Genomics 106 (2015) 43–51

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Comparison of *Brucella canis* genomes isolated from different countries shows multiple variable regions

Miryan Margot Sánchez-Jiménez [a], Juan Pablo Isaza [b,c], Juan F. Alzate [b,c], Martha Olivera-Angel [d,*]

[a] Grupo Biogénesis-Vericel, DrSci in Animal Sciences, Escuela de Medicina Veterinaria, Facultad de Ciencias Agrarias, Universidad de Antioquia, Medellín, Colombia
[b] Centro Nacional de Secuenciación Genómica—CNSG, Universidad de Antioquia, Medellín, Colombia
[c] Grupo de Parasitología, Facultad de Medicina, Universidad de Antioquia, Medellín, Colombia
[d] Grupo Biogénesis-Vericel, Escuela de Medicina Veterinaria, Facultad de Ciencias Agrarias, Universidad de Antioquia, Medellín, Colombia

ARTICLE INFO

Article history:
Received 27 November 2014
Accepted 18 March 2015
Available online 24 March 2015

Keywords:
Genome
Sequencing
Brucella canis

ABSTRACT

Brucella canis is a pathogenic bacterium for dogs and its zoonotic potential has been increasing in recent years. In this study, we report the sequencing, annotation and analysis of the genome of *Brucella canis* strain Oliveri isolated from a dog in a breeding kennel in Medellín, Colombia, South America.

Whole genome shotgun sequencing was carried out using the ROCHE 454 GS FLX Titanium technology at the National Center for Genomic Sequencing—CNSG in Medellin, Colombia. The assembly procedure was performed using Newbler v2.6. In the genome annotation process, each contig was analyzed independently using as reference *Brucella suis* ATCC 1330 chromosomes.

This new genome could be useful for the development of diagnostic tools and for vaccines search as well, in order to reduce the health impact of this infection in both, dogs and humans. The sequence was deposited in EMBL-EBI with accession numbers HG803175 and HG803176 for chromosomes 1 and 2, respectively.

1. Introduction

Brucella is a genus of Gram-negative intracellular coccobacilli facultative intracellular that belongs to the Proteobacteria phylum, comprised of ten well-characterized species. Some species of the genus *Brucella* can infect a wide range of animal hosts, including humans. Due to the relevance of the genus in public health and the need for basic evolutionary studies, a great amount of research is being performed including genome sequencing and analysis of reference strains.

Reports of genomic analyses have included *Brucella melitensis* [1], *Brucella suis* [2–4], *Brucella abortus* [5], *Babesia microti* [6], *Brucella ovis* [7], *Brucella canis* [8] and *Brucella pinnipedialis* [9].

B. canis is a veterinary pathogen that affects the reproductive tract of dogs and can be isolated from blood and other body fluids or tissues of infected animals [10–12].

This bacterium can be transmitted to humans exposed to infected dog secretions, or bacterial laboratory cultures [13], inducing many symptoms from mild flu-like to severe complications [14–19].

The phylogenomic analysis of *B. canis* has shown that this agent is closely related to *B. suis* and that the former originated from the latter around 22,000 years ago [20]. *B. canis* genome sequencing isolated in China and United States were previously analyzed [8,21–23].

Brucella genomes lack plasmidic DNA and contain two chromosomes of approximately 2.1 and 1.2 Mbp in length. Both carry ribosomal gene clusters and approximately 3200 protein-coding genes have been detected in each species. In general, *Brucella* genomes are highly conserved, with less than 6% nucleotide sequence variation, attributed to the recent origin of the genus [5]. Methodologies for sequencing *Brucella* genomes have evolved from the Sanger capillary technique [5], to the 454 WGS (Whole Genome Sequencing) methodology combined with Sanger sequencing to fill the gaps [24], up to date, when *Brucella* genome studies of different strains have been done using the Illumina platform [25].

In the present article we report the full genome sequence of a Colombian isolate of *B. canis* str. Oliveri (accession number HG803175 and HG803176) previously reported as Group 2 [26], using the 454 FLX titanium technologies with the whole genome shotgun strategy. After read assembly, 34 contigs were obtained with an average coverage of 28X. Chromosome 2 was finished using PCR and Sanger sequencing; chromosome 1 was partially degapped using the same strategy. Comparative analysis with *Brucella* reference genomes showed several indel events, some being *B. canis*-specific and others specific of the Colombian strain.

* Corresponding author at: Carrera 75 # 65-87 (Ciudadela Robledo U. de A.) Bloque 46, Oficina 202, Medellín, Colombia. Fax: +57 42199127.
E-mail address: martha.olivera@udea.edu.co (M. Olivera-Angel).

http://dx.doi.org/10.1016/j.ygeno.2015.03.004
0888-7543/

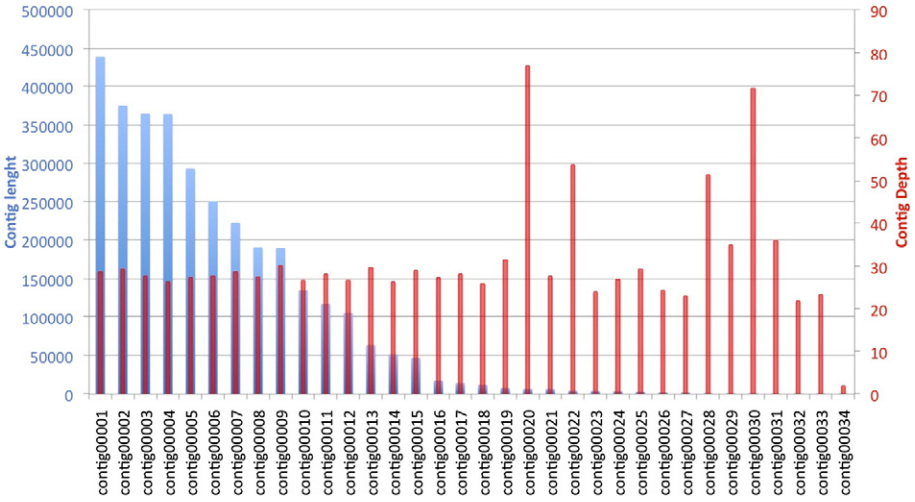


Fig. 1. Contigs length and depth generated using the Newbler v2.6 with an average coverage of 28X.

