequence analysis of the *rnpB* gene, 16S rDNA and 23S rDNA. *Int. J. Syst. Evol. Microbiol.* **52**: 2075–2080. [Medline]
43. Rudoler, N., Rasis, M., Sharir, B., Novikov, A., Shapira, G. and Giladi, M. 2014. First description of *Bartonella bovis* in cattle herds in Israel. *Vet. Microbiol.* **173**: 110–117. [Medline] [CrossRef]
44. Richter, M. and Rosselló-Móra, R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. U.S.A.* **106**: 19126–19131. [Medline] [CrossRef]
45. Saenz, H. L., Engel, P., Stoeckli, M. C., Lanz, C., Raddatz, G., Vayssier-Taussat, M., Birtles, R., Schuster, S. C. and Dehio, C. 2007. Genomic analysis of *Bartonella* identifies type IV secretion systems as host adaptability factors. *Nat. Genet.* **39**: 1469–1476. [Medline] [CrossRef]
46. Sander, A., Zagrosek, A., Bredt, W., Schiltz, E., Piémont, Y., Lanz, C. and Dehio, C. 2000. Characterization of *Bartonella clarridgeiae* flagellin (FlaA) and detection of antiflagellin antibodies in patients with lymphadenopathy. *J. Clin. Microbiol.* **38**: 2943–2948. [Medline]
47. Scherer, D. C., DeBuron-Connors, I. and Minnick, M. F. 1993. Characterization of *Bartonella bacilliformis* flagella and effect of antiflagellin antibodies on invasion of human erythrocytes. *Infect. Immun.* **61**: 4962–4971. [Medline]
48. Schröder, G. and Dehio, C. 2005. Virulence-associated type IV secretion systems of *Bartonella*. *Trends Microbiol.* **13**: 336–342. [Medline] [CrossRef]
49. Schulein, R. and Dehio, C. 2002. The VirB/VirD4 type IV secretion system of *Bartonella* is essential for establishing intraerythrocytic infection. *Mol. Microbiol.* **46**: 1053–1067. [Medline] [CrossRef]
50. Simpson, J. T. and Durbin, R. 2012. Efficient *de novo* assembly of large genomes using compressed data structures. *Genome Res.* **22**: 549–556. [Medline] [CrossRef]
51. Takahashi, M., Kryukov, K. and Saitou, N. 2009. Estimation of bacterial species phylogeny through oligonucleotide frequency distances. *Genomics* **93**: 525–533. [Medline] [CrossRef]
52. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**: 2731–2739. [Medline] [CrossRef]
53. Tay, S. T., Kho, K. L., Wee, W. Y. and Choo, S. W. 2016. Whole-genome sequence analysis and exploration of the zoonotic potential of a rat-borne *Bartonella elizabethae*. *Acta Trop.* **155**: 25–33. [Medline] [CrossRef]
54. Teeling, H., Waldmann, J., Lombardot, T., Bauer, M. and Glöckner, F. O. 2004. TETRA: a web-service and a stand-alone program for the analysis and comparison of tetranucleotide usage patterns in DNA sequences. *BMC Bioinformatics* **5**: 163. [Medline] [CrossRef]
55. Watanabe, S., Takagi, A., Tada, U., Kabir, A. M., Koga, Y., Kamiya, S., Osaki, T. and Miwa, T. 1997. Cytotoxicity and motility of *Helicobacter pylori*. *J. Clin. Gastroenterol.* **25** Suppl 1: S169–S171. [Medline] [CrossRef]
56. Zhang, P., Chomel, B. B., Schau, M. K., Goo, J. S., Droz, S., Kelminson, K. L., George, S. S., Lerche, N. W. and Koehler, J. E. 2004. A family of variably expressed outer-membrane proteins (Vomp) mediates adhesion and autoaggregation in *Bartonella quintana*. *Proc. Natl. Acad. Sci. U.S.A.* **101**: 13630–13635. [Medline] [CrossRef]