



LAWRENCE  
LIVERMORE  
NATIONAL  
LABORATORY

# Whole-genome analyses of the speciation events in the pathogenic *Brucellae*

Patrick Chain, Diego Comerci, Marcelo Tolmasky,  
Frank Larimer, Stephanie Malfatti, Lisa Vergez, Fernan  
Aguero, Miriam Land, Rodolfo Ugalde, Emilio Garcia

July 15, 2005

Infection and Immunity

## **Disclaimer**

---

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

# Whole-genome analyses of the speciation events in the pathogenic

## *Brucellae*

Patrick S. G. Chain<sup>1</sup>, Diego J. Comerci<sup>2</sup>, Marcelo E. Tolmasky<sup>3</sup>, Frank W. Larimer<sup>4</sup>, Stephanie A. Malfatti<sup>1</sup>, Lisa M. Vergez<sup>1</sup>, Fernan Aguero<sup>2</sup>, Miriam L. Land<sup>4</sup>, Rodolfo A. Ugalde<sup>2</sup>, and Emilio Garcia<sup>1</sup>.

<sup>1</sup> Biosciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA 94551

<sup>2</sup> Instituto de Investigaciones Biotecnológicas, Universidad Nacional de General San Martín (IIB-INTECH-CONICET), Av. Gral. Paz 5445, P.O. Box 30, 1650, San Martín, Buenos Aires, Argentina

<sup>3</sup> Department of Biological Science, California State University, Fullerton, CA 92834-6850

<sup>4</sup> Life Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831

**Address of corresponding author:**

7000 East Ave. L-441, Livermore, California, USA 94551.

**Email:** [garcia12@llnl.gov](mailto:garcia12@llnl.gov)

**Abbreviations:** ChrI, chromosome I; ChrII, chromosome II; PC,

phosphatidylcholine; PE, phosphatidyletanolamine; Pmta, phosphatidyl-N-

methylethanolamine N-methyltransferase; Pcs, phosphatidylcholine synthase; ECF,

extracytoplasmic function sigma factor.

## ABSTRACT

1  
2 Despite their high DNA identity and a proposal to group classical *Brucella* species as  
3 biovars of *B. melitensis*, the commonly recognized *Brucella* species can be  
4 distinguished by distinct biochemical and fatty acid characters as well as by a marked  
5 host range (e.g. *B. suis* for swine, *B. melitensis* for sheep and goats, *B. abortus* for  
6 cattle). Here we present the genome of *B. abortus* 2308, the virulent prototype biovar  
7 1 strain, and its comparison to the two other human pathogenic *Brucellae* species and  
8 to the *B. abortus* field isolate 9-941. The global distribution of pseudogenes, deletions  
9 and insertions support previous indications that *B. abortus* and *B. melitensis* share a  
10 common ancestor that diverged from *B. suis*. With the exception of a dozen genes,  
11 the genetic complement of both *B. abortus* strains is identical, whereas the three  
12 species differ in gene content and pseudogenes. The pattern of species-specific gene  
13 inactivations affecting transcriptional regulators and outer membrane proteins suggest  
14 that these inactivations may play an important role in the establishment of host-  
15 specificity and may have been a primary driver of speciation in the *Brucellae*. Despite  
16 being non-motile, the *Brucellae* contain flagellum gene clusters and display species-  
17 specific flagellar gene inactivations, which lead to the putative generation of different  
18 versions of flagellum-derived structures, and may contribute to differences in host-  
19 specificity and virulence. Metabolic changes such as the lack of complete metabolic  
20 pathways for the synthesis of numerous compounds (e.g. glycogen, biotin, NAD, and  
21 choline) are consistent with adaptation of *Brucellae* to an intracellular lifestyle.

22

## INTRODUCTION

Brucellosis is a major emerging infectious disease affecting animals and humans (11). Several *Brucella* species, such as *B. abortus*, *B. melitensis*, and *B. suis*, have been isolated from a variety of animals. Of these, *B. abortus*, the causative agent of bovine brucellosis is the most widespread (11). All three *Brucella* species cause a severe human disease characterized in its acute phase by undulant fever and in its chronic phase by localization of the pathogen and damage of different organs (11, 37). If localized in the brain or the heart, this can result in fatal meningitis or fatal endocarditis, respectively. *Brucella* infection is treated with a combination of antibiotics, however in its chronic phase, eradication is difficult since *Brucella* spp. are intracellular pathogens, which puts them out of reach of humoral immunity and several antibiotics (25, 37). The lack of a safe and efficacious human vaccine underscores the importance of understanding the biology of brucellosis to develop human vaccines and effective therapeutic agents. Furthermore, all three *Brucella* species are listed as potential bio-weapons by the Centers for Disease Control and Prevention (22, 26). This due to the highly infectious nature of all three species, the fact that they can be readily aerosolized and that an outbreak would be difficult to detect because the initial symptoms are easily confused with flu.

*Brucella* spp. belong to a monophyletic branch of the alpha 2 subgroup of proteobacteria whose members share the ability to engage in intimate or sometimes intracellular associations with eukaryotic cells (45). However, little is known regarding the biological basis of *Brucella* spp. host-specificity, or about the mechanisms involved in the intracellular multiplication and persistence of members of this group. The availability of the *B. abortus* strain 2308 sequence reported here, together with the sequences of *B. suis*, *B. melitensis* and *B. abortus* strain 9-941 (12,

1 18, 35) enabled us to perform a comprehensive examination and comparison of the  
2 gene composition, mutations, structural arrangement, and other characteristics of  
3 these genomes. Detailed comparisons have confirmed the *B. abortus* deletions  
4 identified using *B. melitensis*-derived microarrays (36), and have uncovered a large  
5 number of additional differences in *B. abortus*, as well as in *B. melitensis* and *B. suis*.  
6 Overall, our analyses confirm the striking similarities that exist between the three  
7 species and reveal a number of important, though subtle features that may be  
8 important in host specificity and the adaptation of these organisms to an intracellular  
9 life-style.

## MATERIALS AND METHODS

**Genome Sequencing and Annotation.** Genomic DNA was isolated from *B. abortus* strain 2308, a standard laboratory Biovar I strain, virulent for humans, cattle and certain other domestic animals (38). Whole genome sequencing was performed by a shotgun method as previously described (6). All sequencing reactions were performed using BigDye® Terminator v3.0 cycle sequencing chemistry and resolved using 3730xl DNA Analyzers (Applied Biosystems, Foster City, CA). The whole genome sequence of *B. abortus* 2308 was obtained from 40,000 reads generating roughly 7.0 fold redundancy. Sequence finishing and polishing, assessment of final assembly quality, as well as gene modeling and annotation was performed as previously described (6).

**Genome Comparison and Analysis.** Whole chromosome alignments of *B. abortus* 2308 with *B. melitensis* strain 16M and *B. suis* strain 1330 were visualized using the ACT program (which can be found at <http://www.sanger.ac.uk/Software/ACT/>). Due to the similarity and colinearity of the genomes, orthologs can be found by gene order and were confirmed by identifying best reciprocal BLAST hits. Multiple alignments were performed using ClustalW. Pseudogenes were identified by comparing the *Brucella* orthologs to each other and to those of other species. In this fashion, premature stops (nonsense mutations), frameshifts and deletion events were identified. Inframe deletion events were not included in the list of pseudogenes unless the deletion event perturbed a non-repetitive functional domain, as identified via annotation.

**Genome Accessions.** The sequences reported in this paper have been deposited in the EMBL Nucleotide Sequence Database; accession numbers AM040264 (Chromosome I) and AM040265 (Chromosome II).

## RESULTS AND DISCUSSION

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25

**General Features of the *B. abortus* 2308 Genome.** The genome of *B. abortus* strain 2308 consists of two circular chromosomes ChrI (2,121,359 bp) and ChrII (1,156,950bp) coding for 2280 and 1214 annotated open reading frames, respectively. The salient genome features and its comparison with *B. suis*, *B. melitensis* and *B. abortus* 9-941 genomes are summarized in Fig. 1 and Table 1. Although the global characteristics of the *B. abortus* 2308 genome are quite similar to the recently published sequence of the field isolate of *B. abortus* (18) a number of strain-specific deletions and polymorphisms were identified (Supplemental material, Tables 1 and 2). The intra-species comparison is analyzed and discussed in the context of the inter-species comparisons to two previously sequenced *Brucella* genomes (12, 35). Comparative analysis has uncovered a number of species- and strain-specific genomic features and putative physiological and metabolic differences that may help explain their host-specificities and the molecular events underlying their speciation.

