RESEARCH LETTER – Pathogens & Pathogenicity

Typing and comparative genome analysis of Brucella melitensis isolated from Lebanon

Natalia Abou Zaki1, Tamara Salloum1, Marwan Osman2, Rayane Rafei2, Monzer Hamze2 and Sima Tokajian1,∗,†

1Department of Natural Sciences, School of Arts and Sciences, Lebanese American University, Byblos 36, Lebanon and 2Health and Environmental Microbiology Laboratory, Doctoral School for Sciences and Technology and Faculty of Public Health-Lebanese University, Tripoli 1300, Lebanon

∗Corresponding author: Department of Natural Sciences, School of Arts and Sciences, Lebanese American University, Byblos Campus, P.O. Box 36, Lebanon. Tel: +961-9-547262; Fax: +961-9-546262; E-mail: stokajian@lau.edu.lb
†Sima Tokajian, http://orcid.org/0000-0002-3653-8940

ABSTRACT

Brucella melitensis is the main causative agent of the zoonotic disease brucellosis. This study aimed at typing and characterizing genetic variation in 33 Brucella isolates recovered from patients in Lebanon. Bruce-ladder multiplex PCR and PCR-RFLP of omp31, omp2a and omp2b were performed. Sixteen representative isolates were chosen for draft-genome sequencing and analyzed to determine variations in virulence, resistance, genomic islands, prophages and insertion sequences. Comparative whole-genome single nucleotide polymorphism analysis was also performed. The isolates were confirmed to be B. melitensis. Genome analysis revealed multiple virulence determinants and efflux pumps. Genome comparisons and single nucleotide polymorphisms divided the isolates based on geographical distribution but revealed high levels of similarity between the strains. Sequence divergence in B. melitensis was mainly due to lateral gene transfer of mobile elements. This is the first report of an in-depth genomic characterization of B. melitensis in Lebanon.

Keywords: Brucella melitensis; Brucellosis; PCR; RFLP; draft-genome sequencing; SNPs

INTRODUCTION

Brucella spp. are Gram-negative, facultative intracellular, α-proteobacteria that lack classical virulence factors (VFs) possessed by other pathogens such as endospores, native plasmids and capsules (Corbel 1997; Moreno and Moriyón 2002). Apparently descending from a free-living organism, Brucella spp. underwent evolutionary events including the loss and acquisition of genes to subsequently become an animal parasite (Ficht 2010). Brucellosis is transmitted to humans through the consumption of infected meat or dairy products, direct contact or inhalation of droplets (Mantur, Amarnath and Shinde 2007). Only 100 to 1000 organisms are sufficient to cause an infection. Macrophages then transport Brucella to the lymph nodes, spleen, liver, bone marrow, mammary glands and sex organs. Common symptoms of brucellosis include fever, nausea, myalgia, arthralgia of large joints, headache, hot flushes, diarrhea and vomiting. Meningitis, endocarditis, spondylitis and arthritis are also reported (Ko and Splitter 2003). The genus Brucella has 12 recognized species, all of which exhibit distinct host preferences (Moreno 2014) (http://www.bacterio.net/brucella.html). Brucella pathogenesis lies in its ability to survive intracellularly and multiply within host cells such as macrophages.
through utilization of factors that enhance its fitness such as lipopolysaccharides (LPS), type IV secretion system (T4SS) and BvrR/S two-component system (TCS) (Martínez-Núñez et al. 2010).

The Middle East is by far the most endemic region for brucellosis worldwide (Pappas and Memish 2007). In Lebanon and the Middle East, brucellosis is mainly caused by Brucella melitensis with reported cases of brucellosis following contamination of dairy products, in addition to occupational risk (Dajani, Maysoud and Barakat 1989; Araj and Azzam 1996; Pappas and Memish 2007; Alwan et al. 2010). Species identification after isolation from blood, tissue or body fluids is performed either through biochemical tests, serological tests or PCR-based assays (Sakran, Chazan and Koren 2006). Serological assays are the most commonly used in the diagnosis of brucellosis (Araj 2010). However, in the absence of a standard reference antigen, the source of the antigens and the methods followed could influence the obtained results (Araj 2010).

