Journal of Animal and Veterinary Advances 13 (9): 582-588, 2014
ISSN: 1680-5593

Complete Genome Sequence of *B. abortus* RB51 and Genomic Comparison with *B. abortus* 2308

[1]Wolong Ma, [2]Fei Guo, [1]Antao Wang, [1]Zhanhui Song and [1]Chuangfu Chen
[1]College of Animal Science and Technology, [2]College of Medicine,
Shihezi University, 832003 Shihezi, China

Abstract: A rifampin-resistant mutant of *Brucella abortus*, designated as RB51 was derived by repeated passage of strain 2308 on Trypticase soy supplemented with 1.5% agar and varying concentrations rifampin or penicillin. In this study, 325 Mb data was produced for RB51 using Illumina HiSeq 2000 sequencing platform. Based on the assemble result of RB51, researchers obtained that 36 scaffolds with GC content of 57.25% which contains 150 contigs and has genome size was about 3.26 Mb. Detail bioinformatics annotation of RB51 revealed that the genome contained 3,283 genes, the total length of genes was 2,797,347 bp which makes up 85.91% of genome. Comparative genomics analysis found that there was a tandem repeat sequence of 339 nucleotide in YP_413569.1 protein. In addition, RB51 encoded a glycosyltransferase wboA in scaffold12, an enzyme essential in the biosynthesis of O antigen which has been disrupted by IS711. In comparison with the reference genome *B. abortus* 2308, researchers found that RB51 contained 135 SNPs and 5 InDels among which 1 SNP and 2 InDels may be associated with virulent related to lipopolysaccharide.

Key words: *B. abortus* RB51, genome sequence, functional annotation, *wboA* gene, scaffold

INTRODUCTION

A rifampin-resistant mutant of *Brucella abortus*, designated RB51 was derived by repeated passage of strain 2308 on Trypticase soy supplemented with 1.5% agar and varying concentrations rifampin or penicillin. The RB51 colonies absorbed crystal violet and RB51 cell suspensions autoagglutinated indicating a rough type colonial morphology for this strain. No O-chain component was detected in Lipopolysaccharide (LPS) extracted from RB51 on SDS-PAGE gels stained with silver (Ciuchini *et al*., 2002). Western blot analysis with the monoclonal antibody BRU 38 which is specific for the perosamine homopolymer O-chain of smooth Brucella LPS indicated that the LPS of RB51 is highly deficient in O-chain when compared with the parenteral smooth strain 2308 or rough strain 45/20. Biochemically, RB51 resembles parental strain 2308 in its ability to utilize erythritol. Intraperitoneal inoculation of RB51 into mice results in a splenic colonization which is cleared within 4 weeks post infection (Vemulapalli *et al*., 1999). RB51 does not revert to smooth colony morphology upon passage *in vivo* (mice) or *in vitro*. Mice infected with RB51 produce antibodies against *B. abortus* antigens including class 2 and 3 outer membrane proteins but not against the O-chain (De Bagues *et al*., 1994). Furthermore, rabbits, goats and cattle hyperimmunized with sonicates of RB51 develop antibodies to *B. abortus* cellular antigens but do not develop antibodies specific for the O-chain. Immunization of mice with 1×10 (8) viable RB51 organisms confers significant protection against challenge with virulent *B. abortus* strain 2308 (Cloeckaert *et al*., 2002).

In this study for the first time, researchers determined the full genome sequence of RB51 using Illumina HiSeq 2000 sequencing platform. By comparing the strain RB51 with *B. abortus* 2308, researchers identified all chromosomal components specific to each other strain as well as those conserved in both strains. These results provided a broad array of whole genome-level information not only for obtaining a complete set of genes potentially related to the virulence of Brucellus but also for understanding the evolution of Brucella (Schurig *et al*., 1991; CDC, 1998).

MATERIALS AND METHODS

Bacterial strain: *B. abortus* strain RB51 was from the culture collection and grown either in Tryptic Soy Broth (TSB) or on Tryptic Soy Agar (TSA) plates.