**Comparative Genomics.** The genomes of the three *Brucella* species display an average of >94% identity at the nucleotide level. Despite the highly conserved genome backbone shared between the three *Brucellae*, several important species-specific differences have been uncovered in this work and are described and discussed below. In addition, intraspecies comparisons between *B. abortus* strains 2308 and 9-941 revealed a small number of strain-specific deletions and polymorphisms (Supplemental material, Tables 1 and 2). Specific differences between strains 9-941 and 2308 are also discussed below in the context of cross species comparisons.

**i) Insertions/deletions.** Only *B. suis* possesses laterally-acquired species-specific genes. Located within a single 18.3 kb cluster on ChrII, this *B. suis*-specific region

1 encompasses 18 genes (BRA0362-BRA0379), and appears to have been acquired  
2 recently via a phage-mediated integration event. It does not encode any obvious  
3 virulence factor, but encodes many hypothetical proteins, a putative transcriptional  
4 regulator, and a type-IV conjugal transfer cluster of genes previously described by  
5 Paulsen *et al.* (35) and is flanked by 15 bp direct repeats with a phage integrase on  
6 one end.

7         The only other integration event observed that differentiates between *Brucella*  
8 species (*B. suis* lacks this region) has already been described in *B. melitensis* ChrI  
9 (35), and is shared with both *B. abortus* strains 2308 (BAB1\_0250- BAB1\_0279) and  
10 9-941 (BruAb1\_0245-BruAb1\_0274). This chromosomal region represents the only  
11 gene-containing region unique to *B. melitensis* and *B. abortus* absent from *B. suis*.

12         In *B. abortus*, four deletion events have led to the loss of 26 complete genes  
13 and two 5' or 3' gene segments. The two larger deletions (25.1 and 2.8 kb) have been  
14 described in other comparative genomics analyses (18, 36), while the two smaller  
15 deletions (1.2 and 0.9 kb), also observed using DNA microarray comparative  
16 hybridization (36), encode a prokaryotic signaling, diguanylate  
17 cyclase/phosphodiesterase domain (GGDEF) protein and a ThiJ/PfpI family protease.

18         *B. abortus* strain 2308 harbors 2 additional deletions (5.2 kb in ChrII and 3.7  
19 kb in ChrI) that affect an ABC periplasmic sugar-binding protein (BRA0304 in *B.*  
20 *suis*), a putative sugar binding proteins (BRA0302), a transcriptional regulator  
21 (BRA0301) and NarK (nitrite extrusion protein, BRA0300) in ChrII, and isovaleryl-  
22 CoA dehydrogenase (BR0020), acetoacetyl-CoA synthase (BR0021), and an  
23 acyltransferase (BR0023) in ChrI. These two recent strain 2308-specific deletions  
24 appear to have occurred in regions already in a state of decay, as suggested by the  
25 many gene inactivations found in these regions in both *B. abortus* strains and also *B.*

1 *melitensis* 16M. For example, other components of the sugar ABC transport system,  
2 *narK* and *narG* (found in the same operon as *narK*), and the acetoacetyl-CoA  
3 synthase gene, all appear inactivated in two or all three strains. Though the 3.7 kb  
4 region was found to be absent from the *B. abortus* S2308 isolate used in the DNA  
5 microarray study by Rajashekara et al. (36), the 5.2 kb was not identified and may  
6 represent a very recent, isolate-specific deletion.

7         In *B. melitensis*, 16 complete genes, four 5' or 3' gene segments and a tRNA-  
8 Glu are missing as the result of seven deletions (five in ChrI and two in ChrII - see  
9 Supplementary material, Table 1). Interestingly, four of the five ChrI deletions in *B.*  
10 *melitensis* represent most of the few observed homologous recombination events  
11 within the three *Brucella* species. In one case, a transcriptional regulator  
12 (BAB1\_1860) and an amino acid permease (BAB1\_1861) were deleted as the result  
13 of a recombination event between two flanking and highly similar transposase-  
14 encoding genes. In another, a homologous recombination between two nearby tRNA-  
15 Glu (tRNA-Glu1 and tRNA-Glu2 in ChrI of *B. abortus*) resulted in the *B. melitensis*  
16 loss of one tRNA and the intervening region.

17         In addition, a 2.7 kb region is missing from both *B. abortus* and *B. melitensis*.  
18 This *B. suis*-specific region is located on ChrI, encompasses BR0952-BR0954 as well  
19 as parts of the flanking genes BR0951 and BR0955 (partly deleted from the two other  
20 strains), which affects an amino acid ABC-type transporter and a glutathione-S-  
21 transferase. We conclude that the specific deletion of this amino acid transporter,  
22 likely involved in the transport of arginine, ornithine and lysine, is thus likely  
23 responsible for the observed inability of *B. melitensis* and *B. abortus* to oxidize these  
24 amino acids. This is the only deletion shared across the *Brucella* species. Together  
25 with the above finding, the common 18.3 kb insertion and shared 2.2 kb inversion

1 between *B. abortus* and *B. melitensis* (see Inversions section ii), and the unequal  
2 distribution of pseudogenes (see Gene Inactivations section iii), strongly supports the  
3 hypothesis that these two species arose from a common ancestor that diverged from  
4 the *B. suis* lineage (18, 29).

5 **ii) Inversions.** Perhaps due to the paucity of repeated elements in the *Brucellae*  
6 genomes, few rearrangements were observed between the three genomes and none of  
7 the observed inversions have occurred at repeated sequences but instead appear to be  
8 the result of non-homologous recombinations and are likely to represent stable events.  
9 This contrasts sharply with the genomes of other pathogens with closely related  
10 neighbor species, such as members of the *Yersinia* group, in which rearrangements  
11 occur primarily at insertion sequences or other repeated sequences (7, 13).

12 The genome of *B. abortus* 2308 harbors the same three inversions previously  
13 described in strain 9-941 (18), including a small 2.2 kb inversion shared with *B.*  
14 *melitensis* that disrupts the hypothetical gene BAB2\_0749 and BAB2\_0752, which  
15 encodes a glycoside hydrolase (Supplementary material, Fig. 1). *B. melitensis* harbors  
16 two additional small inversions in ChrII, a 420 bp inversion within an IS66 family  
17 element transposase (BAB2\_0684 in *B. abortus*; BMEII0713/4 in *B. melitensis*) and a  
18 204 bp inversion that disrupts a hypothetical gene (BAB2\_0441 in *B. abortus*;  
19 BMEII0494 in *B. melitensis*). *B. suis* harbors no noticeable genome rearrangements  
20 compared with the genomes of *B. abortus* and *B. melitensis*.

21 **iii) Gene inactivations.** Several small deletions in the three *Brucella* species (often  
22 intragenic) result in the generation of partial gene remnants or in the fusion of two  
23 partial genes (see Supplementary material, Table 1). A number of these events  
24 contribute to species-specific alleles, while others support the notion that *B. abortus*  
25 and *B. melitensis* are more closely related to one another. A total of 207 pseudogenes

1 were identified in the *B. abortus* 2308 genome (versus 212 in 9-941) while 152 and 82  
2 were identified in *B. melitensis* and *B. suis*, respectively (Fig. 2 and Supplementary  
3 material, Table 2). Very few gene inactivation differences were observed when  
4 comparing the two strains of *B. abortus*. From the total pool of inactivated genes, 16  
5 pseudogenes are common to all three species, 57 are shared by *B. abortus* and *B.*  
6 *melitensis* and only 6 and 5 are shared between *B. suis* and *B. abortus*, and *B. suis* and  
7 *B. melitensis* respectively. It is important to note that all of these latter shared  
8 pseudogenes (between *B. suis* and either *B. melitensis* or *B. abortus*) are the result of  
9 independent mutations and do not represent commonly inherited traits. Among the 16  
10 pseudogenes common to the three species, 9 were generated by the same event, thus  
11 strongly suggesting being acquired by the common ancestor of all *Brucellae*.  
12 Likewise, most of the 56 pseudogenes shared by *B. abortus* and *B. melitensis* share  
13 the same mutation. The remainder of the pseudogenes were species-specific,  
14 suggesting that a large number of independent mutations have either contributed to, or  
15 occurred after, their divergence from one another.