2. Material and methods

2.1. Bacterial Culture and DNA extraction

The *B. canis* strain from a blood culture in tryptic soy broth (Becton Dickinson, Franklin Lakes, NJ, USA), was isolated in tryptic soy agar (Becton Dickinson, Franklin Lakes, NJ, USA), from a dog of a kennel in Medellín, Colombia. The strain was confirmed as *B. canis* using a biochemical test such as urease production and molecular tests [26]. One colony was then inoculated in tryptic soy broth and incubated for 2 days at 37 °C; this liquid culture was used for genomic DNA extraction. A column-based method was used following the manufacturer's instructions (QIAGEN, DNeasy Blood & Tissue Kit, CAT# 69504). DNA concentration was measured using UV light absorption at 260 nm and Picogreen fluorescence (INVITROGEN, Quant-iT™ PicoGreen® dsDNA Assay Kit, CAT# 69504).

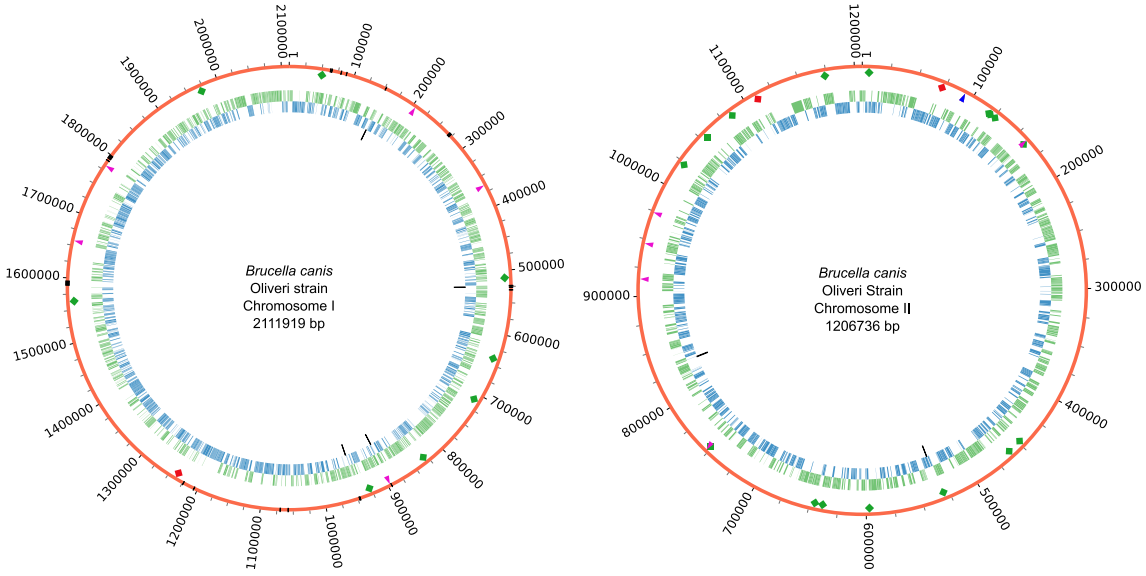
2.2. Whole genome shotgun sequencing and assembly

Whole genome shotgun sequencing was carried out using the ROCHE 454 GS FLX TITANIUM technology at the National Center for Genomic Sequencing—CNSG, Universidad de Antioquia, Medellin, Colombia, following all the standard protocols. One fourth of PTP (picotiter plate) was used for sequencing and 289,912 reads were obtained, representing 93,261,046 raw bases. Read dataset quality was analyzed using FASTQC software, to quickly obtain some summary statistics to check the quality of the run [27]. The assembly procedure was performed using Newbler v2.6 with default options for *de novo* genome assembly. Contig scaffolding was carried out using the ABACAS algorithm [28] with *B. suis* 1330 chromosomes, as reference.

Chromosomes 2 and 1 were completed using PCR, cloning and capillary sequencing for both strands. Gap flanking regions were used to design the primers (see the Table S1 in ref [29]).

2.3. Genome annotation

For the genome annotation process, each chromosome was analyzed independently. The RATT algorithm was used for automatic annotation using the *B. suis* ATCC 1330 genome with Genbank accession numbers NC_004310.3 and NC_004311.2 as reference. Each chromosome was then manually curated using the ARTEMIS software. Some of the



Figs. 2 and 3. Graphical circular representation of chromosomes I and II of *B. canis* str. Oliveri built using Circos software. Blue triangle: specific insertion in str. Oliveri; Red diamond: specific deletion in str. Oliveri; Green diamond: deletion in all *B. canis* strains; Pink triangle: insertion in all *B. canis* strains. Inner chromosome: Green: CDS Forward; Blue: CDS reverse. Black lanes: Position of mobile genetic element IS711.

Table 1
List of deletions and insertions common to *B. canis* species in chromosomes 1 and 2.

Deletions common to *B. canis* species				
Region coordinates in *B. suis* 1330				
Start base	End base	Length (bp)	Affected feature	Predicted protein
Chromosome 1				
50,944	51,027	83	BR0047	SH3 type 3 domain-containing protein
511,446	511,803	357	BR0510	Capsular polysaccharide biosynthesis protein capD
639,583	639,681	98	BR0648	Uracil-DNA glycosylase, family 4
709,618	709,914	296	BR0725	N-acetyltransferase GCN5
830,809	830,936	127	BCA52141_I0356	Hypothetical protein
928,033	928,046	13	BR0956	Molybdopterin-guanine dinucleotide biosynthesis protein MobB
1,563,635	1,563,759	124	BR1617	Hypothetical protein
1,971,863	1,971,866	3	BR2047	Coenzyme A transferase
Chromosome 2				
5834	5842	8	BRA0008	Inner-membrane translocator
119,509	119,517	8	BRA0120	Hypothetical protein
120,884	120,885	1	BCA52141_II1041	Flagellar motor switch protein G
125,838	125,867	29	BRA0128	Flagellar basal body rod protein FlgF
159,502	159,522	20	BRA0173	Outer membrane autotransporter
448,126	448,167	41	Intergenic	
461,191	461,202	11	BRA0474	Transcriptional regulator
530,060	530,071	11	BRA0549	Hypothetical protein
597,027	597,059	32	BRA0609	50S ribosomal protein L33
638,523	638,529	6	Intergenic	
645,465	645,481	16	Intergenic	
752,583	752,584	1	BRA0772	Hypothetical protein
1,023,225	1,023,232	7	Intergenic	
1,055,237	1,055,238	1	Intergenic	
1,084,170	1,084,171	1	BRA1097	Glutathione-binding protein GsiB
1,173,676	1,173,687	11	BRA1172	Amidohydrolase 3
Insertions common to *B. canis* species				
Region coordinates in *B. canis* str. Oliveri				
Chromosome 1				
Start base	End base	Length (bp)	Affected feature	Predicted protein
204,542	204,543	1	BR0192	Hypothetical protein
365,915	365,916	1	BR0353	Cysteine desulfurase
485,335	485,394	59	BR0485	Hypothetical protein
896,401	896,417	16	BR0923	Rhamnosyltransferase
1,657,489	1,657,490	1	BR1711	Inositol-phosphate phosphatase
1,782,052	1,782,469	189	BR1846	Membrane protein
Chromosome 2				
159,323	159,328	5	BRA0173	Outer membrane autotransporter
753,091	753,104	13	BRA0772	Hypothetical protein
915,093	915,094	1	BRA0929	Bifunctional imidazolonepropionase/histidine ammonia-lyase
946,208	946,209	1	BRA0957	Inner-membrane translocator
973,988	973,992	4	BRA0987	Cobalamin synthesis protein P47K