Sequencing and assembly: Genomic DNA is extracted and fragmented randomly and then required length DNA fragments are retained by electrophoresis. And after this researchers ligate adapters to DNA fragments then

Corresponding Author: Chuangfu Chen, College of Animal Science and Technology, Shihezi University, 832003 Shihezi, China

conduct cluster preparation sequencing finally. After the DNA sample (s) was (were) delivered, researchers did a sample quality test first. Then, researchers used this (those) qualified DNA sample (s) to construct BS library: Purified DNA sample such as genomic DNA, Bacterial artificial chromosome or long-length PCR productions is sheared into smaller fragments with a desired size by Covaris S/E210 or Bioruptor firstly. Then, the overhangs resulting from fragmentation are converted into blunt ends by using T4 DNA polymerase, Klenow Fragment and T4 Polynucleotide Kinase. After adding an A base to the 3' end of the blunt phosphorylated DNA fragments, adapters are ligated to the ends of the DNA fragments. The desired fragments can be purified though gel-electrophoresis then selectively enriched and amplified by PCR. The index tag could be introduced into the adapter at the PCR stage as appropriate and researchers did a library quality test. At last, the qualified BS library would be used for sequencing.

ORF prediction and functional annotation: Bioinformatics analysis will be proceeding after data filtering.

- Data filtering: The raw data is filtered and generate clean data
- Assembly: Use SOAPdenovo Software to assemble the reads after filtering
- Genomic component analysis:
  - Analysis on repeat sequences which includes tandem repeats sequence, minisatellite DNA and microsatellite DNA
  - CRISPER prediction
  - Non-coding RNA prediction. Non-coding RNA includes rRNA, tRNA and sRNA
  - Genomic Islands (GIs) prediction
  - Prophage prediction
  - Gene prediction
- Analysis on gene function:
  - Gene annotation: The predicted ORFs are annotated by GO, KEGG, Swiss-Prot (default), NR database and COG database, respectively
  - Analysis on animal pathogens including: T3SS effector protein, PHI, VFDB, ARDB database annotation
  - Pathogenicity analysis on plant pathogens including: T3SS effector protein, CAZy, PHI
- Comparative genomic:
  - Structural variation (Synteny)
  - Core-pan gene analysis
  - Evolution analysis: construction of phylogenetic tree and ka/ks analysis
  - Gene family analysis
- Report accomplishment

RESULTS AND DISCUSSION

General features: Using Illumina HiSeq 2000 sequencing platform, 325 Mb data was produced for RB51. Based on the assemble result of RB51, researchers found that the genome size was 3.26 Mb, GC content was 57.25%, the number of scaffold was 36 and the number of contig was 150 (Table 1). The genome information of sequenced bacterium was obtained using the methods of gene prediction, repeat region prediction and non-coding RNA prediction. From the genome analysis results of RB51, researchers found that the genome contained 3,283 genes, the total length of genes was 2,797,347 bp which makes up 85.91% of genome, the number of tandem repeat sequence was 80, the total length of tandem repeat sequence was 6,564 bp which makes up 0.2016% of genome, the number of minisatellite DNA was 54, the number of microsatellite DNA was 12, the number of tRNA was 49 and the number of rRNA was 3. The detail information was shown in Table 1.

Functional annotation: To gain insights into the functions of genes set of RB51, Gene Ontology (GO) enrichment analysis was performed (Fig. 1). The GO analysis revealed several ontology terms which were enriched in three ontology domains, respectively. In the terms of biological process, several genes were enriched in metabolism (GO:0008152) and other cellular process (GO:0009987). In the terms of molecular function most genes of RB51 were involved in catalytic activity (GO:0003824) binding (GO:0005488) and transporter activity (GO:0005215) which were likely to be related to the virulence activity of Brucellus. In the term of cell component, a large percent of genes were identified in macromolecular complex (GO: 0032991) and organelle (GO: 0043226) of cell part (GO: 0043226).

Based on COG annotation, researchers found that there were 381 genes involved in amino acid transport and metabolism, 204 genes involved in carbohydrate transport and metabolism and 203 genes involved in transcription (Table 2).