16 Analysis of the functional categories of the inactivated genes in *Brucellae*  
17 revealed that one third belongs to the transport and cell envelope category, whereas  
18 transcriptional regulation and signal transduction accounted for 10%, and energy  
19 production and conversion categories accounted for 6% of the total (Fig. 3). This  
20 process of reduction of transport and cell envelope components might be the  
21 consequence of adaptation to an intracellular lifestyle in a protected and more stable  
22 local environment/niche that provides a constant supply of nutrients, while the  
23 dissimilar accumulation of inactivated transcriptional regulators (Fig. 3) and outer  
24 membrane structures present in the *Brucella* might represent a contribution to, or be  
25 the consequence of, the process of adaptation to a specific animal host. This would

1 suggest that the virulence and host specificity differences observed among the three  
2 species may be the result of divergence driven by differential accumulation of  
3 pseudogenes within regulatory functions in the three species. Thus, the pattern of gene  
4 inactivation observed in *Brucellae* would resemble that observed in the genomes of  
5 *Bordetella bronchiseptica*, *B. parapertussis* and *B. pertussis*. In these species,  
6 differences in virulence and host adaptations were attributed to inactivation and loss  
7 of cell surface structural genes, transport functions and regulatory functions (33). A  
8 similar situation has been recently observed in *Francisella tularensis* (27) and *Y.*  
9 *pestis* (7, 34), in which gene loss may contribute to virulence differences among  
10 species within the same genus. Thus, it appears that gene inactivation rather than gene  
11 deletion or gene acquisition (via horizontal transfer) may play a more important role  
12 in the speciation and host adaptation of this pathogenic group.

13

14 **Metabolism.** Most of the metabolic pathways that characterize prototrophic species  
15 are functional in *Brucellae* (perhaps surprising for a well-adapted intracellular  
16 pathogen), with some interesting exceptions:

17 **i) Energy storage metabolites.** All three *Brucella* species have lost the genes  
18 involved in the metabolism of glycogen. It is thus interesting that in the closely  
19 related alpha-proteobacteria *Rhizobium* and *Agrobacterium*, glycogen genes are  
20 conserved and organized as a single operon (*glg*) that contains the  
21 phosphoglucosemutase (*pgm*) gene (44). The only gene of the *glg* operon conserved  
22 in *Brucella* is *pgm* which is responsible for the synthesis of UDP-glucose, ADP-  
23 glucose or any other sugar nucleotide derivative from them. In *Brucella* as in  
24 *Agrobacterium* and *Rhizobium*, *pgm* is required for effective bacterium-host  
25 interaction. *Brucella pgm* mutants contain a defective LPS (devoid of O-antigen) and

1 are avirulent even though they are able to replicate inside non-phagocytic cells. It is  
2 thus, interesting to note that the only glycogen related gene conserved in *Brucellae* is  
3 the one that is required for virulence (43).

4 A search for the presence of other genes involved in energy storage  
5 compounds revealed that all sequenced *Brucella* have also lost the ability to  
6 synthesize and degrade poly-hydroxybutyrate (PHB). It is likely that, as with other  
7 intracellular pathogens, *Brucella* has adapted to its intracellular lifestyle and no longer  
8 requires the accumulation of energy-storing molecules.

9 **ii) Sugar metabolism.** In most organisms, the conversion of  $\alpha$ -D-galactose to the  
10 more metabolically useful glucose-1-phosphate is accomplished by the action of four  
11 enzymes that constitute the Leloir pathway. In *Brucella* species the only enzyme of  
12 the Leloir pathway present is the UDP-galactose 4-epimerase (BAB1\_0734), which  
13 implies that in *Brucella*, galactose is metabolized through an oxidative pathway  
14 encoded by a gene cluster located in ChrII (BAB2\_0293-BAB2\_0296). Since  
15 *Agrobacterium* and *Rhizobium* also lack the Leloir pathway it is likely that the  
16 absence of this pathway in *Brucella* represents the loss of these enzymes in an earlier  
17 ancestor of the alpha proteobacteria.

18 A 25.1 kb region containing 23 genes, of which most are involved in sugar  
19 metabolism and biosynthesis of a hypothetical polysaccharide, is absent in *B. abortus*  
20 yet is present in both *B. melitensis* and *B. suis* (see Supplementary Material, Table 1).  
21 Conversely, a 9 kb gene encoding Cgs, the cyclic beta 1-2 glucan synthetase  
22 (BAB1\_0108), known to be involved in virulence (5, 8, 21) is strictly conserved in the  
23 three species. Thus, it is tempting to speculate that the set of genes absent from *B.*  
24 *abortus* are involved in species-specific host recognition and dispensable in this  
25 species, while the cyclic glucan synthase, a protein required for the synthesis of a

1 virulence factor that controls the intracellular trafficking (1), must be conserved in all  
2 three species. Interestingly, Cgs is also conserved in the symbiont *Rhizobium* and is  
3 required for effective nodule invasion.

4 **iii) Nicotinamide synthesis.** As mentioned above, although *Brucellae* are relatively  
5 fastidious in terms of growth in the laboratory, they actually display a limited number  
6 of nutrient auxotrophies. All three species are auxotrophic for nicotinic acid.  
7 Sequence data confirm the three species have an incomplete pathway for the  
8 metabolism of nicotinamide, which is metabolized through the action of *pncA* (EC  
9 3.5.1.19) into nicotinic acid. Since in the *Brucella*, the *de novo* synthesis of  
10 quinolinate and NAD are absent, the only precursor for the synthesis of NAD in these  
11 organisms is nicotinic acid (23). Sequence analysis confirms that all three species lack  
12 the genes for L-aspartate oxidase (EC 1.4.3.16), quinolinate synthase and nicotinate-  
13 nucleotide pyrophosphorylase (EC 2.4.2.19), confirming the auxotrophy (17, 23) for  
14 nicotinic acid in these organisms.

15 **iv) Biotin synthesis.** *Brucella* species are considered biotin auxotrophs(17), however  
16 the reason for this requirement is unknown. Genome analysis reveals that the three  
17 species of *Brucella* possess all the genes required for the synthesis of biotin  
18 (BAB2\_0744-BAB2\_0748) in a highly conserved operon arrangement in ChrII. Since  
19 no hits were obtained for the *bioW*, *bioI*, *bioC* and *bioH* orthologs, we concluded that  
20 these genes are not present in the *Brucella* genome. However, the genes *bioW*, *bioI*  
21 and *bioC* are known to be dispensable and their absence should not affect their ability  
22 to synthesize biotin (15). Furthermore, the last gene of the *B. abortus* bio operon  
23 (BAB2\_0748) displays high homology to *bioZ*, a fatty acid synthase gene which, in  
24 *Mesorhizobium loti*, is involved in the synthesis of pimeloyl-CoA. Thus in *Brucella*  
25 this gene could compensate for the absence of the pimeloyl-CoA synthase gene, BioH

1 (41). Our genome comparisons failed to uncover the reason for biotin auxotrophy in  
2 the *Brucellae*.

3 It is interesting to note the high similarities of the *Brucella* and the *M. loti bio*  
4 operons and the fact that the *M. loti bio* operon is located within its symbiotic island, a  
5 region that can be lost or horizontally acquired (41). The possibility that the *Brucella*  
6 *bio* operon may have been horizontally acquired is supported by a drastic change in  
7 G+C content (32-45% compared with 57.3% for ChrII) observed at both ends  
8 flanking the *bio* operon (ChrII, bp739700-739900 and bp744920-745140).

9 **v) Phosphatidylcholine biosynthesis.** One of the outstanding characteristics of the  
10 *Brucella* cell envelope is the presence of phosphatidylcholine (PC) as one of the  
11 major membrane-forming phospholipids (32), a feature characteristic of eukaryotic  
12 membranes. Though no present in most prokaryotes, PC has also been found in a  
13 diverse array of bacteria, including photosynthetic bacteria, symbionts, and  
14 pathogenic bacteria capable of causing persistent infections, such as *Pseudomonas*  
15 *aeruginosa*, *Legionella pneumophila*, *Borrelia burgdorferi* (39). The genomes of  
16 *Brucellae* encode orthologs of the enzymes responsible for methylating  
17 phosphatidylethanolamine (PE) and for directly condensing choline with CDP-  
18 diacylglycerol into PC, phosphatidyl-N-methylethanolamine N-methyltransferase  
19 (Pmta) and phosphatidylcholine synthase (Pcs) respectively. The Pmta coding  
20 sequence is located on ChrI in all three species (BMEI2000, BR2127 and  
21 BAB1\_2131). However, a closer inspection of the sequences indicates unique amino  
22 acid substitutions spanning the conserved consensus motif (VLELGXGXG) in the S-  
23 adenosylmethionine-utilizing methyltransferases in all three species. This suggests  
24 that the methylation pathway may not be functional in this genus.