features had to be edited since the RATT tool incorrectly identified many non-ATG starting codons or skipped several genes.

2.4. Genome comparison

The program MAUVE v2.3.1 served for entire genome alignment and comparative analysis of Indels and SNPs between *B. canis* str. Oliveri, *B. canis* ATCC 23365, *B. canis* HSK A52141 and *B. suis* 1330. Output tables were filtered and edited using custom PERL and Python scripts. FASTA package was used for global comparisons of genes and predicted protein sequences.

3. Results

3.1. Genome assembly and annotation

Whole genome shotgun in one quarter PTP of 454 FLX Titanium produced a total of 289,912 reads with a GC content of 57%, an average length of 321 and a Phred quality score of 31. Total read bases summed 93,326,476 with no ambiguous nucleotides. Genome assembly with Newbler v2.6 produced 34 contigs with an average coverage of 28X, only one contig presented an aberrant read depth of 1.9X (Fig. 1). This contig was also very small with only 141 bases; it might represent an aberrant product of the assembler and therefore was excluded from further analyses (see the Table S2 in ref [29]). The length of the 33 remaining contigs ranged from 552 to 439,538 bases, the N50 genome assembly value was of 294,016 bases with 99.91% of the bases in the assembly with a Q40 quality value. Repetitive elements within the genome suffer compression during the assembly process, resulting in contigs with a higher read depth, often multiples, of the average contig depth. Contigs 20, 22, 28, and 30 clearly presented such phenomena (see the Table S2 in ref [29]). BLAST comparisons of such contigs

Table 2
List of *B. canis* str. Oliveri specific indels in chromosomes 1 and 2.

B. canis str. Oliveri specific indels					
Chromosome	Start base	End base	Length	Affected feature	Indel
1	1,241,093	1,241,311	218	del tRNA-GLU	Deletion
1	1,241,090	1,241,092	2	Intergenic	Deletion
2	1,110,224	1,110,225	1	Intergenic	Deletion
2	73,078	73,085	7	Intergenic	Deletion
2	92,554	92,557	2	BRA0095	Insertion

Table 3
List of divergent proteins between *B. canis* strain Oliveri vs *B. canis* HSK A52141.

Divergent proteins between *B. canis* strain Oliveri vs *B. canis* HSK A52141			
Chromosome 1			
B. canis strain Oliveri	*B. canis HSK A52141*	Percentage divergence	Predicted protein
BR0013	BCAN_A0992	16.95	Hypothetical protein
BR0019	BCAN_A2193	19.84	Ligase
BR0030	BCAN_A1832	23.88	Hypothetical protein
BR0046	BCAN_A0356	25	Hypothetical protein
BR0069	BCAN_A1989	20.25	Error prone DNA polimerase
BR0073	BCAN_A0348	19.7	Invasion associated locus B family protein
BR0138	BCAN_A1331	25	Hypothetical protein
BR0189	BCAN_A0544	28.57	Hypothetical protein
BR0192	BCAN_A1585	21.11	Hypothetical protein
BR0198	BCAN_A1505	22.22	Hypothetical protein
BR0208	BCAN_A0675	20.59	Hypothetical protein
BR0209	BCAN_A0688	20.69	Hypothetical protein
BR0242	BCAN_A1997	11.43	Hypothetical protein
BR0257	BCAN_A0101	24.44	Hypothetical protein
BR0266	BCAN_A1893	26.32	Hypothetical protein
BR0332	BCAN_A0472	25	Hypothetical protein
BR0338	BCAN_A1822	21.74	Hypothetical protein
BR0343	BCAN_A0335	21.37	Tripartite ATP-independent periplasmic transporter DctQ component
BR0377	BCAN_A0835	28.57	Hypothetical protein
BR0379	BCAN_A0576	29.73	Hypothetical protein
BR0388	BCAN_A1947	22.82	Glycosyltransferase
BR0390	BCAN_A1059	17.78	Hypothetical protein
BR0419	BCAN_A1644	27.27	Hypothetical protein
BR0422	BCAN_A0396	19.18	Hypothetical protein
BR0423	BCAN_A1551	41.89	Cold shock protein CspA
BR0451	BCAN_A1521	19.42	Cyclopropane-fatty-acyl-phospholipid synthase
BR0474	BCAN_A1196	19.15	Surf1 protein
BR0502	BCAN_A0776	22.45	Hypothetical protein
BR0503	BCAN_A0185	23.33	Hypothetical protein
BR0511	BCAN_A0148	23.21	Glycoside hydrolase
BR0513	BCAN_A0537	21.9	Transposase
BR0515	BCAN_A1781	22.5	Transposase
BR0530	BCAN_A0535	35.26	Transposase
BR0551	BCAN_A1993	24.32	Hypothetical protein
BR0576	BCAN_A1288	26.09	Hypothetical protein
BR0631	BCAN_A1196	33.33	Hypothetical protein
BR0632	BCAN_A1361	23.53	Hypothetical protein
BR0638	BCAN_A0300	26.53	Hypothetical protein
BR0641	BCAN_A0079	27.14	Hypothetical protein
BR0645	BCAN_A1210	19.44	Hypothetical protein
BR0655	BCAN_A0181	21.56	Oxygen-independent coproporphyrinogen III oxidase
BR0676	BCAN_A1262	19.15	Hypothetical protein
BR0712	BCAN_A1672	40	Hypothetical protein
BR0748	BCAN_A1689	19.34	Polyphosphate kinase
BR0763	BCAN_A0472	24.24	Hypothetical protein
BR0764	BCAN_A1210	27.66	Hypothetical protein
BR0775	BCAN_A1296	26.32	Hypothetical protein
BR0781	BCAN_A1946	21.66	CPS biosynthesis glycosyltransferase
BR0784	BCAN_A1210	22.22	Hypothetical protein
BR0795	BCAN_A1433	11.76	Hypothetical protein
BR0828	BCAN_A1822	19.57	Extensin family protein
BR0846	BCAN_A1154	19.51	Hypothetical protein
BR0847	BCAN_A1893	20.51	Hypothetical protein
BR0864	BCAN_A0654	18.81	Aspartyl protease
BR0901	BCAN_A0103	22.92	Hypothetical protein
BR0903	BCAN_A1205	23.53	Hypothetical protein
BR0914	BCAN_A1225	28.57	Hypothetical protein
BR0921	BCAN_A1004	17.65	Hypothetical protein
BR0922	BCAN_A1780	24.47	Peptidase
BR0939	BCAN_A0889	36.07	Hesb protein
BR0941	BCAN_A1997	17.14	hypothetical protein
BR0947	BCAN_A0814	21.43	hypothetical protein
BR0956	BCAN_A2164	23.61	Molybdopterin-guanine dinucleotide biosynthesis protein MobB
BR0957	BCAN_A0888	22.36	Molybdopterin-guanine dinucleotide biosynthesis protein MobA
BR0963	BCAN_A1262	24.49	Hypothetical protein
BR0999	BCAN_A2084	20	Hypothetical protein
BR1007	BCAN_A0394	23.47	Hypothetical protein
BR1009	BCAN_A1360	27.5	Hypothetical protein
BR1023	BCAN_A0312	31.58	Hypothetical protein
BR1037	BCAN_A1296	30.23	Hypothetical protein
WP_006199373	BCAN_A0262	27.56	Hypothetical protein
BR1107	BCAN_A1997	26.19	Hypothetical protein