Comparative genomics analysis of RB51 with *B. abortus* 2308: Comparative genomics analysis found that there

Table 1: General genomi feature of *B. abortus* RB51-1

Feature/property	*B. abortus* RB51-1
Sequence length (bp)	3,256,158
G+C content (%)	57.25
G+C content of protein coding region (%)	58.42
G+C content of non-coding region (%)	50.11
Coding content (%)	85.91
Predicted protein Coding Sequences (CDSs)	3,283
rRNAs	3
tRNAs	49
Small *RNA* genes	3
Average gene length (bp)	852

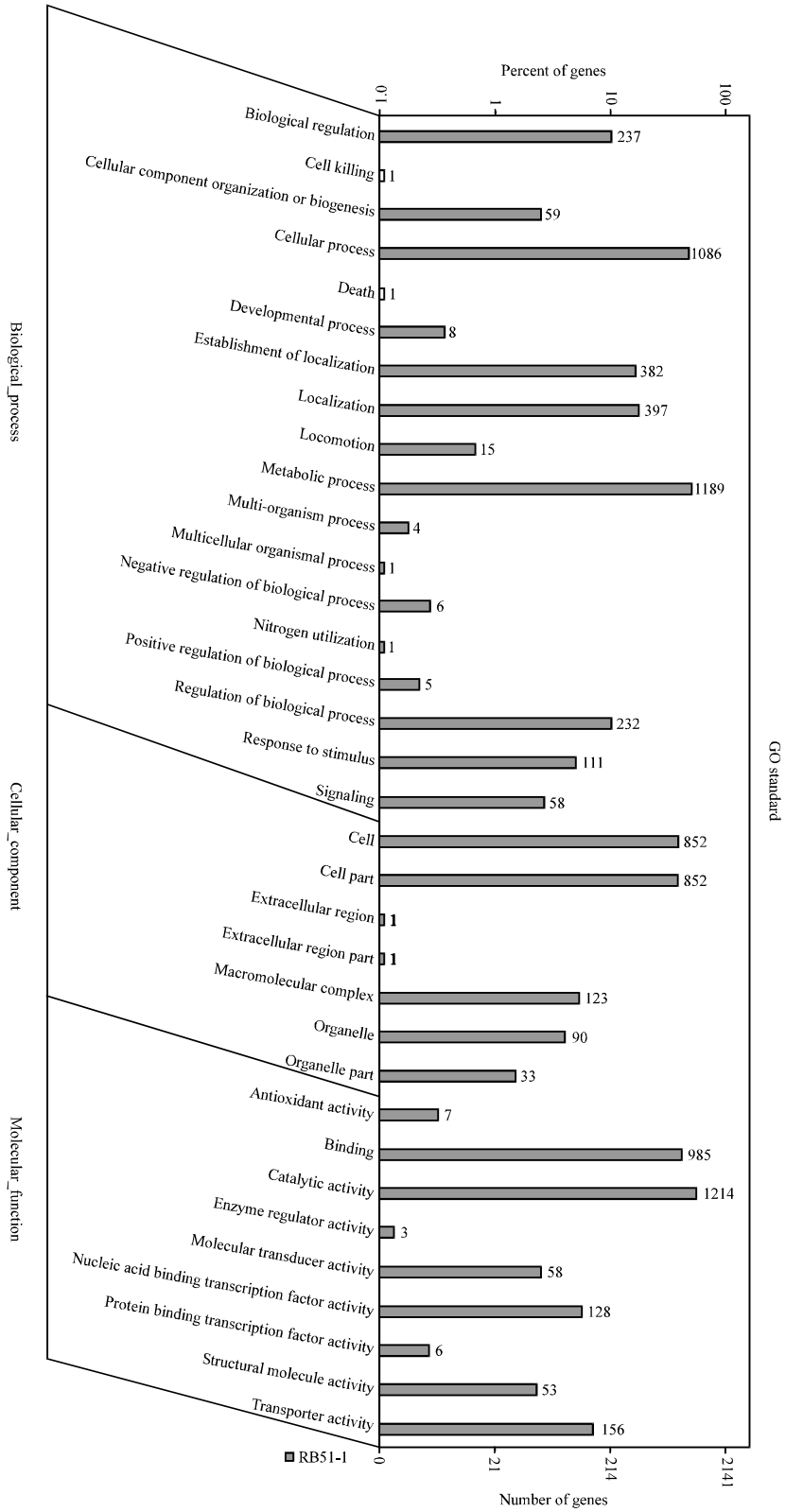

Fig. 1: Gene ontology annotation of the ORFs obtained from the assembled RB51

was a tandem repeat sequence of 339 nucleotide in YP_413569.1 protein (Fig. 2). Six copies of repeated sequence occur in the full-size ORFs of strains 2308 and 9-941. In S19, only one copy was left. In RB51, the tandem repeat also was deleted only three copies were left. The length of repeat copies, 339 bp is consistent with the report. The start and stop positions of report are 805-2,499 and in RB51 are 724-2, 758 predicted by Tandem repeat finder (Fig. 2). The presence or similar deletion in other virulent species so researchers could not confirm whether the deletion associated virulence or not. The 6 repeats alignment result was shown in Fig. 3 and researchers found that the similarity between the first repeat and other repeats is smaller.