Several molecular typing methods have been applied to search for DNA polymorphisms between Brucella spp. and biovars, such as PCR-RFLP (Wattiau et al. 2011; Mirnejad et al. 2013). Moreover, the omp2 locus composed of two gene copies (omp2a and omp2b) coding for porin proteins has been found particularly useful for Brucella identification at the species, biovar and strain levels (Zerva et al. 2001). However, one of the challenges of using molecular typing techniques for the differentiation of various Brucella spp. and strains is their high level of genetic homology (~90%) (Verger et al. 1985). Targeting loci containing variable number of tandem repeats (VNTRs) and the corresponding MLVA (multiple locus VNTR analysis) proved useful in discriminating human Brucella isolates in correlation with their geographic origins (Al Dahouk et al. 2007). MLVA-16 distinguished 56 genotypes among B. melitensis isolates in Italy mostly belonging to the West Mediterranean lineage (Garofolo et al. 2013). Similarly, MLVA-15 largely separated genotypes of B. melitensis from Egypt, Qatar and Libya (Tiller et al. 2009). In another study, MLVA was successfully used in the genetic analysis of camel B. melitensis infections in the United Arab Emirates (Gyuranecz et al. 2016). More recently genome-based approaches such as single nucleotide polymorphism (SNP) analysis confirmed whole-genome sequencing (WGS) as a powerful tool for the accurate typing and better resolution in the Brucella genus (Georgi et al. 2017). In closely related species such as in Brucella, SNP-based phylogenetics provide enough discrimination and ease of comparison (Foster et al. 2009).

However, the lack of whole-genome-based molecular analysis is a caveat in the understanding of the patterns of evolution of Brucella and its patterns of transmission associated with travelling and recent population migrations (Dean et al. 2012; Georgi et al. 2017). To our knowledge, this is the first in-depth genetic investigation of a collection of B. melitensis strains obtained from human cases in Lebanon.

In this study, the species of 33 Brucella isolates recovered from blood and articular fluid were determined using the Bruce-ladder multiplex PCR, IS711 gene PCR, omp31, omp2a and omp2b PCR-RFLP and Rev 1 rpsL PCR to differentiate the clinical isolates from Rev 1 vaccine strain. Sixteen selected B. melitensis isolates were then subjected to draft-genome sequencing to characterize their resistance patterns, VFs and repertoire of mobile genetic elements. Evolutionary and geographic relatedness of the isolates compared to publicly available genomes were determined by whole-genome SNPs (wgSNPs), comparative and phylogenetic analysis.

**MATERIALS AND METHODS**

**Ethical approval**

Ethical approval was not required as clinical isolates were collected and stored as part of routine clinical care. Clinical isolates and patient records/information were anonymized and de-identified prior to analysis.

**Clinical isolates**

A total of 33 Brucella isolates were kindly provided by Azm Center for Research and Biotechnology. The isolates are referred to as Bru003–Bru011, Bru013–Bru036 throughout the manuscript. Isolates Bru003–011 and Bru013–036 were recovered from blood; Bru016 was recovered from the articular fluid and Bru027 from the ascitic fluid. For genome sequencing, 16 representative isolates were chosen: Bru003–004, Bru008, Bru010, Bru013–015, Bru018, Bru025–Bru027, Bru029-Bru030 and Bru034–036.

**Antimicrobial testing**

Antimicrobial susceptibility testing was performed using the disk diffusion method for ofloxacin, ciprofloxacin, tetracycline, minocycline, rifampicin, gentamicin and sulfamethoxazole/trimethoprim (Oxoid, England). All antimicrobial testing was performed on Mueller-Hinton agar with 5% (v/v) blood and the Clinical and Laboratory Standards Institute guidelines were used for data interpretation (CLSI 2013).

**DNA extraction**

DNA was extracted using the Nucleospin Tissue kit (Macherey-Nagel, Germany) according to the manufacturer’s instructions.

**Bruce-ladder multiplex PCR**

Bruce-ladder multiplex PCR for species determination of Brucella isolates was performed following Garcia-Yoldi et al. (2006).

**IS711 gene and Rev 1 rpsL PCR**

Genomic DNA was amplified using IS711 primers specific for B. melitensis. In order to differentiate the obtained B. melitensis clinical isolates from Rev 1 vaccine strain, rpsL-PCR was performed using rpsL primers specific for Rev 1 vaccine strain, as previously described by Awwad et al. (2015).

