Supplemental Material (Supplemental Figures S1 – S3)

Core genome multilocus sequence typing scheme for improved characterization and epidemiological surveillance of pathogenic *Brucella*

Mostafa Y. Abdel-Glil ^{a,b,c}, Prasad Thomas ^d, Christian Brandt^c, Falk Melzer^a, Anbazhagan Subbaiyan^d, Pallab Chaudhuri^d, Dag Harmsen^e, Keith A. Jolley^f, Anna Janowicz^g, Giuliano Garofolo^g, Heinrich Neubauer^a, Mathias W. Pletz^c

^aInstitute of Bacterial Infections and Zoonoses, Friedrich-Loeffler-Institut, Naumburger Str. 96A, 07743 Jena, Germany.

^bFaculty of Veterinary Medicine, Zagazig University, Sharkia province, Egypt.

^cInstitute for Infectious Diseases and Infection Control, Jena University Hospital – Friedrich Schiller University, Jena, Germany.

^dDivision of Bacteriology and Mycology, ICAR-Indian Veterinary Research Institute, Izatnagar, 243122, India.

^eDepartment of Periodontology and Operative Dentistry, University Hospital Muenster, Muenster, Germany

^fDepartment of Zoology, University of Oxford, Oxford, United Kingdom.

gIstituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise "G. Caporale", Teramo, Italy

Supplemental Figures S1 – S3

Supplemental Figure S1: A maximum-likelihood (ML) phylogenetic tree computed with RAxML based on 159,572 SNP positions identified in the 612 *Brucella* genomes using the NASP pipeline. Branch coloration highlights bootstrap values as indicated in the legend

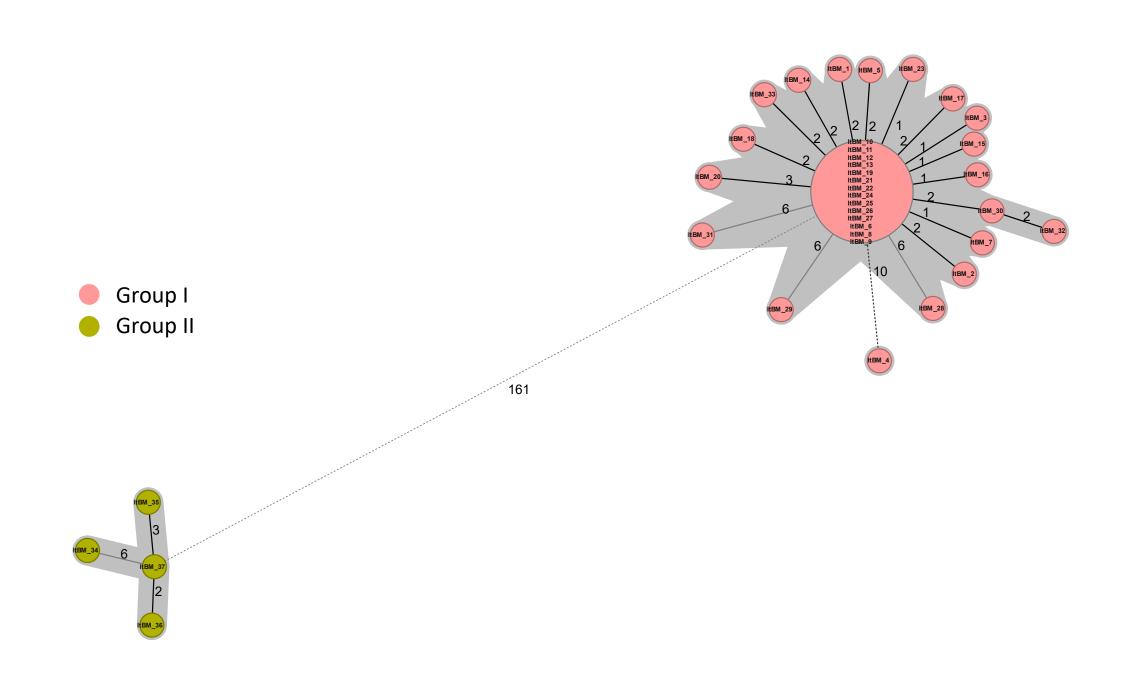
0% 100%

Brucella species

Brucella suis

Brucella species

Supplemental Figure S2: Minimum spanning tree (MST) calculated for 37 *B. melitensis* genomes with known epidemiological relation using the *B. melitensis*-specific core genome MLST targets (n= 2,704). The MST was generated using Ridom SeqSphere v07 with missing values being ignored in pairwise comparisons. Unique cgMLST profiles are represented with circles. The colors of the circles represent the genetic group of strains, while the circle size is related to the number of genomes assigned in each cgMLST genotype. The number of different alleles between cgMLST profiles is indicated on the connecting lines between the circles. The solid and dashed connecting lines represent allele differences below and above 10, respectively. Clusters were identified based on fewer than three allele mismatches and highlighted with grey shaded areas.



pplemental Figure S3: A maximum-likelihood (ML) phylogenetic tree computed with RAxML the 1,322 <i>Brucella</i> genomes. Branch coloration highlights bootstrap values as in Fig S1.