Table 3 (*continued*)

Divergent proteins between *B. canis* strain Oliveri vs *B. canis* HSK A52141			
Chromosome 1			
B. canis strain Oliveri	*B. canis HSK A52141*	Percentage divergence	Predicted protein
BR1155	BCAN_A0690	26.67	Hypothetical protein
BR1240	BCAN_A1296	28.95	Hypothetical protein
BR1244	BCAN_A1670	20	Hypothetical protein
BR1253	BCAN_A0879	26	Hypothetical protein
BR1257	BCAN_A1360	20	Hypothetical protein
BR1263	BCAN_A1059	21.74	Hypothetical protein
BR1279	BCAN_A0725	26.09	Hypothetical protein
BR1280	BCAN_A1004	28.57	Hypothetical protein
BR1316	BCAN_A0364	26.42	Entericidin EcnAB
BR1317	BCAN_A1720	18.92	Hypothetical protein
BR1333	BCAN_A0574	22.86	Hypothetical protein
BR1335	BCAN_A0067	26.03	Hypothetical protein
BR1341	BCAN_A0472	27.27	Hypothetical protein
BR1350	BCAN_A1711	20.75	ABC transporter
BR1353	BCAN_A0683	21.28	Hypothetical protein
BR1364	BCAN_A1597	22.35	Cobalt ABC transporter substrate-binding protein CbiN
BR1442	BCAN_A0022	22.58	Hypothetical protein
BR1484	BCAN_A0576	55.81	Hypothetical protein
BR1494	BCAN_A1505	20.93	Hypothetical protein
BR1496	BCAN_A1533	30.68	Hypothetical protein
BR1513	BCAN_A0707	18.79	Lytic transglycosylase
BR1515	BCAN_A1924	19.15	Hypothetical protein
BR1520	BCAN_A0808	22.39	Marr family transcriptional regulator
BR1523	BCAN_A1822	54.35	Hypothetical protein
BR1534	BCAN_A1459	22.18	ATP-binding protein
BR1548	BCAN_A1997	27.5	Hypothetical protein
BR1565	BCAN_A1911	24.49	Hypothetical protein
BR1576	BCAN_A0909	23.53	Hypothetical protein
BR1577	BCAN_A0101	21.43	Hypothetical protein
BR1578	BCAN_A1011	23.53	Hypothetical protein
BR1587	BCAN_A1972	19.81	Pyridoxamine 5′-phosphate oxidase-like FMN-binding protein
BR1617	BCAN_A0690	21.57	Hypothetical protein
BR1622	BCAN_A0121	38.2	Outer-membrane immunogenic protein
BR1624	BCAN_A1361	24.24	Hypothetical protein
BR1625	BCAN_A1997	32.43	Hypothetical protein
BR1632	BCAN_A2043	24.55	Enolase
BR1633	BCAN_A1632	22.92	Hypothetical protein
BR1644	BCAN_A0576	21.21	Hypothetical protein
BR1645	BCAN_A0150	19.44	Hypothetical protein
BR1663	BCAN_A2084	23.4	Hypothetical protein
BR1674	BCAN_A1911	18	Hypothetical protein
BR1688	BCAN_A1327	23.29	Bile acid sodium symporter
BR1694	BCAN_A1402	39.13	Hypothetical protein
BR1709	BCAN_A1296	28.21	Hypothetical protein
BR1726	BCAN_A0495	26.47	Hypothetical protein
BR1746	BCAN_A2188	18.87	50S ribosomal protein L36
BR1760	BCAN_A0675	24.24	Hypothetical protein
BR1770	BCAN_A1124	20.04	ATP-dependent helicase
BR1771	BCAN_A0982	23.94	Hypothetical protein
BR1773	BCAN_A1517	28.07	Hypothetical protein
BR1786	BCAN_A0543	22.22	Hypothetical protein
BR1795	BCAN_A1440	21.65	GNTR family transcriptional regulator
BR1797	BCAN_A0312	17.65	Hypothetical protein
BR1818	BCAN_A1563	19.27	Hypothetical protein
BR1827	BCAN_A1872	22.06	Hypothetical protein
BR1840	BCAN_A0180	21.07	Ribosomal rna large subunit methyltransferase H
BR1863	BCAN_A1121	29.79	Hypothetical protein
BR1883	BCAN_A0150	28.57	Hypothetical protein
BR1950	BCAN_A0715	14.29	Hypothetical protein
BR2000	BCAN_A1265	21.43	Hypothetical protein
BR2002	BCAN_A0725	22.22	Hypothetical protein
BR2012	BCAN_A0868	19.09	Hypothetical protein
BR2013	BCAN_A0050	19.63	Outer membrane autotransporter
BR2020	BCAN_A0805	20.12	Hypothetical protein
BR2047	BCAN_A1650	19.64	Coenzyme A transferase
BR2054	BCAN_A0585	21.5	Hypothetical protein
BR2079	BCAN_A1127	24.8	ATP-dependent protease ATP-binding subunit HslU
BR2088	BCAN_A1644	20.41	Hypothetical protein
BR2093	BCAN_A1456	20.21	Hypothetical protein
BR2098	BCAN_A1813	25	Hypothetical protein
BR2104	BCAN_A0700	30.36	Hypothetical protein
BR2109	BCAN_A2132	20.49	Hypothetical protein