**PCR-RFLP of omp31, omp2a and omp2b**

PCR of omp31 was performed using 31sd (5’-TGACA-GCATTTTTCGGCAA-3’) and 31ter (5’-CATCAGGACAA-TTCCGCC- 3’) primers (Singh et al. 2013). omp2a and omp2b were amplified as previously reported by Cloackaert et al. (2002).

Each PCR product (5 μl) was cleaved in a 20-μl volume reaction with 5 U of restriction enzymes HaeIII, HinfI and PstI for omp31, omp2a and omp2b, respectively (Thermo Fisher Scientific, MA, USA) for 2 h at 37 °C according to the manufacturer’s instructions.

**Genome sequencing**

One nanogram of genomic DNA from each isolate was used as input for library preparation using the Nextera XT library prep kit (Illumina, Inc., San Diego, CA, USA). The subsequent
clean-up steps were performed using the AMPure XP PCR purification beads (Agencourt, Brea, CA, USA) following the manufacturer's instructions. The resulting libraries were quantified by quantitative PCR in triplicate at 1:1000 and 1:2000 and using the Kapa library quantification kit (Kapa Biosystems, Woburn, MA, USA) and as recommended by the manufacturer’s protocol. The samples were pooled together and then sequenced on an Illumina MiSeq for paired-end 250-bp reads.

**Genome assembly and annotation**

De novo genome assembly was performed using the A5-miseq assembly pipeline keeping default parameters (Tritt et al. 2012). Data cleaning, error correction, contig assembly, scaffolding and quality control processes are all automated by this pipeline. The resulting de novo assemblies were annotated using the RAST server [http://rast.nmpdr.org](http://rast.nmpdr.org) (Aziz et al. 2012).

KmerFinder tool was used for species prediction based on the number of co-occurring k-mers in DNA (www.genomicepidemiology.org). Core components of the T3SS and T4SS were searched using the T346Hunter application (Martinez-Garcia, Ramos and Rodriguez-Palenzuela 2015). Phage Search Tool (PHAST) [http://phast.wishartlab.com/index.html](http://phast.wishartlab.com/index.html) was used to detect phages, sites of integration and closest identity match (Zhou et al. 2011). Insertion sequences (IIS) were detected using Biotoul IS-Finder [http://www.is-biotoul.fr/](http://www.is-biotoul.fr/) (Siguer et al. 2006). IslandViewer [http://pathogenomics.sfu.ca/islandviewer](http://pathogenomics.sfu.ca/islandviewer) was used for the identification of pathogenicity islands (Dhillon et al. 2015).

**Comparative genome analysis**

CGview Server V 1.0 was used to visualize the sequence feature information and circular genome representation of sequenced isolates through comparison with B. melitensis 16M reference genome (accession #: NC_003318 and NC_003318) (DelVecchio et al. 2002; Grant and Stothard 2008). DNAPlotter was used for circular genome maps visualization (Carver et al. 2009).

In order to determine the phylogenetic relationship of the sequenced strains, 37 concatenated marker genes were used to construct a maximum-likelihood tree using seven B. melitensis reference genomes downloaded from the NCBI database (16M, ATCC 23457, Ether, NI, M5–90, M28 and strain 63/9, in addition to other Brucella spp. (B. abortus 2308, B. abortus 9–941, B. canis, B. ceti, B. pinnipediais, B. microti, B. suis, B. ovis and B. vulpis). Ochrobactrum anthropi was used as an outgroup to root the tree. The genomes were processed with PhyloSift, and Dendroscope was used to edit and visualize the tree (Huson and Scornavacca 2012; Darling et al. 2014). Phylosift accounts for a set of 37 marker genes having parallel phylogenetic histories (Darling et al. 2014).

All genome sequences were aligned against the reference 16M using Mauve 2.2.0 (Darling et al. 2004).

**SNP-based phylogenetic analysis**

The sequence data were aligned against the 16M reference sequence using the Burrows–Wheeler Aligner (Li and Durbin 2010). SNP calling was done by mapping the paired-end reads of the 16 genomes and 7 B. melitensis reference genomes downloaded from NCBI to the genome of 16M, using the BioNumerics v7.6.1 beta software (Applied Maths, Sint-Martens-Latem, Belgium). A neighbor-joining tree was drawn in BioNumerics by using the wgSNP data as input. The tree was rooted by using the O. anthropi as an outgroup.