1           The Pcs coding sequence is located on ChrII of the three species (BMEII0695,  
2 BRA0572, BAB2\_0668). The Pcs consensus motif characteristic of CDP-alcohol  
3 phosphatidyltransferases (DGX<sub>2</sub>ARX<sub>12</sub>GX<sub>3</sub>DX<sub>3</sub>D) is absolutely conserved suggesting  
4 that this gene could be the one responsible for PC synthesis in the genus. This is  
5 supported by experimental data which indicates that the *Brucellae* are unable to  
6 synthesize PC *de novo* using PE as a precursor; suggesting that the pathway for PC  
7 formation is via condensation of choline provided by the host with CDP-diaclycerol  
8 (D.J. Comerci, unpublished results).

9 **vi) Electron transfer.** Genome analysis revealed that the respiratory chain of  
10 *Brucella* consists of a branched pathway. Electrons may be transferred from the  
11 quinone pool to the terminal oxidases through either the cytochrome bc<sub>1</sub> complex or  
12 the alternative route using quinol oxidase. Genes BAB1\_1557-BAB1\_1559 encode  
13 the Rieske Fe-S protein, and cytochrome *b* and *c* of the bc<sub>1</sub> complex. This complex  
14 transfers electrons from the oxidized quinol pool to the *c*-type cytochromes, which  
15 subsequently transfer them to the terminal cytochrome *c* oxidase. Indeed, the terminal  
16 branch for the bc<sub>1</sub> complex is represented by the cytochrome *c* oxidase complex aa<sub>3</sub>  
17 (BAB1\_0492-BAB1\_0497) and cbb<sub>3</sub> (BAB1\_0385-BAB1\_0392) but only the latter  
18 seems to be functional since the cytochrome *c* oxidase assembly protein CtaG is  
19 inactivated by a frameshift (BAB1\_0495/BAB1\_0496). The cbb<sub>3</sub>-type complex is  
20 encoded by the *fixNOQP* operon present in several rhizobia, species in which this is  
21 required for microaerobic respiration during root nodule symbiosis. Experimental data  
22 indicates that the cytochrome bc<sub>1</sub> pathway is dispensable for growth under vegetative  
23 or intracellular conditions suggesting that there is an alternate route for respiration  
24 (24).

1 Examination of the *B. abortus* genome indicates that it harbors genes encoding  
2 the terminal quinol oxidase *bo<sub>3</sub>* complex (BAB1\_0038-BAB1\_0042) and *bd* complex  
3 (BAB2\_0727-BAB2\_0730). However, as in the previous example, a critical gene in  
4 the pathway is missing. In this case, the *qoxB* gene encoding the subunit II of the *bo<sub>3</sub>*  
5 complex is inactivated by a frameshift (BAB1\_0039/BAB1\_0040) suggesting that  
6 only the cytochrome-D-ubiquinol oxidase *bd* complex is functional. Previous  
7 experimental data has revealed that the *bd* complex is essential for intracellular  
8 survival and virulence in *B. abortus* 2308, further suggesting that this is the main  
9 electron transfer route (14) in the *Brucellae*. In other bacteria, the cytochrome *bd*  
10 complex has high affinity for oxygen (42) and is the main cytochrome expressed  
11 under microaerobic conditions in *E. coli* and in nitrogen-fixing bacteria. It can thus be  
12 speculated that the presence of cytochromes with high oxygen affinity may represent  
13 an important adaptation of *Brucella* to its intracellular life style.

14

15 **Potential Virulence Modulators.** *Brucellae* do not possess classic virulence factors,  
16 rather, those already identified play roles facilitating cell invasion and survival by  
17 subverting the innate cellular defense mechanisms (1, 10, 40). Genetic elements and  
18 cellular systems that fall in this category include:

19 **i) Type IV secretion system.** In pathogenic bacteria, secretion systems are usually  
20 involved in pathogen-host interactions by delivering signals (including toxins) into  
21 host eukaryotic cells. The three species of *Brucella* have a complete and functional  
22 type IV secretion system, which is highly conserved at the nucleotide level,  
23 suggesting high selective pressure for its conservation (4). This key virulence  
24 determinant is 99% conserved at the level of DNA between the three species, a

1 feature that contrasts with the high variation observed in the putative Type III-  
2 secretion system and flagellar-export system region described below.

3 **ii) Flagella.** *Brucellae* are non-motile bacteria however their genomes contain a large  
4 complement of class II and class III flagellar genes (12, 28). The organization of these  
5 genes is identical in the three species, with 44 genes arranged in three loci on ChrII.  
6 Another two genes (a second copy of *motB* and the flagellum-specific muramidase,  
7 *flgJ*) are located in different regions of ChrI. Closer examination revealed that out of  
8 the 44 genes, seven are pseudogenes in *B. melitensis*, compared with four each in *B.*  
9 *suis* and in *B. abortus* (Supplementary material, Table 3). Only *fliG* (encoding the  
10 flagellar motor switch) is a pseudogene in all three *Brucella* species while the  
11 distribution of the other flagellar pseudogenes varies across the species. The inactive  
12 genes contain either single point mutations or small deletions and are essential for the  
13 activity (*fliG*, *fliM*, *motC*) or the assembly (*fliI*, *flgA*, *flhA*) of the flagellum, or are key  
14 structural proteins of this organelle (*flgF*, *flgI*) (3). Thus, the lack of motility in  
15 *Brucella* is a reflection of inactivation of key flagellar genes and the absence of  
16 chemotactic systems.

17         Recently however, it has been shown that under specific growth conditions, *B.*  
18 *melitensis* is able to build-up a sheathed flagellum-like structure that seems to be  
19 necessary for establishing a chronic infection process in mice by a yet unknown  
20 mechanism (16). Besides the putative function that the flagellum could play in  
21 virulence, there is no evidence to date for *Brucella* motility. Our comparative analysis  
22 of the flagellar clusters in *Brucellae* reveals that these species possess the potential to  
23 generate structures derived from the flagellum but that are not related to motility since  
24 most genes encoding basal body or motor components have been inactivated. These  
25 putative functional flagellum-related genes could thus be responsible for the

1 flagellum-like structure reported by Fretin, *et al.* (16). In summary, the differential  
2 distribution of flagellar pseudogenes across the three species, plus the potential  
3 expression of different flagellar-related components in the *Brucellae* could play a role  
4 in host-specificity and the virulence process in these organisms.

5 **iii) Regulatory elements.** Bacterial alternative RNA polymerase sigma factors are  
6 key global adaptive response regulators that have been shown to regulate virulence  
7 mechanisms in several bacterial pathogens including *Salmonella typhimurium* (20),  
8 *Legionella pneumophila* (2) and *Mycobacterium tuberculosis* (9). In particular,  
9 extracytoplasmic function (ECF) sigma factors, which mainly control cell envelope  
10 synthesis, secretory functions and periplasmic proteins, may play an important role in  
11 *Brucella* pathogenesis.

12 All three species of *Brucella* carry a highly conserved ECF gene in ChrI while  
13 a 211 amino acid paralog encoded in ChrII is present only in *B. abortus* and *B. suis*. In  
14 *B. melitensis*, a GGGG insertion in the 5' region of ECF gene has generated a fusion  
15 with a downstream hypothetical gene, which now putatively encodes a fusion protein  
16 of 413 amino acids. Though the significance of this difference is not known, it is  
17 interesting to note that other species of the  $\alpha$ -proteobacteria phylogenetically close to  
18 the *Brucella* have a variable number of ECFs. The free living species *Caulobacter*  
19 *crescentus* and *Agrobacterium tumefaciens* have 13 and 11 genes respectively. In  
20 contrast, *Brucella* spp. and the facultative endosymbionts *M. loti* and *S. meliloti*, have  
21 only two ECF paralogs, while the strict intracellular pathogens such as *Rickettsia* spp.  
22 lack ECF sigma factors entirely. This correlation suggests that in the alpha  
23 proteobacteria, these sigma factors may be counter-selected or simply not required as  
24 the organisms evolve toward a more host-dependent lifestyle.