(*continued on next page*)

Table 3 (*continued*)

Divergent proteins between *B. canis* strain Oliveri vs *B. canis* HSK A52141			
Chromosome 1			
B. canis strain Oliveri	*B. canis HSK A52141*	Percentage divergence	Predicted protein
BR2133	BCAN_A1513	20.42	Transglutaminase
BR2134	BCAN_A1425	18.37	Hypothetical protein
BR2147	BCAN_A1154	19.51	Hypothetical protein
BR2154	BCAN_A1203	31.25	Hypothetical protein
Chromosome 2			
B. canis strain Oliveri	*B. canis* HSK *A52141*	Percentage divergence	Predicted function
BRA0008	BCAN_B0459	24.82	Inner-membrane translocator
BRA0009	BCAN_B0982	29.03	ABC transporter
BRA0016	BCAN_B1076	18.55	Isochorismate synthase
BRA0027	BCAN_B0529	20.59	Branched-chain amino acid ABC transporter substrate-binding protein
BRA0082	BCAN_B0647	21.24	Response regulator containing CheY-like receiver, AAA-type ATPase, and DNA-binding domains
BRA0130	BCAN_B0545	19.79	Hypothetical protein
BRA0138	BCAN_B0665	27.08	Hypothetical protein
BRA0140	BCAN_B0741	19.51	Hypothetical protein
BRA0142	BCAN_B0509	28.95	Hypothetical protein
BRA0165	BCAN_B0243	30.3	Hypothetical protein
BRA0169	BCAN_B0290	33.87	Hypothetical protein
BRA0173	BCAN_B0730	18.02	Outer membrane autotransporter
BRA0198	BCAN_B0179	16.67	Hypothetical protein
BRA0200	BCAN_B1042	23.91	Hypothetical protein
BRA0219	BCAN_B0876	20.96	MucK, cis,cis-muconate transport protein
BRA0235	BCAN_B0612	22.73	Hypothetical protein
BRA0241	BCAN_B0793	21.95	Hypothetical protein
BRA0253	BCAN_B1083	22.17	Oxidoreductase
BRA0273	BCAN_B0529	25.71	Hypothetical protein
BRA0295	BCAN_B0557	21.47	Peptidyl-prolyl cis-trans isomerase C
BRA0303	BCAN_B0260	20	Hypothetical protein
BRA0312	BCAN_B0741	20.93	Hypothetical protein
BRA0329	BCAN_B0156	21.76	Binding-protein-dependent transport system inner membrane protein
BRA0342	BCAN_B0873	20.59	Hypothetical protein
BRA0400	BCAN_B0134	19.79	Branched-chain amino acid ABC transporter substrate-binding protein
BRA0441	BCAN_B0948	24.07	Hypothetical protein
BRA0458	BCAN_B0547	24.32	Membrane protein
BRA0475	BCAN_B0735	19.64	Hypothetical protein
BRA0495	BCAN_B0793	19.51	Hypothetical protein
BRA0498	BCAN_B0612	25.53	Hypothetical protein
BRA0506	BCAN_B0376	19.82	Hypothetical protein
BRA0517	BCAN_B0837	24	Hypothetical protein
BRA0523	BCAN_B0593	18.03	Hypothetical protein
BRA0541	BCAN_B0460	23.96	Hypothetical protein
BRA0552	BCAN_B0627	20.95	Hypothetical protein
BRA0617	BCAN_B0936	24.39	Hypothetical protein
BRA0619	BCAN_B0529	20	Hypothetical protein
BRA0620	BCAN_B0179	22.5	Hypothetical Protein
BRA0638	BCAN_B0665	26.09	3-oxoadipate CoA-transferase
BRA0656	BCAN_B1213	22.49	sn-glycerol-3-phosphate transport system permease ugpA
BRA0674	BCAN_B0179	20	Hypothetical protein
BRA0680	BCAN_B0509	21.21	Hypothetical protein
BRA0698	BCAN_B0623	20	Hypothetical protein
BRA0721	BCAN_B0697	22	Hypothetical protein
BRA0724	BCAN_B1061	20.45	Hypothetical protein
BRA0747	BCAN_B0936	17.5	Hypothetical protein
BRA0757	BCAN_B0755	20.45	Transport protein; periplasmic binding protein
BRA0756	BCAN_B0687	25	Hypothetical protein
BRA0762	BCAN_B0509	24.24	Hypothetical protein
BRA0772	BCAN_B0547	22.16	Unknown function; bacterial intein-like
BRA0822	BCAN_B0086	18.98	Contractile protein
BRA0829	BCAN_B0610	25	Hypothetical protein
BRA0846	BCAN_B0429	19.56	Hypothetical protein
BRA0855	BCAN_B0901	22.49	4-hydroxybenzoate polyprenyltransferase-related prenyltransferase
BRA0862	BCAN_B0699	25.64	Hypothetical protein
BRA0863	BCAN_B0006	20	Hypothetical protein
BRA0864	BCAN_B0270	24.95	Erythritol kinase
BRA0907	BCAN_B0309	20.97	Hypothetical protein
BRA0917	BCAN_B0709	21.82	Hypothetical protein
BRA0929	BCAN_B1002	21.9	Imidazolonepropionase
BRA0946	BCAN_B0169	19.64	Hypothetical protein
BRA0965	BCAN_B0339	29.03	Hypothetical protein
BRA0966	BCAN_B0604	25.86	30S ribosomal protein S21
BRA0975	BCAN_B1149	27.08	Hypothetical protein

Table 3 (*continued*)