**Nucleotide sequence accession numbers**

The whole-genome shotgun sequences of Brucella spp. Bru003–004, Bru008, Bru010, Bru013–015, Bru0018, Bru025–027, Bru029–030 and Bru034–036 are available in GenBank respectively under accession numbers: MKP00000000, MKPW00000000, MJL00000000, MKX100000000, MKXJ00000000, MKPX00000000, MKPY00000000, MNBN00000000, MKF00000000, MNK00000000, MRUO00000000, MRUP00000000, MRUQ00000000 and MKQB00000000.

**RESULTS**

**Bruce-ladder multiplex PCR**

In all the tested isolates, six fragments of 152, 450, 590, 794, 1071 and 1682 bp were amplified. The fragment sizes obtained were in accordance with the unique pattern of Brucella melitensis (Data not shown).

**PCR analysis**

IS711 gene (731 bp) specific for B. melitensis 16M was amplified in all 33 isolates. None of the isolates was positive for Rev 1 vaccine strain rpsL, indicating that these were all naturally occurring isolates.

**PCR-RFLP of omp31, omp2a and omp2b**

*omp31* gene (900 bp) and *omp2a* (1100 bp) genes were amplified in all the tested isolates. Three bands of 210 bp, 280 bp and 320 bp were observed following enzymatic digestion of *omp31* with HaellI. Three bands of 210 bp, 270 bp and 550 bp were observed with HinfI restriction digestion of *omp2a* (P2 and P3 patterns) and three bands of 250 bp, 480 bp and 520 bp were obtained with HinfI restriction digestion of *omp2b* (P1 pattern). All three patterns corresponded to B. melitensis biovar 1 (data not shown).

**Genome statistics**

The average genome size of the 16 sequenced isolates was 3297 621 bp, ranging from 3290 179 bp to 3322 963 bp. The number of contigs ranged from 34 to 80. The G + C content was 57.2 ± 0.1% and the number of detected subsystems was 434 ± 4. The average number of coding sequences was 3327, ranging from 3314 to 3339, with 56 RNAs detected (tRNAs and rRNAs) (Fig. 1; Table S1, Supporting Information). All the 16 sequenced genomes were confirmed to be B. melitensis using the KmerFinder 2.0 tool along with the NCBI Blast (Hasman et al. 2014).

**Abundance of VFs**

Among the detected virulence determinants in the 16 sequenced genomes were those linked to secretion systems (T4SS and flagellar genes), LPS genes linked to O-antigen biosynthesis (genes encoding phosphomannomutase, mannos 6-phosphate isomerase and mannose guanylyltransferase), ureases (*ureA, ureB, ureC, ureD, ureE, ureF* and *ureG*), BvrR/S TCS (encoded by *chol* and *chuG*), regulation of gene expression (*hfq*), cyclic β-1,2-glucan (*cgs*) and stringent response (*spoT*/*rsh*).

For T4SS components, *virB 1–11* were detected in all the 16 sequenced genomes using RAST and T346Hunter except *virB7*. 
Figure 1. Circular maps of Bru004 chromosomes 1 and 2. From outside to inside, the tracks represent forward coding sequence (CDS) (red); reverse CDS (blue); tRNAs (green); G + C plot; G + C skew (black above average and gray below average). Image was generated using DNAPlotter (Carver et al. 2009).

Table 1. Genes found on GIs in the sequenced B. melitensis isolates as detected by Islandviewer 3.

<table>
<thead>
<tr>
<th>Component</th>
<th>Bru003</th>
<th>Bru004</th>
<th>Bru008</th>
<th>Bru010</th>
<th>Bru013</th>
<th>Bru014</th>
<th>Bru015</th>
<th>Bru018</th>
<th>Bru025</th>
<th>Bru026</th>
<th>Bru027</th>
<th>Bru029</th>
<th>Bru030</th>
<th>Bru034</th>
<th>Bru035</th>
<th>Bru036</th>
</tr>
</thead>
<tbody>
<tr>
<td>wbkE</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>per</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Gmd</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>wboA</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>wboB</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>wbkC</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>wbkA</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Outer membrane protein: − = absent, + = present.

Exonuclease: − = absent, + = present.

wbkE = glycosyltransferase, per = perosamine synthetase, gmd = GDP-mannose 4-6 dehydratase, wboA = glycosyltransferase, wboB = glycosyltransferase, wbkC = formyltransferase, wbkA = mannosyltransferase.