Table 2: Number of gene pre COG group in the of *B. abortus* RB51

Function category (NCBI COGs)	Class ID	No. of RFs
Information storage and processing		
Chromatin structure and dynamics	B	1
Translation, ribosomal structure and biogenesis	J	162
Transcription	K	203
Replication, recombination and repair	L	109
Celular processing and signaling		
Cell cycle control, cell division, chromosome partitioning	D	27
Cell wall/membrane/envelope biogenesis	M	158
Cell motility	N	31
Posttranslational modification, protein turnover, chaperones	O	122
Signal transduction mechanisms	T	84
Intracellular trafficking, secretion and vesicular transport	U	50
Defense mechanisms	Y	40
Extracellular structures	W	1
Metabolism		
Energy production and conversion	C	180
Amino acid transport and metabolism	E	381
Nucleotide transport and metabolism	F	69
Carbohydrate transport and metabolism	G	204
Coenzyme transport and metabolism	H	129
Lipid transport and metabolism	I	112
Inorganic ion transport and metabolism	P	185
Secondary metabolites biosynthesis, transport and catabolism	Q	72
Poorly characterized		
General function prediction only	R	363
Function unknown	S	254
No similarity to COGs with an e-value <1e-5	-	346

Brucella abortus Strain RB51 (SRB51) is a Lipopolysaccha-ride (LPS) O-antigen-deficient mutant of the virulent Strain 2308 (S2308) of *B. abortus*. SRB51 is being evaluated as an alternative to *B. abortus* Strain 19 (S19) as a vaccine for preventing brucellosis and abortions in cattle because unlike S19, it does not induce antibodies to the Brucella LPS O antigens that are detected by serodiagnostic tests for brucellosis. Consequently, the replacement of S19 with SRB51 for use as a vaccine may facilitate the identification of cattle with brucellosis and their removal from vaccinated herds.

Analysis of *wboA* gene: *wboA* gene of *B. abortus* encodes glycosyltransferase, an enzyme essential in the biosynthesis of O antigen. Disruption of the *wboA* gene in smooth strain resulted in conversion to a rough phenotype (4). This gene was located on scaffold 12 in RB51 and also disrupted by IS711 from 709 bp and this fragment is absence in *B. abortus* 2308, *B. abortus* 9-941 and *B. abortus* S19 (Fig. 4). The length of IS711 is 844 bp. Alignment result between RB51 IS711 and *Brucella ovis* IS711 is shown in Fig. 5 (Van Metre *et al*., 1999; Michaux-Charachon *et al*., 1997).

wboA gene of *B. abortus* encodes glycosyltransferase an enzyme essential in the biosynthesis of O antigen. Disruption of the *wboA* gene in smooth strain resulted in conversion to a rough phenotype (Rutherford *et al*., 2000). This gene was located on scaffold 12 in RB51 and also disrupted by IS711 from 709 bp and this fragment is absence in *B. abortus* 2308, *B. abortus* 9-941 and *B. abortus* S19 (Fig. 4). The length of IS711 is 844 bp (Sanchez *et al*., 2001). Alignment result between RB51 IS711 and *Brucella ovis* IS711 is shown in Fig. 5.