1 **iv) Proteins containing autotransporter domains.** Autotransporters are members of  
2 a large family of exported proteins that encode an integral outer membrane pore that  
3 enables them to cross the outer membrane (19). Members of this family have been  
4 implicated as important virulence factors in many Gram-negative pathogens. The  
5 three *Brucella* genomes contain four genes with putative autotransporter  $\beta$ -domains  
6 at their C-terminus (Supplementary material, Table 4), however these genes encode  
7 four different proteins since they show low or no sequence similarity outside of this  
8 domain. The four autotransporters contain a number of haemagglutinin (Pfam  
9 PF05662) and heptapeptide (Pfam PF05658) motifs that are present in bacterial  
10 adhesins and invasins. Based on the analysis of the gene sequences, it is clear that  
11 many of them are likely inactive (Supplementary material, Table 4). We have found  
12 that all four autotransporters in *B. melitensis* are pseudogenes, each with a variety of  
13 different inactivating mutations. As a consequence of a frameshift *B. suis* BR0072 lost  
14 the signal peptide sequence and is thus predicted to start a few bases downstream at  
15 an alternative ATG codon giving an almost complete CDS lacking the signal peptide  
16 sequence that is present in *B. abortus* (BAB1\_0069). Thus, though BR0072 seems  
17 likely to be a pseudogene in *B. suis*, it remains to be seen if it is functional or able to  
18 be secreted into the periplasm via a sec-independent mechanism. The gene  
19 BAB1\_0069 also differs from its *B. suis* (BR0072) and *B. melitensis* (BMEI1873/4)  
20 orthologs in the number of internal repeated tandem sequences, which contain  
21 hemagglutinin motifs and are commonly found in autotransporter genes. Similar  
22 differences in repeated motifs was observed in the *B. abortus*  
23 BAB2\_0167/BAB2\_0168 autotransporter, which is much larger than its *B. suis*  
24 (BRA0173) and *B. melitensis* (BMEII1069/70, a pseudogene) counterparts, but carries  
25 a number of mutations which likely render it inactive. A third autotransporter appears

1 to be functional in *B. abortus* (BAB2\_1107) and *B. suis* (BRA1148) but a pseudogene  
2 in *B. melitensis* (BMEI0148/9). As was previously described, the fourth  
3 autotransporter carries inactivating nonsense mutations in *B. abortus*  
4 (BAB1\_2013/BAB1\_2014) and *B. melitensis* (BMEI0058) but is full length in *B. suis*  
5 (BR2013). Since almost all of the autotransporters known to date are involved in  
6 functions related to the invasion process, either by direct interaction with the host or  
7 with the extracellular matrix, it is tempting to speculate that the difference in the  
8 number of active autotransporters, and the variation within them, may play a role on  
9 the ability of each species to interact with its host and thus be an important contributor  
10 to the host-specificity displayed by this group.

11

## 12 **Concluding Remarks**

13 Twenty years ago, DNA-DNA hybridization studies revealed the high degree  
14 of homology shared by the six recognized *Brucella* species. This led the proposition  
15 that *Brucella* constitute a monospecific genus and that *B. melitensis* be used as the  
16 exclusive species, while the others should be considered only biovars of *B. melitensis*  
17 (46). However, this single-species organization does not reflect accurately the  
18 observed differences in pathogenicity, host-preference and evolutionary history and  
19 by and large, the classical scheme of six species is still preferred. The genomic  
20 comparisons presented in this work support the notion that the three main *Brucella*  
21 species have a singular evolutionary history, and that species-specific DNA sequences  
22 and pseudogene distribution might correlate with different host-preference.

23 Recently, it has been hypothesized that *B. melitensis* and *B. abortus* shared a  
24 common ancestor and became isolated about 20 millions years ago when radiation of  
25 artiodactyls occurred (30). The genomic rearrangements, species-specific DNA

1 sequences and distinct patterns of gene inactivation independently suggest that *B.*  
2 *abortus* and *B. melitensis* share the same lineage, which differs from the *B. suis*  
3 lineage that has undergone fewer genetic mutations since it diverged from the most  
4 recent common ancestor of all *Brucellae*. Indeed, both *B. melitensis* and *B. abortus*  
5 can cause abortion in cattle, sheep and goat, and seem to be more restricted in host-  
6 range whereas *B. suis* only causes abortions in swine and is the most diverse in  
7 genomic structure and host preference (31).

8         This evolutionary history is highly reminiscent of the radiation of  
9 Cetartiodactyla during the Paleocene-Eocene period (between 50-20 million years  
10 ago), where the family Suidae (pigs) is the deeper clade and the suborder  
11 Ruminantia, which include both the subfamilies Bovinae (cows and buffaloes) and  
12 Caprinae (goats and sheeps) evolved later. In this regard, the completion of the  
13 genome sequence of the marine *Brucellae* (from both pinnipeds and cetaceans) and of  
14 *B. ovis* would provide valuable information necessary to complete this hypothesis.

15         The genomic analyses in these work show that the majority of the identified  
16 virulence genes were conserved among the three species and thus, differences in  
17 pathogenicity could not be attributed to variations in gene complement, but rather  
18 may be the consequence of dissimilar gene inactivation of cell surface components  
19 and of transcriptional regulators. It is tempting to speculate that species-specific gene  
20 inactivations, such as in outer membrane proteins, auto-transporters, ABC transporters  
21 and regulatory genes, may account for cell surface variations and may be involved in  
22 environmental adaptation and nutrient scavenging. Interestingly, the most frequent  
23 number of species variations, including pseudogenes, occurs precisely within these  
24 functional categories. Accordingly, whereas the 11.7 kb *virB* operon encoding the  
25 important type IV secretion system is highly conserved in all four genomes, high

1 variability and differential gene inactivations characterizes what appears to be a  
2 remnant flagellar type III secretion system implicated in the production of a  
3 proteinaceous species with a role in the chronic form of brucellosis infections. Such  
4 dichotomy reaffirms the nature of the *virB* operon as an essential component of the  
5 virulence process in members of this group, while the high variation and  
6 heterogeneity observed in the flagellar loci would be more consistent with a putative  
7 role in host specificity.

8         Finally, a numerous losses affecting general metabolic processes in these  
9 organisms such as the elimination of a gene cluster involved in sugar nucleotide  
10 synthesis, sugar modification and polysaccharide synthesis as well as the loss of  
11 several other pathways such as the synthesis of energy and carbon storage compounds  
12 glycogen and poly-hydroxybutyrate, the vitamin biotin, choline plus the apparent  
13 selection for a functional high affinity respiratory would be consistent with adaptation  
14 of *Brucellae* to the protected, nutrient-poor, low oxygen-tension environment of its  
15 intracellular niche

## **ACKNOWLEDGEMENTS**

We thank Brian Souza and Warren Regala for their technical contributions. This work was performed under the auspices of the U.S. DOE by Lawrence Livermore National Laboratory under Contract W-7405-ENG-48 and Oak Ridge National Laboratory under Contract DE-AC05-00OR22725, the California State University Program for Education and Biotechnology, and the Agencia Nacional de Promoción Científica y Tecnológica, Argentina, PICT 01-11882 and 01-9194.