Chromosome 2			
B. canis strain Oliveri	*B. canis* HSK *A52141*	Percentage divergence	Predicted function
BRA0977	BCAN_B1149	23.4	Hypothetical protein
BRA0980	BCAN_B0137	17.57	Hypothetical protein
BRA0981	BCAN_B1061	21.43	Hypothetical protein
BRA0982	BCAN_B0723	20.43	Hypothetical protein
BRA0999	BCAN_B0516	20.56	TonB-dependent copper receptor
BRA1030	BCAN_B1084	18.42	Hypothetical protein
BRA1033	BCAN_B0206	18.65	tRNA pseudouridine synthase A
BRA1059	BCAN_B0783	21.51	Endonuclease/exonuclease/phosphatase family protein
BRA1070	BCAN_B0095	23.61	Hypothetical protein
BRA1091	BCAN_B0529	25	Hypothetical protein
BRA1097	BCAN_B0549	22.84	Glutathione-binding protein GsiB
BRA1101	BCAN_B1101	44.81	Oligopeptide ABC transporter ATP-binding protein
BRA1112	BCAN_B0565	18.78	Hydrolase
BRA1117	BCAN_B0793	14.29	Hypothetical protein
BRA1121	BCAN_B0374	25.17	Hypothetical protein
BRA1126	BCAN_B0515	22.73	Hypothetical protein
BRA1133	BCAN_B0484	18.75	Flagellar biosynthesis protein FliQ
BRA1153	BCAN_B1011	18.55	Hypothetical protein
BRA1191	BCAN_B0793	18.6	Hypothetical protein
BRA1198	BCAN_B0651	21.71	Cadmium-translocating P-type ATPase

showed them to be ribosomal, IS3 (chr2) and IS711 (portion) sequences, repeated two or more times in the genome.

The following step involved contig reordering based on its position within the chromosome. For this purpose genome alignment with the close reference strain, *B. suis* 1330, was carried out using the MUMMER genome aligner. Once the contigs had been assigned and positioned, the ABACAS automatic algorithm reordered and constructed pseudochromosome of both molecules. Contigs 28, 29, and 32 were excluded by ABACAS in the pseudochromosome (see the Table S2 in ref [29]).

For further confirmation of the contig scaffolding with the reference chromosomes, PCR amplification and capillary sequencing was carried out for both chromosomes. Chromosome 2 of *B. canis* str. Oliveri was fully validated and neither ambiguous bases nor gaps remained. In chromosome 1, comparative analysis with other *B. canis* genomes, showed that gaps between contigs were related to ribosomal gene clusters, to part of the IS711 insertion sequence, to the translation elongation factor Tu gene and to some repeated sequences at the ends of neighboring contigs. These gaps were partially removed and those belonging to rDNA, repetitive motifs, *tuf1* and *tuf2* and part of the IS711 element were not filled. Twenty gaps remained in this chromosome. The final sizes of the chromosomes I and II were 2,111,919, and 1,206,736 bp, respectively.

Genome annotation was carried out using the automated annotation transfer tools of the Sanger institute, RATT [30]; the *B. suis* 1330 genome was used as reference. The complete chromosome manually curated allowed the annotation of features that were lost by the RATT script or wrongly transferred. Finally, a total of 2130 and 1158 CDS were annotated in chromosomes 1 and 2 of *B. canis* str. Oliveri, respectively (Figs. 2 and 3) built using Circos [31].

3.2. B. canis str. Oliveri genome structure

The genome showed perfect synteny with other *B. canis* and *B. suis* strains whose genomes are completely annotated in GenBank. The total length of the genome is 3,318,655 bp. A total of 5 insertion elements, including IS711, were identified in chromosomes 1 and 2 in chromosome 2. These are identical to the ones previously reported in *B. canis* ATCC 23365. *B. canis* species-specific deletions and insertions were detected in both chromosomes, which ranged from 1 to 357 bp, most of them involving coding regions (Table 1). Each chromosome in *B. canis* str. Oliveri presented two specific sequence deletions. In chromosome 1, deletions involved 218 and 2 bases affecting the tRNAGLU gene and one intergenic region, respectively. In chromosome 2, deletions were of 1 and 7 bases in intragenic regions (Table 2. and Figs. 2 and 3).

3.3. Protein comparison

Predicted peptides of *B. canis* str. Oliveri were aligned (global alignment) with its respective ortholog in *B. canis* HSK A52141, *B. canis* ATCC 23365 and *B. suis* 1330. This analysis showed that proteins sharing 100% identity were 82% with *B. canis* ATCC 23365, 63% with *B. canis* HSK A52141 and 71% when compared to *B. suis* 1330 (Table 3). The lower protein identity with the HSK A52141 is noteworthy. When compared to other *B. canis* strains and *B. suis*, inspection of the annotated CDS features in the HSK A52141 genome showed several differences at the start codon of several genes; this could explain the protein identity differences observed between the two isolates (Table 3).

We considered as very divergent proteins those that had an amino acid identity below 60% with its respective orthologs between *B. canis* str. Oliveri and the other reference strains compared. In this subclass, we observed 7% of divergence when compared to *B. canis* ATCC 23365; 14% against *B. canis* HSK A52141 and 2.25% against *B. suis* 1330 (Table 4).

3.4. Single nucleotide polymorphisms (SNPs)

SNPs between *B. canis* str. Oliveri and the other three reference genomes studied were calculated. For chromosome 1; 90, 108, and 1408 SNPs were found compared with *B. canis* ATCC 23365, *B. canis* HSK A52141 and *B. suis* 1330, respectively. In chromosome 2, in the same order, 82, 81, and 918 SNPs were detected. This data correlates with the percentage of proteins that showed 100% identity, indicating that *B. canis* str. Oliveri is closer to *B. canis* ATCC 23365. Forty-eight SNPs were unique to the Colombian *B. canis* str. Oliveri (see the Table S3 in ref [29]).

4. Discussion

Bacterial genome sequencing has opened a new era in the analysis of pathogenic bacteria.

Regarding genome structure, *B. canis* str. Oliveri has a GC-content of 57%, similar to that of both Chinese isolates: the dog strain 118 (57, 27%) [21] and the human strain BCB018 [23].

Table 4
Predicted protein comparison between *B. canis* str. Oliveri and other *B. canis* isolates and *B. suis 1330*.