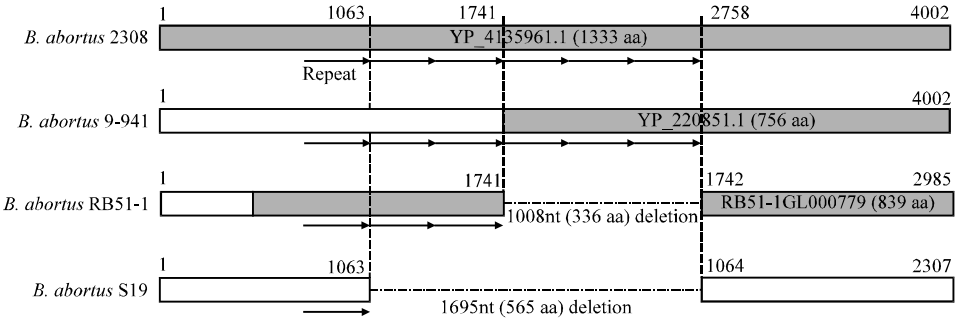

Fig. 2: Diagramatic representation of alignment of YP_413569.1

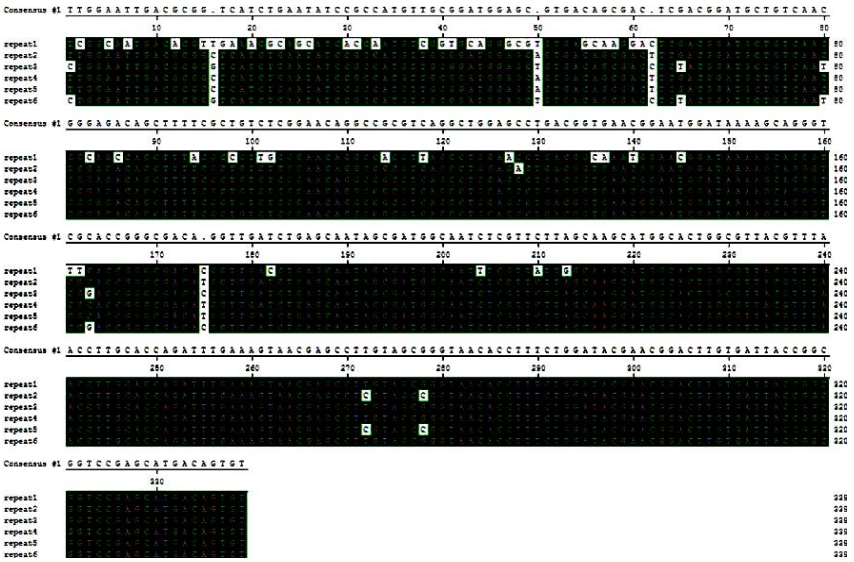

Fig. 3: Multiple sequences alignments of six repeats using the ClustalX program

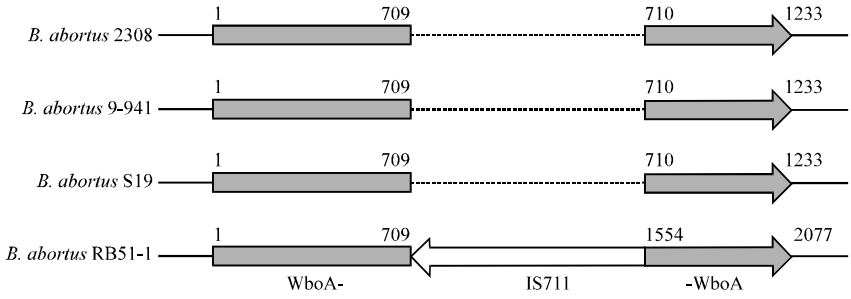

Fig. 4: Diagrammatic representation of alignment of WboA

Fig. 5: Pair-wise sequence alignment result of IS711 using the BLAST program

Table 3: SNPs may be associated with virulent

SNP pos	Ref base<-> sample base	Ref code<-> sample code	Mutation type	Gene ID	Function	VFDB identity	VFDB anno
NC_007618.1_529447	G<->A	GAG<->AAG	Nonsyn	BAB1_0534	Polysaccharide biosynthesis protein CapD	43.60	Lipopolysaccharide biosynthesis protein

Table 4: InDels may be associated with virulent

InDelpos	InDel type	length	Pos type	Gene ID	Function	VFDB identity	VFDB anno
Scaffold 2	Insertion	93	CDS	BAB2_0582	Glycerol-3-phosphate transporter	40.85	Iron (III) ABC transporter, ATP-binding protein
Scaffold 13	Insertion	8	Intergenic, 5' 74 bp	BAB1_0055	ATP-binding subunit Phosphoglucomutase	99.82	Phosphoglucomutase

Table 5: Genetic variation detected in *B. abortus* RB51 genome

Genetic variation	*B. abortus* 2308 chr I	*B. abortus* 2308 chr II	Total
SNPs			
Intergenic	9	11	20
Synonymous	9	7	16
Nonsynonymous	29	12	41
Nonsense	47	30	77
INDELs			
Insertion	2	3	5
Deletion	2	1	3
Total	4	4	8