The components of T4SS were within one gene cluster. Table 2 shows the 11 core components of T4SS and their corresponding annotations. Three gene clusters linked to flagellar components were also detected using T346Hunter (Table S2, Supporting Information).

Islandviewer 3 was used to detect genomic islands (GIs) by aligning the genomes with reference strain B. melitensis 16M [Illumina + 454] (Accession #: GCA_00740415). LPS-related genes, exonucleases and outer membrane proteins were found in several GIs within the sequenced isolates. Additionally, genes linked to the virB operon (virB2, virB3, virB4, virB5, virB6 and virB8) encoding for T4SS components were detected in Bru015 (Table 1).

Antibiotics resistance profiles
The antimicrobial susceptibility against ofloxacin, minocycline, rifampicin, tetracycline, gentamicin, ciprofloxacin and sulfamethoxazole/trimethoprim was determined. All isolates were susceptible to tetracyclines (minocycline and tetracycline), aminoglycosides (gentamicin) and rifamycins (rifampicin). Resistance to ofloxacin (fluoroquinolone) was detected in Bru007. Bru028 was resistant to both ofloxacin and ciprofloxacin. The NorMII efflux pump, which confers resistance to fluoroquinolones, was detected in all the sequenced genomes.

Resistance to trimethoprim and sulfonamide was observed in 87.5% (n = 14) of the tested isolates (Bru003, Bru014–015, Bru017, Bru019–024, Bru028 and Bru030–032), all of which were positive for the RND efflux pumps, conferring resistance to trimethoprim.

Phages and mobile genetic elements
At least one prophage was detected per genome with sizes ranging between 6 and 69 kb providing some diversity between the
Table 2. Core components of the T4SS.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Corresponding protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>virB1</td>
<td>Peptidoglycan hydrolase involved in T-DNA transfer</td>
</tr>
<tr>
<td>virB2</td>
<td>Major pilus subunit of type IV secretion complex</td>
</tr>
<tr>
<td>virB3</td>
<td>Inner membrane protein forms channel for type IV secretion of T-DNA complex</td>
</tr>
<tr>
<td>virB4</td>
<td>ATPase required for both assembly of type IV secretion complex and secretion of T-DNA complex</td>
</tr>
<tr>
<td>virB5</td>
<td>Minor pilin of type IV secretion complex</td>
</tr>
<tr>
<td>virB6</td>
<td>Inner membrane protein of type IV secretion of T-DNA complex</td>
</tr>
<tr>
<td>virB8</td>
<td>Inner membrane protein forms channel for type IV secretion of T-DNA complex</td>
</tr>
<tr>
<td>virB9</td>
<td>Outer membrane and periplasm component of type IV secretion of T-DNA complex, has secretin-like domain</td>
</tr>
<tr>
<td>virB10</td>
<td>Inner membrane protein of type IV secretion of T-DNA complex, TonB-like</td>
</tr>
<tr>
<td>virB11</td>
<td>ATPase required for both assembly of type IV secretion complex and secretion of T-DNA complex</td>
</tr>
</tbody>
</table>

Figure 2. Maximum-likelihood concatenated marker tree of the Brucella strains used in this study compared to reference strains downloaded from NCBI database based on the concatenated alignment of 37 core genes as generated by PhyloSift (Darling et al. 2014).

isolates. Brucellaphage BiPBO1 was detected in isolates Bru003, Bru004, Bru013, Bru015, Bru018, Bru025, Bru026, Bru027 and Bru034 with sizes ranging between 15.1 (incomplete) and 24.1 kb (intact). An intact phage (Paracoccus vB Fmas) was found in all isolates except Bru013 and Bru034. A 6-kb Gordonia Phage Splinter was detected in Bru010 and Bru030. Staphylococcus beta-like phage ranging in size between 20.4 and 38.8 kb was found in Bru003, Bru025 and Bru036. However, phages Lactococcus jj50 and Mycobacterium phage Makemake were found only in Bru015 and Bru018, respectively (Table S3, Supporting Information).

Insertion sequences

ISL3, IS3, IS481, IS5 and IS711 were detected using IS Finder in all the sequenced genomes. IS66 was only present in Bru018, Bru029, Bru030 and Bru035 and IS110 in Bru004, Bru008, Bru010, Bru014, Bru029, Bru030 and Bru034 and IS6 in Bru003, Bru014, Bru015, Bru025, Bru027 and Bru036.