## REFERENCES

1. **Arellano-Reynoso, B., N. Lapaque, S. Salcedo, G. Briones, A. E. Ciocchini, R. Ugalde, E. Moreno, I. Moriyon, and J. P. Gorvel.** 2005. Cyclic beta-1,2-glucan is a brucella virulence factor required for intracellular survival. *Nat Immunol* **6**:618-25.
2. **Bachman, M. A., and M. S. Swanson.** 2001. RpoS co-operates with other factors to induce *Legionella pneumophila* virulence in the stationary phase. *Mol Microbiol* **40**:1201-14.
3. **Berg, H. C.** 2003. The rotary motor of bacterial flagella. *Annu Rev Biochem* **72**:19-54.
4. **Boschiroli, M. L., S. Ouahrani-Bettache, V. Foulongne, S. Michaux-Charachon, G. Bourg, A. Allardet-Servent, C. Cazevielle, J. P. Lavigne, J. P. Liautard, M. Ramuz, and D. O'Callaghan.** 2002. Type IV secretion and *Brucella* virulence. *Vet Microbiol* **90**:341-8.
5. **Briones, G., N. Inon de Iannino, M. Roset, A. Vigliocco, P. S. Paulo, and R. A. Ugalde.** 2001. *Brucella abortus* cyclic beta-1,2-glucan mutants have reduced virulence in mice and are defective in intracellular replication in HeLa cells. *Infect Immun* **69**:4528-35.
6. **Chain, P., J. Lamerdin, F. Larimer, W. Regala, V. Lao, M. Land, L. Hauser, A. Hooper, M. Klotz, J. Norton, L. Sayavedra-Soto, D. Arciero, N. Hommes, M. Whittaker, and D. Arp.** 2003. Complete genome sequence of the ammonia-oxidizing bacterium and obligate chemolithoautotroph *Nitrosomonas europaea*. *J Bacteriol* **185**:2759-73.
7. **Chain, P. S., E. Carniel, F. W. Larimer, J. Lamerdin, P. O. Stoutland, W. M. Regala, A. M. Georgescu, L. M. Vergez, M. L. Land, V. L. Motin, R. R. Brubaker, J. Fowler, J. Hinnebusch, M. Marceau, C. Medigue, M. Simonet, V. Chenal-Francisque, B. Souza, D. Dacheux, J. M. Elliott, A. Derbyshire, L. J. Hauser, and E. Garcia.** 2004. Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis*. *Proc Natl Acad Sci U S A* **101**:13826-31.
8. **Ciocchini, A. E., M. S. Roset, N. Inon de Iannino, and R. A. Ugalde.** 2004. Membrane topology analysis of cyclic glucan synthase, a virulence determinant of *Brucella abortus*. *J Bacteriol* **186**:7205-13.
9. **Collins, D. M., R. P. Kawakami, G. W. de Lisle, L. Pascopella, B. R. Bloom, and W. R. Jacobs, Jr.** 1995. Mutation of the principal sigma factor causes loss of virulence in a strain of the *Mycobacterium tuberculosis* complex. *Proc Natl Acad Sci U S A* **92**:8036-40.
10. **Comerci, D. J., M. J. Martinez-Lorenzo, R. Sieira, J. P. Gorvel, and R. A. Ugalde.** 2001. Essential role of the VirB machinery in the maturation of the *Brucella abortus*-containing vacuole. *Cell Microbiol* **3**:159-68.
11. **Corbel, M. J.** 1997. Brucellosis: an overview. *Emerg Infect Dis* **3**:213-21.
12. **DelVecchio, V. G., V. Kapatral, R. J. Redkar, G. Patra, C. Mujer, T. Los, N. Ivanova, I. Anderson, A. Bhattacharyya, A. Lykidis, G. Reznik, L. Jablonski, N. Larsen, M. D'Souza, A. Bernal, M. Mazur, E. Goltsman, E. Selkov, P. H. Elzer, S. Hagijs, D. O'Callaghan, J. J. Letesson, R. Haselkorn, N. Kyrpides, and R. Overbeek.** 2002. The genome sequence of the facultative intracellular pathogen *Brucella melitensis*. *Proc Natl Acad Sci U S A* **99**:443-8.

13. **Deng, W., V. Burland, G. Plunkett, 3rd, A. Boutin, G. F. Mayhew, P. Liss, N. T. Perna, D. J. Rose, B. Mau, S. Zhou, D. C. Schwartz, J. D. Fetherston, L. E. Lindler, R. R. Brubaker, G. V. Plano, S. C. Straley, K. A. McDonough, M. L. Nilles, J. S. Matson, F. R. Blattner, and R. D. Perry.** 2002. Genome sequence of *Yersinia pestis* KIM. *J Bacteriol* **184**:4601-11.
14. **Endley, S., D. McMurray, and T. A. Ficht.** 2001. Interruption of the *cydB* locus in *Brucella abortus* attenuates intracellular survival and virulence in the mouse model of infection. *J Bacteriol* **183**:2454-62.
15. **Entcheva, P., D. A. Phillips, and W. R. Streit.** 2002. Functional analysis of *Sinorhizobium meliloti* genes involved in biotin synthesis and transport. *Appl Environ Microbiol* **68**:2843-8.
16. **Fretin, D., A. Fauconnier, S. Kohler, S. Halling, S. Leonard, C. Nijskens, J. Ferooz, P. Lestrade, R. M. Delrue, I. Danese, J. Vandenhoute, A. Tibor, X. DeBolle, and J. J. Letesson.** 2005. The sheathed flagellum of *Brucella melitensis* is involved in persistence in a murine model of infection. *Cell Microbiol* **7**:687-98.
17. **Gerhardt, P.** 1958. The nutrition of brucellae. *Bacteriol Rev* **22**:81-98.
18. **Halling, S. M., B. D. Peterson-Burch, B. J. Bricker, R. L. Zuerner, Z. Qing, L. L. Li, V. Kapur, D. P. Alt, and S. C. Olsen.** 2005. Completion of the genome sequence of *Brucella abortus* and comparison to the highly similar genomes of *Brucella melitensis* and *Brucella suis*. *J Bacteriol* **187**:2715-26.
19. **Henderson, I. R., F. Navarro-Garcia, and J. P. Nataro.** 1998. The great escape: structure and function of the autotransporter proteins. *Trends Microbiol* **6**:370-8.
20. **Humphreys, S., A. Stevenson, A. Bacon, A. B. Weinhardt, and M. Roberts.** 1999. The alternative sigma factor, *sigmaE*, is critically important for the virulence of *Salmonella typhimurium*. *Infect Immun* **67**:1560-8.
21. **Inon de Iannino, N., G. Briones, M. Tolmasky, and R. A. Ugalde.** 1998. Molecular cloning and characterization of *cgs*, the *Brucella abortus* cyclic beta(1-2) glucan synthetase gene: genetic complementation of *Rhizobium meliloti* *ndvB* and *Agrobacterium tumefaciens* *chvB* mutants. *J Bacteriol* **180**:4392-400.
22. **Kaufmann, A. F., M. I. Meltzer, and G. P. Schmid.** 1997. The economic impact of a bioterrorist attack: are prevention and postattack intervention programs justifiable? *Emerg Infect Dis* **3**:83-94.
23. **Kim, S., D. Kurokawa, K. Watanabe, S. Makino, T. Shirahata, and M. Watarai.** 2004. *Brucella abortus* nicotinamidase (*PncA*) contributes to its intracellular replication and infectivity in mice. *FEMS Microbiol Lett* **234**:289-95.
24. **Ko, J., and G. A. Splitter.** 2000. Residual virulence of *Brucella abortus* in the absence of the cytochrome *bc(1)* complex in a murine model in vitro and in vivo. *Microb Pathog* **29**:191-200.
25. **Kohler, S., S. Michaux-Charachon, F. Porte, M. Ramuz, and J. P. Liautard.** 2003. What is the nature of the replicative niche of a stealthy bug named *Brucella*? *Trends Microbiol* **11**:215-9.
26. **Kortepeter, M. G., and G. W. Parker.** 1999. Potential biological weapons threats. *Emerg Infect Dis* **5**:523-7.
27. **Larsson, P., P. C. Oyston, P. Chain, M. C. Chu, M. Duffield, H. H. Fuxelius, E. Garcia, G. Halltorp, D. Johansson, K. E. Isherwood, P. D.**

- Karp, E. Larsson, Y. Liu, S. Michell, J. Prior, R. Prior, S. Malfatti, A. Sjostedt, K. Svensson, N. Thompson, L. Vergez, J. K. Wagg, B. W. Wren, L. E. Lindler, S. G. Andersson, M. Forsman, and R. W. Titball. 2005. The complete genome sequence of *Francisella tularensis*, the causative agent of tularemia. *Nat Genet* **37**:153-9.
28. Letesson, J. J., P. Lestrade, R. M. Delrue, I. Danese, F. Bellefontaine, D. Fretin, B. Taminiau, A. Tibor, A. Dricot, C. Deschamps, V. Haine, S. Leonard, T. Laurent, P. Mertens, J. Vandenhautte, and X. De Bolle. 2002. Fun stories about *Brucella*: the "furtive nasty bug". *Vet Microbiol* **90**:317-28.
29. 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**:3244-9.
30. Moreno, E., A. Cloeckert, and I. Moriyon. 2002. *Brucella* evolution and taxonomy. *Vet Microbiol* **90**:209-27.
31. Moreno, E., and I. Moriyon. 2001. Genus *Brucella*. In Dworkin (ed.), *The procaryotes: an evolving microbiological resource for the microbiological community*. Springer, New York.
32. Moriyon, I., and I. Lopez-Goni. 1998. Structure and properties of the outer membranes of *Brucella abortus* and *Brucella melitensis*. *Int Microbiol* **1**:19-26.
33. Parkhill, J., M. Sebahia, A. Preston, L. D. Murphy, N. Thomson, D. E. Harris, M. T. Holden, C. M. Churcher, S. D. Bentley, K. L. Mungall, A. M. Cerdeno-Tarraga, L. Temple, K. James, B. Harris, M. A. Quail, M. Achtman, R. Atkin, S. Baker, D. Basham, N. Bason, I. Cherevach, T. Chillingworth, M. Collins, A. Cronin, P. Davis, J. Doggett, T. Feltwell, A. Goble, N. Hamlin, H. Hauser, S. Holroyd, K. Jagels, S. Leather, S. Moule, H. Norberczak, S. O'Neil, D. Ormond, C. Price, E. Rabinowitsch, S. Rutter, M. Sanders, D. Saunders, K. Seeger, S. Sharp, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, L. Unwin, S. Whitehead, B. G. Barrell, and D. J. Maskell. 2003. Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. *Nat Genet* **35**:32-40.
34. Parkhill, J., B. W. Wren, N. R. Thomson, R. W. Titball, M. T. Holden, M. B. Prentice, M. Sebahia, K. D. James, C. Churcher, K. L. Mungall, S. Baker, D. Basham, S. D. Bentley, K. Brooks, A. M. Cerdeno-Tarraga, T. Chillingworth, A. Cronin, R. M. Davies, P. Davis, G. Dougan, T. Feltwell, N. Hamlin, S. Holroyd, K. Jagels, A. V. Karlyshev, S. Leather, S. Moule, P. C. Oyston, M. Quail, K. Rutherford, M. Simmonds, J. Skelton, K. Stevens, S. Whitehead, and B. G. Barrell. 2001. Genome sequence of *Yersinia pestis*, the causative agent of plague. *Nature* **413**:523-7.
35. Paulsen, I. T., R. Seshadri, K. E. Nelson, J. A. Eisen, J. F. Heidelberg, T. D. Read, R. J. Dodson, L. Umayam, L. M. Brinkac, M. J. Beanan, S. C. Daugherty, R. T. Deboy, A. S. Durkin, J. F. Kolonay, R. Madupu, W. C. Nelson, B. Ayodeji, M. Kraul, J. Shetty, J. Malek, S. E. Van Aken, S. Riedmuller, H. Tettelin, S. R. Gill, O. White, S. L. Salzberg, D. L. Hoover, L. E. Lindler, S. M. Halling, S. M. Boyle, and C. M. Fraser. 2002. The *Brucella suis* genome reveals fundamental similarities between animal and plant pathogens and symbionts. *Proc Natl Acad Sci U S A* **99**:13148-53.