Predicted protein comparison between *B. canis* str. Oliveri and other *B. canis* isolates and *B. suis 1330*			
Global percentage of identity. Number of proteins within ranges: 100%/99.9–80%/79.9–60%/<59.9%			
	B. canis ATCC23365	*B. canis* HSK A52141	*B. suis* 1330
Chromosome 1	82.02/10.47/0.52/6.99	63.36/21.99/1.46/14.19	70.55/26.54/0.66/2.25
Chromosome 2	81.07/10.71/0.95/7.26	59.6/21.87/1.47/17.03	61.66/32.82/1.38/4.15

Regarding sizes, when these three strains were compared, str. Oliveri showed a larger genome size; 3,318,655 versus 3,234,827 bp for strain 118 and 3,247,324 bp for strain BCB018.

As was described in earlier investigations, within the *Brucella* genome, it is common to find deletion events, more frequently than insertions [21–23,32,33]. Apart from the indels common to all the *B. canis* genomes examined, there were specific mutations of the Colombian isolate. This map of variations could be used as candidate for molecular epidemiology studies.

The differences found may be explained by the bacterial adaptation to hosts and environments that produce genetic changes and therefore loss of genomic material unnecessary in the pathogenesis process, or produce genetic polymorphisms, as has been reported by other authors [21,22,34].

Genome sequencing of several *B. canis* strains around the world, and also from different hosts, such as dogs and humans, it is very important to establish which characteristics are conserved or different between the strains. Perhaps, depending on their environment and host, the pathogenic mechanisms and co-evolution processes could have generated small differences in the genetic material of the bacterium. These changes could be useful in the future to determine what generates the differences in virulence and host specificity [35].

The genome sequence of *B. canis* str. Oliveri can be used as the starting point in the development of specific diagnostic tools for early detection of infection in dogs and humans, as well as in the development of vaccines, all of which could help avoid the epidemiological, public health and economic complications caused by the disease.

5. Conclusions

We report here the annotated genome sequence of *B. canis* str. Oliveri, isolated from a dog in Medellín, Colombia. It shows unique genomic characteristics that indicate that within a species, there are differences in genome structure associated to its geographical origin. This genome could also be useful in the development of diagnostic tools and vaccines, in order to reduce health complications of this infection in dogs and humans.

Conflict of interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by Universidad de Antioquia sustainability strategy 2013–2014, COLCIENCIAS through a contractual research grant Generación del Bicentenario 2010 for PhD students to MM Sánchez-Jiménez and Convenio Especial de Cooperación No. 767 from Colciencias-Colombia and Vicerrectoría de Investigación, Universidad de Antioquia.

We also thank July Duque, Veronica Bermudez, and Juan Vélez for their technical assistance and biologist Juan Jacobo de la Cuesta for his help in the translation of this paper.

References

[1] V.G. DelVecchio, V. Kapatral, R.J. Redkar, G. Patra, C. Mujer, T. Los, et al., The genome sequence of the facultative intracellular pathogen *Brucella melitensis*, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 443–448, http://dx.doi.org/10.1073/pnas.221575398.
[2] I.T. Paulsen, R. Seshadri, K.E. Nelson, J.A. Eisen, J.F. Heidelberg, T.D. Read, et al., The *Brucella suis* genome reveals fundamental similarities between animal and plant pathogens and symbionts, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 13148–13153, http://dx.doi.org/10.1073/pnas.192319099.
[3] H. Tae, S. Shallom, R. Settlage, D. Preston, L.G. Adams, H.R. Garner, Revised genome sequence of *Brucella suis* 1330, J. Bacteriol. 193 (2011) 6410, http://dx.doi.org/10.1128/JB.06181-11.
[4] H. Tae, S. Shallom, R. Settlage, G.N. Hawkins, L.G. Adams, H.R. Garner, Complete genome sequence of *Brucella suis* VBI22, isolated from bovine milk, J. Bacteriol. 194 (2012) 910, http://dx.doi.org/10.1128/JB.06556-11.
[5] S.M. Halling, B.D. Peterson-burch, J. Betsy, R.L. Zuerner, Z. Qing, L. Li, et al., Completion of the Genome Sequence of *Brucella abortus* and Comparison to the Highly Similar Genomes of *Brucella melitensis* and *Brucella suis* Completion of the Genome Sequence of *Brucella abortus* and Comparison to the Highly Similar Genomes of *Brucella meli*, 2005. http://dx.doi.org/10.1128/JB.187.8.2715.
[6] S. Audic, M. Lescot, J.-M. Claverie, H.C. Scholz, *Brucella microti*: the genome sequence of an emerging pathogen, BMC Genomics 10 (2009) 352, http://dx.doi.org/10.1186/1471-2164-10-352.
[7] R.M. Tsolis, R. Seshadri, R.L. Santos, F.J. Sangari, J.M.G. Lobo, M.F. de Jong, et al., Genome degradation in *Brucella ovis* corresponds with narrowing of its host range and tissue tropism, PLoS One 4 (2009) e5519, http://dx.doi.org/10.1371/journal.pone.0005519.
[8] A.R. Wattam, K.P. Williams, E.E. Snyder, N.F. Almeida Jr, M. Shukla, A.W. Dickerman, et al., Analysis of ten *Brucella* genomes reveals evidence for horizontal gene transfer despite a preferred intracellular lifestyle, J. Bacteriol. 191 (2009) 3569–3579, http://dx.doi.org/10.1128/JB.01767-08.
[9] S. Audic, M. Lescot, J.-M. Claverie, A. Cloeckaert, M.S. Zygmunt, The genome sequence of *Brucella pinnipedialis* B2/94 sheds light on the evolutionary history of the genus *Brucella*, BMC Evol. Biol. 11 (2011) 200, http://dx.doi.org/10.1186/1471-2148-11-200.
[10] M. Wanke, Canine brucellosis, Anim. Reprod. Sci. 82–83 (2004) 195–207, http://dx.doi.org/10.1016/j.anireprosci.2004.05.005.
[11] M. Gyuranecz, L. Szeredi, Z. Rónai, B. Dénes, L. Dencso, Á. Dán, et al., Detection of *Brucella canis*-induced reproductive diseases in a kennel, J. Vet. Diagn. Investig. 23 (2011) 143–147, http://dx.doi.org/10.1177/104063871102300127.
[12] M. Corrente, D. Franchini, N. Decaro, G. Greco, M. D'Abramo, M.F. Greco, et al., Detection of *Brucella canis* in a dog in Italy, New Microbiol. 33 (2010) 337–341 (http://www.ncbi.nlm.nih.gov/pubmed/21213592).
[13] J.C. Wallach, G.H. Giambartolomei, P.C. Baldi, C. a Fossati, Human infection with M− strain of *Brucella canis*, Emerg. Infect. Dis. 10 (2004) 146–148, http://dx.doi.org/10.3201/eid1001.020622.
[14] N. Lucero, G. Escobar, S. Ayala, N. Jacob, Diagnosis of human brucellosis caused by *Brucella canis*, J. Med. Microbiol. 54 (2005) 457–461, http://dx.doi.org/10.1099/jmm. 0.45927-0.
[15] N.E. Lucero, P. Maldonado, S. Kaufman, G. Escobar, E. Boeri, N.R. Jacob, *Brucella canis* causing infection in an HIV-infected patient, Vector Borne Zoonotic Dis. 10 (2010) 527–529.
[16] S. Marzetti, C. Carranza, M. Roncallo, G.I. Escobar, N.E. Lucero, Recent trends in human *Brucella canis* infection, Comp. Immunol. Microbiol. Infect. Dis. 36 (2013) 55–61, http://dx.doi.org/10.1016/j.cimid.2012.09.002.
[17] V. Manias, A. Nagel, A. Mollerach, M. Mendosa, H. Freyre, A. Gómez, et al., Endocarditis por *Brucella canis*: primer caso documentado en un paciente adulto en Argentina, Rev. Argent. Microbiol. 45 (2013) 50–53.
[18] W.S. Krueger, N.E. Lucero, A. Brower, G.L. Heil, G.C. Gray, Evidence for unapparent *Brucella canis* infections among adults with occupational exposure to dogs, Zoonoses Public Health (2014) 1–10, http://dx.doi.org/10.1111/zph.12102.
[19] M. Olivera-Angel, P. Ristow, A.I. Ko, C. Di-lorenzo, Serological trail of *Brucella* infection in an urban slum population in Brazil, J. Infect. Dev. Ctries. 6 (2012) 675–679.
[20] J.T. Foster, S.M. Beckstrom-Sternberg, T. Pearson, J.S. Beckstrom-Sternberg, P.S.G. Chain, F.F. Roberto, et al., Whole-genome-based phylogeny and divergence of the genus *Brucella*, J. Bacteriol. 191 (2009) 2864–2870, http://dx.doi.org/10.1128/JB.01581-08.
[21] G. Gao, J. Li, T. Li, Z. Zhang, L. Wang, X. Yuan, et al., Complete genome sequence of *Brucella canis* strain 118, a strain isolated from canine, J. Bacteriol. 194 (2012) 6680, http://dx.doi.org/10.1128/JB.01853-12.
[22] J.-S. Kim, W. Jeong, H.-Y. Jeoung, J.-Y. Song, H. Kim, J.-H. Beak, et al., Complete genome sequence of *Brucella canis* strain HSK A52141, isolated from the blood of an infected dog, J. Bacteriol. 194 (2012) 5134, http://dx.doi.org/10.1128/JB.01125-12.