Phylogenetic analysis

Phylogenetic analysis using PhyloSift based on the concatenated alignment of 37 core genes separated the sequenced isolates in five clades distinct from other B. melitensis reference strains. Bru003, Bru015 and Bru034 clustered separately (Fig. 2). Also, comparative analysis was performed to detect sequence divergence between reference strains 16M (Biovar 1) and Ether (Biovar 3), ATCC23457 and 63/9 (Biovar 2), and strains M5–90 and M28. All reference strains were highly conserved in their core genomes, where divergence was mainly attributed to hypothetical proteins and mobile elements between strains 16M and Ether. As for strains M5–90 and M28, divergence was due to several virulence determinants including efflux pumps, ABC transporters, cold shock proteins and flagellar genes.

Comparative genome analysis

To study the genetic relatedness of the isolates, wgSNP-based phylogenetic analysis was conducted using seven B. melitensis genomes downloaded from NCBI and the genome of 16M as a reference. SNP analysis yielded 38,431 SNP sites and found only silent and intergenic mutations between the sequenced isolates causing no amino acids substitutions. The SNFs were evenly distributed all over the genomes, and their locations were also conserved between the sequenced isolates. Particularly, SNPs were present in methyltransferases, phage integrases, dehydrogenases, SOS-response repressors, various ATPases and enzymes involved in metabolism, among other genes.

SNP-based phylogenetic analysis revealed three main clades associated with the geographic attributions of the genomes, as previously distinguished by Kay et al. (2014), Tan et al. (2015) and more recently by Georgi et al. (2017): the first clade comprised the isolate Ether isolated from Italy and falling in the Western Mediterranean Region; the second clade comprised the reference 16M isolated from the USA and falling in the American Region; the third clade consisted of the genomes NI, M5–90, M28 (China) and 63–9 and ATCC 23457 (India) in addition to the 16...
sequenced genomes and fall under the East Mediterranean Region including the Middle and Far East. The 16 sequenced genomes clumped together in a separate subclade under the East Mediterranean Region (Fig. 3).

Genome alignment of the 16 B. melitensis against 16M using Mauve and BLAST analysis revealed high levels of conservation between the isolates (Fig. 4; Fig. S1, Supporting Information). When minimal variations were observed, they involved hypothetical proteins, sugar binding and transport and glutamate de-carboxylase enzyme.

**DISCUSSION**

*Brucella melitensis* is the major cause of brucellosis in humans in the Middle East and Lebanon, specifically (Araj and Azzam 1996). Recent analysis has concluded that the increase in brucellosis cases witnessed in Germany between 2014 and 2015 is associated with cases originating from the Middle East (Georgi et al. 2017). In Lebanon, the VFs and genotypic characteristics of *Brucella* were not previously studied and there is a lack of data on the pathogenomics and epidemiology of this important human
pathogen. This study aimed at identifying the species of a collection of 33 Brucella isolates by using various molecular typing methods and analyzing the genomic distinctions of 16 selected B. melitensis isolates through wgSNPs and comparative genome analysis.

Brucella lacks the classical virulence determinants usually found in other pathogens such as cytolsins, exotoxins, fimbriae, plasmids and antigenic variation (Moreno and Moriyón 2002). Accordingly, the VFIs detected in our sequenced isolates included secretion systems specifically the T4SS and flagellar genes. All components of the Brucella T4SS were detected in the sequenced isolates except for virB7, encoding lipoprotein of T4SS that spans the outer membrane and periplasm, previously reported as being absent from B. melitensis strains M5–90 and M28 (Sankarasubramanian et al. 2016). Remnants of genes required for flagellar assembly were also detected and were distributed into three gene clusters, as previously described by Fretin et al. (2005). However, chemotactic proteins CheA, CheB, CheW and CheR, were absent confirming that the flagellar system in Brucella is cryptic (Fretin et al. 2005). Finally, the BvrR/S TCS coding for the homolog of BvrR/S TCS found in B. abortus which is needed for a successful intracellular infection and proper vacuolar maturation was also present (Martínez-Núñez et al. 2010).