SNPs and InDels: Comparing with reference sequence *B. abortus* 2308 (Tsolis, 2002), researchers found that RB51 contains 135 SNPs including 26 synonymous mutations and 89 non-synonymous mutations. The results are shown in Table 2. Comparing with reference sequence, researchers found 5 inDels in RB51 contains 3 insertion mutations, 2 deletion mutations. The 1 SNP and 2 InDels may be associated with virulent (Table 3 and 4). The genes contained mutations could alignmemt with the virulence facters database. Lipopolysaccharide (LPS) is an important component of the outer membrane in gram-negative bacteria. LPS has three domains: the lipid A, the core oligosaccharide and the O antige or O side chain. The absence of side chain from LPS determines the rough phenotype. Generally these mutants are less virulent than the wild type. The LPS may play a role in extracellular survival in the animal (1). Wild type will become rough type if delete a part of PLS and the mutation strain induced protection levels comparable to those induced by S19 (2). There are 8 bp insert in the 5' upstream 74 bp. Mutation experiment needs to be tested whether the insertion regulates the LPS transcription. The results are shown in Table 3 and 5.

CONCLUSION

The complete genome sequence of *B. abortus* RB51 obtained in this study can provide considerable insights into genetic, evolution, functional investigation of its genome and biological roles.

ACKNOWLEDGEMENTS

This research was supported by grants from the National Basic Research Program of China (973 Program) (2010CB530203) and National Twelfth Five-Year Plan for Scie-Nc and Technology Support Program (2013BAI05B05) and the International Cooperation in Science and Technology (2013DFA32380). Wolong Ma, Fei Guo, Antao Wang and Zhanhui Song contributed equally to the research.

REFERENCES

CDC, 1998. Human exposure to *Brucella abortus* strain RB51-Kansas, 1997. Morbidity Mortal. Weekly Rep., 47: 172-175.

Ciuchini, F., R. Adone and P. Pasquali, 2002. Coombs antiglobulin test using *Brucella abortus* 99 as antigen to detect incomplete antibodies induced by *B. abortus* RB51 vaccine in cattle. Clin. Diagn. Lab. Immunol., 9: 1398-1399.

Cloeckaert, A., M.S. Zygmunt and L.A. Guilloteau, 2002. *Brucella abortus* vaccine strain RB51 produces low levels of M-like O-antigen. Vaccine, 20: 1820-1822.

De Bagues, M.J., P.H. Elzer, S.M. Jones, J.M. Blasco, F.M. Enright, G.G. Schurig and A.J. Winter, 1994. Vaccination with *Brucella abortus* rough mutant RB51 protects BALB/c mice against virulent strains of *Brucella abortus*, *Brucella melitensis* and *Brucella ovis*. Infect. Immun., 62: 4990-4996.

Michaux-Charachon, S., G. Bourg, E. Jumas-Bilak, P. Guigue-Talet, A. Allardet-Servent, D. O'Callaghan and M. Ramuz, 1997. Genome structure and phylogeny in the genus *Brucella*. J. Bacteriol., 179: 3244-3249.

Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M.A. Rajandream and B. Barrell, 2000. Artemis: Sequence visualization and annotation. Bioinformatics, 16: 944-945.

Sanchez, D.O., R.O. Zandomeni, S. Cravero, R.E. Verdun and E. Pierrou *et al*., 2001. Gene discovery through genomic sequencing of *Brucella abortus*. Infect. Immun., 69: 865-868.

Schurig, G.G., R.M. Roop, T. Bagchi, S. Boyle, D. Buhrman and N. Sriranganathan, 1991. Biological properties of RB51; A stable rough strain of *Brucella abortus*. Vet. Microbiol., 28: 171-188.

Tsolis, R.M., 2002. Comparative genome analysis of the α-proteobacteria: Relationships between plant and animal pathogens and host specificity. Proc. Natl. Acad. Sci. USA, 99: 12503-12505.

Van Metre, D.C., A. Kennedy, C. Olsen, R. Hansen and R. Ewalt, 1999. Brucellosis induced by RB51 vaccine in a pregnant heifer. J. Am. Vet. Med. Assoc., 215: 1491-1493, 1449.

Vemulapalli, R., J.R. McQuiston, G.G. Schurig, N. Sriranganathan, S.M. Halling and S.M. Boyle, 1999. Identification of an IS711 element interrupting the *wboA* gene of *Brucella abortus* vaccine strain RB51 and a PCR assay to distinguish strain RB51 from other *Brucella* species and strains. Clin. Diagn. Lab. Immunol., 6: 760-764.