[23] Y. Wang, Y. Ke, Q. Zhen, X. Yuan, J. Xu, Y. Qiu, et al., Complete genome sequence of *Brucella canis* BCB018, a strain isolated from a human patient, J. Bacteriol. 194 (2012) 6697–6698, http://dx.doi.org/10.1128/JB.01811-12.
[24] O.R. Crasta, O. Folkerts, Z. Fei, S.P. Mane, C. Evans, S. Martino-Catt, et al., Genome sequence of *Brucella abortus* vaccine strain S19 compared to virulent strains yields candidate virulence genes, PLoS One 3 (2008) e2193, http://dx.doi.org/10.1371/journal.pone.0002193.
[25] J. Ding, Y. Pan, H. Jiang, J. Cheng, T. Liu, N. Qin, et al., Whole genome sequences of four *Brucella* strains, J. Bacteriol. 193 (2011) 3674–3675, http://dx.doi.org/10.1128/JB.05155-11.
[26] L.F. Ortiz, C. Muskus, M.M. Sanchez, M. Olivera, Identification of *Brucella canis* Group 2 in Colombian kennels, Rev. Colomb. Cienc. Pec. 25 (2012) 615–619.
[27] S. Babraham, B. Andrews, FastQC a quality control tool for high throughput sequence data, http://www.bioinformatics.babraham.ac.uk/projects/fastqc/2009.
[28] S. Assefa, T.M. Keane, T.D. Otto, C. Newbold, M. Berriman, ABACAS: Algorithm-based Automatic Contiguation of Assembled Sequences, 25 2009, pp. 1968–1969, http://dx.doi.org/10.1093/bioinformatics/btp347.
[29] M.M. Sánchez-Jiménez, J.P. Isaza, J.F. Alzate, M. Olivera-Angel, Data in Brief. Multiple Variable Regions Founded After Comparison of *Brucella canis* Genomes, 2015. (submitted for publication).
[30] T.D. Otto, G.P. Dillon, W.S. Degrave, M. Berriman, RATT: rapid annotation transfer tool, Nucleic Acids Res. 39 (2011) e57, http://dx.doi.org/10.1093/nar/gkq1268.
[31] M.I. Krzywinski, J.E. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, et al., Circos: an information aesthetic for comparative genomics, Genome Res. 19 (2009) 1639–1645, http://dx.doi.org/10.1101/gr.092759.109.
[32] P.S.G. Chain, D.J. Comerci, M.E. Tolmasky, F.W. Larimer, S.A. Malfatti, L.M. Vergez, et al., Whole-Genome Analyses of Speciation Events in Pathogenic *Brucellae* Whole-genome Analyses of Speciation Events in Pathogenic *Brucellae*, 2005. http://dx.doi.org/10.1128/IAI.73.12.8353.
[33] J.T. Foster, R.T. Okinaka, R. Svensson, K. Shaw, B.K. De, R. Robison, et al., Real-time PCR assays of single-nucleotide polymorphisms defining the major *Brucella* clades, J. Clin. Microbiol. 46 (2008) 296–301, http://dx.doi.org/10.1128/JCM.01496-07.
[34] V.L. Atluri, M.N. Xavier, M.F. De Jong, A.B. Den Hartigh, Interactions of the human pathogenic *Brucella* species with their hosts, Annu. Rev. Microbiol. 65 (2011) 523–541, http://dx.doi.org/10.1146/annurev-micro-090110-102905.
[35] K.M. Kim, K.-W. Kim, S. Sung, H. Kim, A genome-wide identification of genes potentially associated with host specificity of *Brucella* species, J. Microbiol. 49 (2011) 768–775, http://dx.doi.org/10.1007/s12275-011-1084-3.