The observed fluoroquinolone resistance in some of the isolates (Table S4, Supporting Information) could be linked to the presence of NorM multidrug efflux pump detected in all of the isolates leading to resistance to ofloxacin and ciprofloxacin (Braibant, Guilhouet and Zygmunt 2002). Resistance could also be linked to mutations in the quinolone resistance-determining regions of gyrA, gyrB, parC and parF genes encoding DNA gyrase subunit A, DNA gyrase subunit B, DNA topoisomerase IV subunit A and DNA topoisomerase IV subunit B, respectively (Valdezate et al. 2010).

A total of 30 prophage-associated regions were detected. Brucellaphage BiPBO1, previously isolated from B. inopinata, was detected in nine of the isolates. Brucellaphage BiPBO1s known to infect B. melitensis species and its presence and similarity to an O. anthropi phage suggests that it might be involved in horizontal gene transfer between Brucella and Ochrobactrum spp. (Hammerl et al. 2016). Lactococcus phage j50, previously isolated from Europe and belonging to the 936 group of virulent phages, was found in all but two of the isolates studied (Mahony et al. 2006). Yet, little is known about the role of prophages in Brucella genome evolution and they should be better addressed to understand their role, if any, in pathogenesis (Godfroid et al. 2000).

LPS-related genes and hypothetical proteins were the most abundant on the GIs within the sequenced isolates. Previously, it has been shown that several transposases were responsible for splitting up the original loci of LPS-related genes on chromosome I (Mancilla 2015). Also, chromosomal deletions leading to the loss of LPS genes are related to Brucella carrying unstable, mobile genetic elements (Mancilla 2015). The detected IS711 is one of the most important insertion sequences found in B. abortus and was previously shown to attenuate the RBS1 strain by interrupting the wboA gene involved in LPS O-side chain synthesis (Vemulapalli et al. 1999).

Phylogenetic analysis clustered the sequenced isolates separately from the chosen reference strains (16M, ATCC 23457, 63/9, Ether, NI, M28 and M5–90) originating from different parts of the world. Since PhyloSift is based on concatenated alignment of 37 core genes of a genome, we used RAST to expand our phylogenetic analysis with a number of closely related reference strains (16M and Ether, ATCC 23457 and 63/9, and MS–90 and M28) (Wang et al. 2013; Azam et al. 2016). Our analysis revealed that the genomes were highly conserved and main sequence divergence between strains 16M and Ether was attributed to hypothetical proteins and mobile elements. Along the same line, Wang et al. (2013) showed that genetic variation between strains M5–90 and M28, where M28 was the parent strain of M5–90, was linked to virulence determinants. Those included cold shock proteins, flagellar genes and efflux pumps. Although minimal variation was observed in the genes involved in sugar binding and transport, previous reports suggest that B. melitensis do not differ significantly in the ability to oxidize various sugar substrates (Croch and Elberg 1967). Other sugar binding proteins, such as ChvE, were associated with enhanced virulence in Agrobacterium tumefaciens, a plant pathogen closely related to Brucella (Doty, Chang and Nester 1993).

As previously reported on the phylogenetic relatedness of B. melitensis, wgSNP-based analysis provided better resolution and revealed spatial clustering by dividing the strains into various genotypes based on their geographic locations (Georgi et al. 2017). Isolates belonging to the Western Mediterranean Region, the American Region and the East Mediterranean Region including the Middle and Far East clustered separately.

CONCLUSION
We report on the draft genome sequences of Brucella melitensis isolated from blood and articular fluid. To the best of our knowledge, this is the first extensive genome-wide analysis of Brucella recovered from clinical specimens in Lebanon. The genome-wide analysis showed, and despite similarities in VFs and gene content, differences in resistance patterns, prophages, ISs, GIs, virulence determinants carried on GIs and the possible role of unstable mobile genetic elements in chromosomal deletions. wgSNP-based analysis provided effective means of discriminating the native geographical origin of the pathogens based on spatial clustering and proved to be a powerful tool for epidemiological studies and tracing outbreaks. The draft genomes generated are a valuable reference for future studies targeting virulence and vaccine efficacy.

SUPPLEMENTARY DATA
Supplementary data are available at FEMSLE online.

Conflicts of interest. None declared.

REFERENCES


Moreno E. Retrospective and prospective perspectives on zoonotic brucellosis. Front Microbiol 2014; 5, DOI: 10.3389/fmicb.2014.00213