36. **Rajashekara, G., J. D. Glasner, D. A. Glover, and G. A. Splitter.** 2004. Comparative whole-genome hybridization reveals genomic islands in *Brucella* species. *J Bacteriol* **186**:5040-51.
37. **Roop, R. M., 2nd, B. H. Bellaire, M. W. Valderas, and J. A. Cardelli.** 2004. Adaptation of the *Brucellae* to their intracellular niche. *Mol Microbiol* **52**:621-30.
38. **Smith, L. D., and T. A. Ficht.** 1990. Pathogenesis of *Brucella*. *Crit Rev Microbiol* **17**:209-30.
39. **Sohlenkamp, C., I. M. Lopez-Lara, and O. Geiger.** 2003. Biosynthesis of phosphatidylcholine in bacteria. *Prog Lipid Res* **42**:115-62.
40. **Sola-Landa, A., J. Pizarro-Cerda, M. J. Grillo, E. Moreno, I. Moriyon, J. M. Blasco, J. P. Gorvel, and I. Lopez-Goni.** 1998. A two-component regulatory system playing a critical role in plant pathogens and endosymbionts is present in *Brucella abortus* and controls cell invasion and virulence. *Mol Microbiol* **29**:125-38.
41. **Sullivan, J. T., S. D. Brown, R. R. Yocum, and C. W. Ronson.** 2001. The bio operon on the acquired symbiosis island of *Mesorhizobium* sp. strain R7A includes a novel gene involved in pimeloyl-CoA synthesis. *Microbiology* **147**:1315-22.
42. **Thony-Meyer, L.** 1997. Biogenesis of respiratory cytochromes in bacteria. *Microbiol Mol Biol Rev* **61**:337-76.
43. **Ugalde, J. E., C. Czibener, M. F. Feldman, and R. A. Ugalde.** 2000. Identification and characterization of the *Brucella abortus* phosphoglucomutase gene: role of lipopolysaccharide in virulence and intracellular multiplication. *Infect Immun* **68**:5716-23.
44. **Ugalde, J. E., V. Lepek, A. Uttaro, J. Estrella, A. Iglesias, and R. A. Ugalde.** 1998. Gene organization and transcription analysis of the *Agrobacterium tumefaciens* glycogen (glg) operon: two transcripts for the single phosphoglucomutase gene. *J Bacteriol* **180**:6557-64.
45. **Ugalde, R. A.** 1999. Intracellular lifestyle of *Brucella* spp. Common genes with other animal pathogens, plant pathogens, and endosymbionts. *Microbes Infect* **1**:1211-9.
46. **Verger, J. M., F. Grimont, P. A. Grimont, and M. Grayon.** 1985. *Brucella*, a monospecific genus as shown by deoxyribonucleic acid hybridization. *Int J Syst Bacteriol* **35**:292-5.

TABLE 1. General features of *B. abortus* 2308 and comparison with *B. suis* and *B. melitensis*

Property	<i>B. abortus</i> 2308		<i>B. suis</i> 1330		<i>B. melitensis</i> 16M	
	ChrI	ChrII	ChrI	ChrII	ChrI	ChrII
Size in bp	2,121,359	1,156,950	2,107,792	1,207,381	2,117,188	1,177,787
G+C	57.16	57.34	57.21	57.32	57.16	57.34
ORFs, total <sup>1</sup>	2,186	1,164	2,185	1,203	2,059	1,138
RNA operons	2	1	2	1	2	1
tRNAs	41	14	41	14	40	14
IS711	5	2 (+1 partial)	5	2	5	2
Pseudogenes	120	87	43	39	62	90

<sup>1</sup> As reported in Paulsen *et al.* (35) and DelVecchio *et al.* (12)

FIG. 1. Circular representations of the two *B. abortus* 2308 chromosomes. Outer rings 1 and 2 represent annotated genes, rings 3 and 4 represent the locations of *B. abortus* 2308 pseudogenes, rings 5 and 6 represent orthologs in *B. melitensis* with 100% identity at the amino acid level, rings 7 and 8 represent orthologs in *B. suis* with 100% identity at the amino acid level, ring 9 displays G+C content and ring 10 displays GC skew.

FIG. 2. Venn diagram displaying the distribution of pseudogenes among the *Brucella* genomes. Total number of pseudogenes for each section are shown in blue; distribution by chromosome is indicated (ChrI/ChrII). The total number within each species genome is shown outside the circles and under the species names. Note: these numbers do not reflect genes that are absent from the genomes.

FIG. 3. Distribution of pseudogenes by functional category. Functional classifications are as follows: J, translation ribosomal structure and biogenesis, K, transcription, L, DNA replication, recombination and repair, D, cell division, V, defense mechanisms; T, signal transduction; M, cell envelope, biogenesis, outer membrane ; N, cell motility and secretion; U, intracellular traffic, secretion, and vesicular transport; O, post-translational modification, protein turnover, and chaperons; C, energy production and conversion; G, carbohydrates metabolism, E, amino acid metabolism; F, nucleotide metabolism; H, coenzyme metabolism; I, lipid metabolism; Q, secondary metabolites biosynthesis and catabolism; R, general function prediction only; S, function unknown; X, no COG; TR, transport.

## SUPPLEMENTARY MATERIAL

### SUPPL. TABLES

Table 1: Insertions and deletions found in the *Brucellae*

Table 2: Pseudogenes found in the *Brucellae* genomes

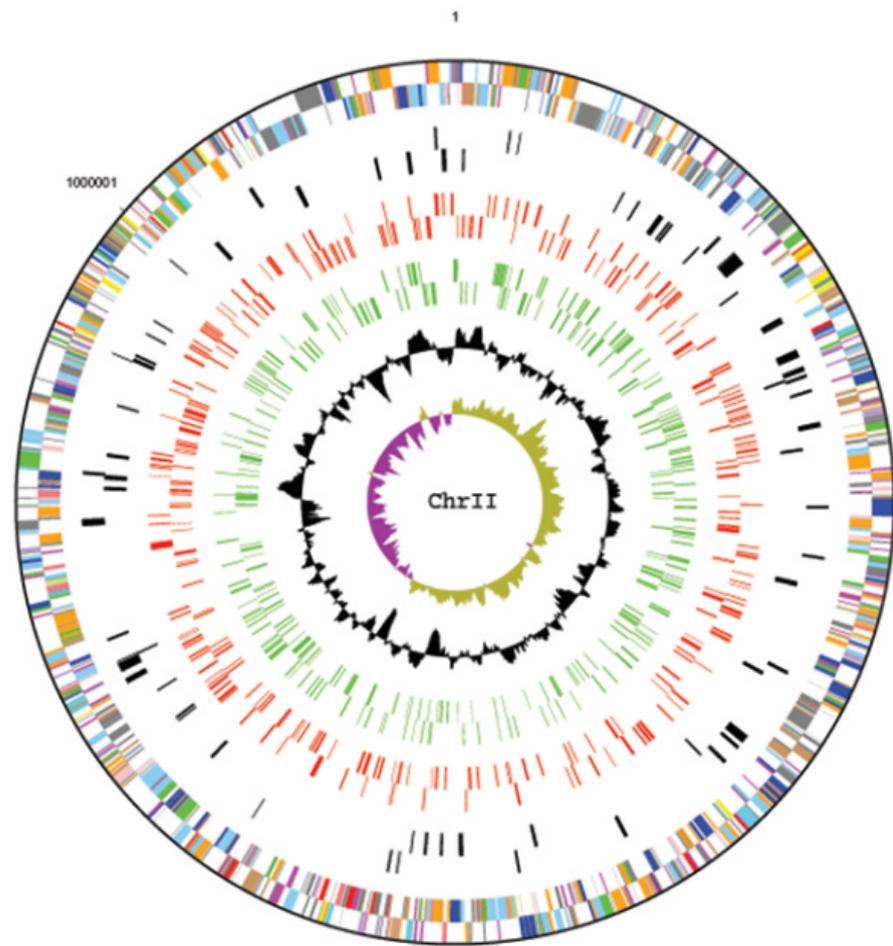
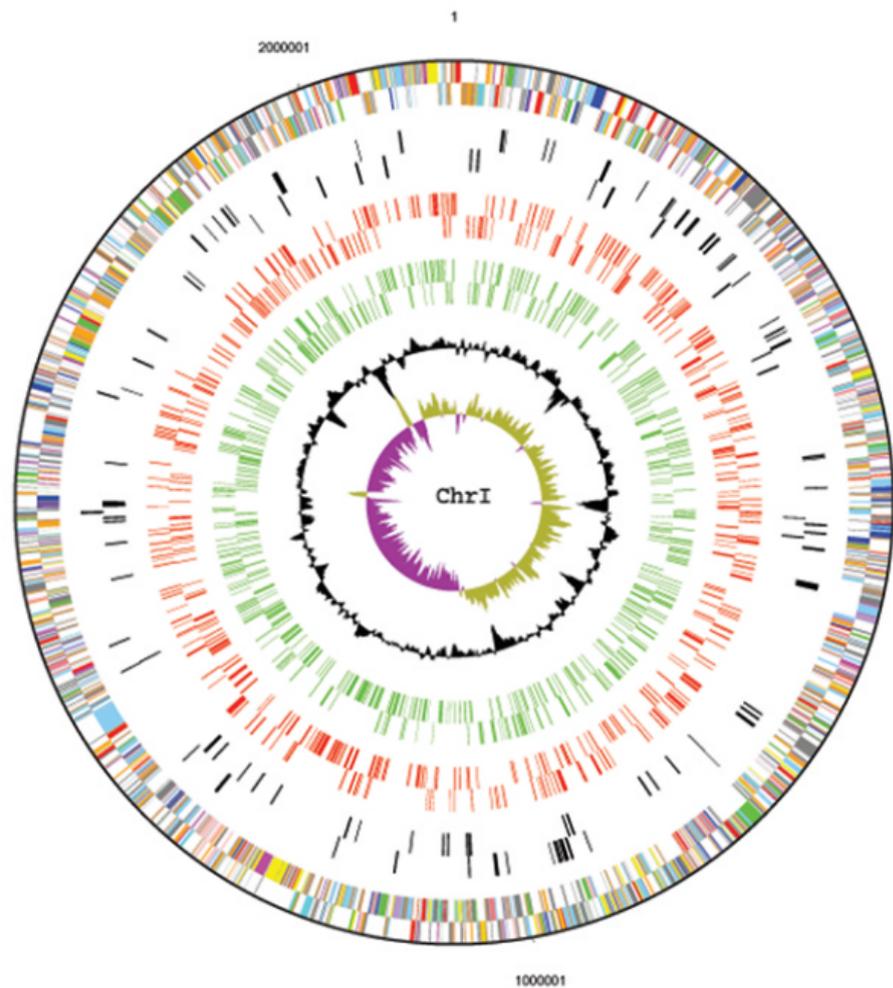
Table 3: Comparison of the *Brucellae* Flagellum Components

Table 4: Comparison of autotransporters in the three *Brucella* species

All tables are as Excel spreadsheets in the Excel Workbook entitled Suppl.Tables.xls

### SUPPL. FIGURES

FIG. 1. Alignment of a section of the genomes of *B. abortus* (top), *B. suis* (middle) and *B. melitensis* (bottom). Each genome is represented by two horizontal lines representing the DNA, flanked above and below by the six frames. Genes and their products are represented by white (gene), cyan (protein product), or red (pseudogene/product) rectangles. Between the linear genomes are blocks of similarity; blue represents locally collinear DNA, yellow represents inverted DNA.



*B. abortus*

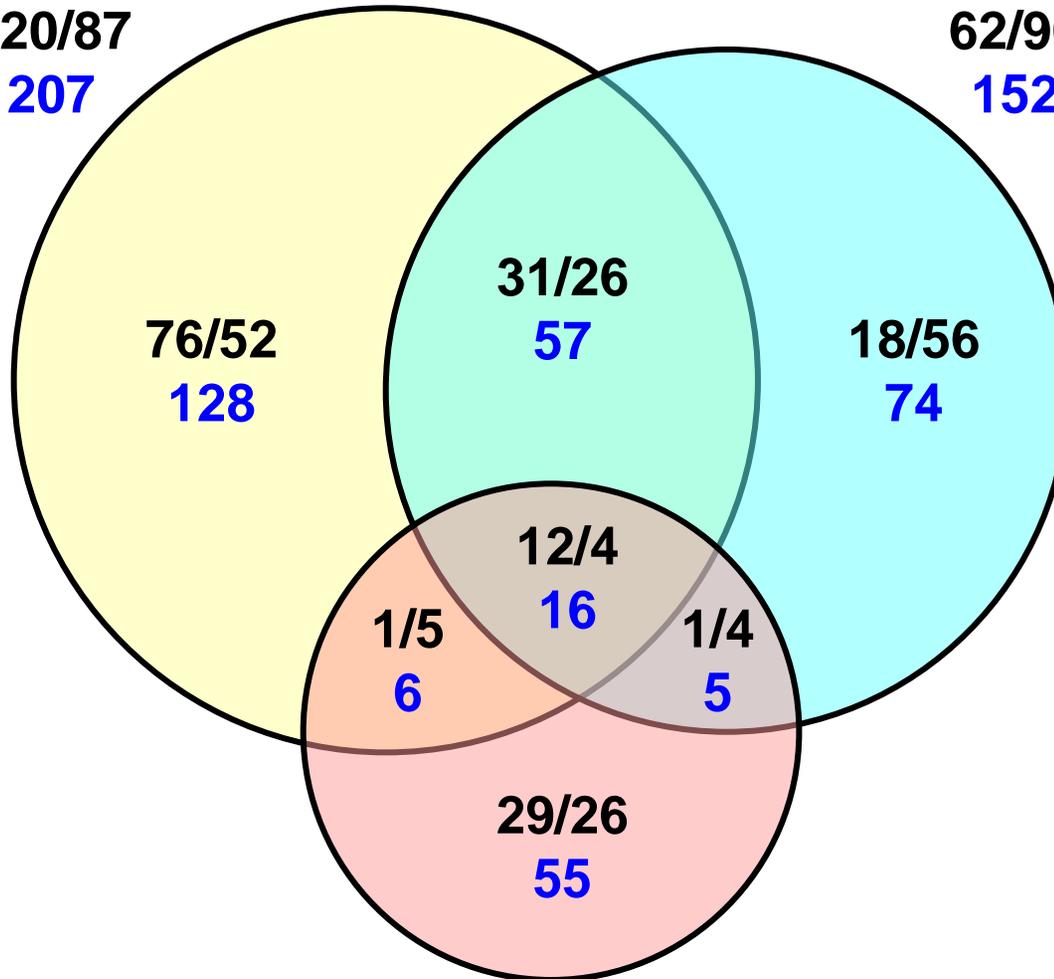
120/87

207

*B. melitensis*

62/90

152



*B. suis*

43/39

82

Genes deleted and inactivated in *Brucella*

